The ghost of introduction past: Spatial and temporal variability in the genetic diversity of invasive smallmouth bass

Diedericks, G.; Henriques, Romina; von der Heyden, S.; Weyl, O.L.F.; Hui, C.

Published in:
Evolutionary Applications

Link to article, DOI:
10.1111/eva.12652

Publication date:
2018

Document Version
Peer reviewed version

Citation (APA):
Diedericks, G., Henriques, R., von der Heyden, S., Weyl, O. L. F., & Hui, C. (2018). The ghost of introduction past: Spatial and temporal variability in the genetic diversity of invasive smallmouth bass. Evolutionary Applications, 11(9), 1609-1629. https://doi.org/10.1111/eva.12652
The ghost of introduction past: Spatial and temporal variability in the genetic diversity of invasive smallmouth bass

Running headline: Genetic variability of invasive smallmouth bass

Authors: G. DiedericksI,II, R. HenriquesII, S. von der HeydenIII, O.L.F. WeylIV,V, C. HuiVI, VII

Affiliations: ICentre for Invasion Biology, Department of Botany and Zoology, Stellenbosch University, Private Bag X1, Matieland, 7602, Stellenbosch, South Africa. Email: diedericks.genevieve@gmail.com

IISection for Marine Living Resources, National Institute of Aquatic Resources, Technical University of Denmark, Kemitorvet, 2800, Kgs. Lyngby, Denmark. Email: romhe@aqua.dtu.dk

IIIEvolutionary Genomics Group, Department of Botany and Zoology, Stellenbosch University, Private Bag X1, 7602, Stellenbosch, South Africa. Email: svdh@sun.ac.za

IVDST/NRF Research Chair in Inland Fisheries and Freshwater Ecology, South African Institute for Aquatic Biodiversity (SAIAB), Grahamstown, South Africa. Email: O.Weyl@saiab.ac.za

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/eva.12652

This article is protected by copyright. All rights reserved.
ABSTRACT

Understanding the demographic history of introduced populations is essential for unravelling their invasive potential and adaptability to a novel environment. To this end, levels of genetic diversity within the native and invasive range of a species are often compared. Most studies, however, focus solely on contemporary samples, relying heavily on the premise that the historic population structure within the native range has been maintained over time. Here, we assess this assumption by conducting a three-way comparison of the genetic diversity of native (historic and contemporary) and invasive (contemporary) smallmouth bass (Micropterus dolomieu) populations. Analyses of a total of 572 M. dolomieu samples, representing the contemporary invasive South African range, contemporary and historical native USA range (dating back to the 1930s when these fish were first introduced into South Africa), revealed that the historical native range had higher genetic diversity levels when compared to both contemporary native and invasive ranges. These results suggest that both contemporary populations experienced a recent genetic bottleneck. Furthermore, the invasive range displayed significant population structure, whereas both historical and contemporary native USA populations revealed higher levels of admixture. Comparison of contemporary and historical samples showed both a historic introduction of M. dolomieu, as well as a more recent introduction,
thereby demonstrating that undocumented introductions of this species have occurred. Although multiple introductions might have contributed to the high levels of genetic diversity in the invaded range, we discuss alternative factors that may have been responsible for the elevated levels of genetic diversity and highlight the importance of incorporating historic specimens into demographic analyses.

Keywords: Genetic diversity; invasive; genetic bottleneck; demographic history; historic DNA; sampling design; multiple introductions

INTRODUCTION

Understanding the demographic history of populations constitutes a fundamental aspect of evolutionary biology. Invasive species are particularly suitable for demographic analyses, as they frequently experience rapid alternations in levels of genetic diversity following introduction (Lee, 2002; Roman & Darling, 2007; Rius & Darling, 2014; Chown et al., 2015; Hui & Richardson, 2017). To this end, the assessment of genetic diversity has become essential for establishing the demographic and adaptive potential of populations in novel environments (Prentis, Wilson, Dormott, Richardson, & Lowe, 2008; Zenni, Bailey, & Simberloff, 2014; Dlugosch, Anderson, Braasch, Cang, & Gillette, 2015; Stapley, Santure, & Dennis, 2015), and provides insight into the role that genetic variability plays in an organisms’ invasive success (Edelaar et al., 2015). Ultimately, this information allows predictions to be made on population viability, aiding in the development of an appropriate, species-specific management strategy (Prentis et al., 2008; Chown et al., 2015; Meyer et al., 2017).
Numerous studies have attempted to assess the effects of invasion dynamics on genetic variation (e.g., founder effects, genetic bottlenecks, admixture, propagule pressure; Mayr, 1963; Baker & Stebbins, 1965; Hui & Richardson, 2017) by comparing populations in the native and invasive ranges (Kolbe et al., 2004; Kelly, Muirhead, Heath, & Macisaac, 2006; Rollins, Woolnough, Wilton, Sinclair, & Sherwin, 2009; Naccarato, Dejarnette, & Allman, 2015). These types of studies aid in unravelling the demographic history of the invasive species in question (Ficetola, Bonin, & Miaud, 2008; Gillis, Walters, Fernandes, & Hoffman, 2009; Neilson & Stepien, 2011; Gray et al., 2014). Yet, despite the wealth of specimens and information housed within Natural History collections, the majority of invasion studies to date have focussed exclusively on contemporary populations, thereby relying heavily on the premise that the historic population structure within the native range has been maintained over time.

Historic DNA serves as a valuable reference when examining contemporary genetic diversity (Bouzat, 2000; Guinand, Scribner, & Page, 2003; Lozier & Cameron, 2009; Dormontt et al., 2014), as it allows for the monitoring of temporal changes in genetic diversity across generations (Guinand et al., 2003; Sefc, Payne, & Sorenson, 2007). This temporal approach increases the chance of detecting subtle changes frequently overlooked by studies focussing only on contemporary data (Lozier & Cameron, 2009), and thus allows us to delineate the most likely invasion scenario (Gillis et al., 2009; Van Kleunen, Weber, & Fischer, 2010; Thompson et al., 2011) and reveal connectivity levels among invasive populations (Funk, Garcia, Cortina, & Hill, 2011; Beneteau, Walter, Mandrak, & Heath, 2012; Snyder & Stepien, 2017). This may be of particular importance in studies conducted on taxa for which there is a priori reason to suspect temporal fluctuations in genetic variation, such as highly exploited (and subsequently stocked) taxa or species often associated with human-mediated dispersal. Hence, from an evolutionary perspective, the incorporation of historic DNA is therefore of fundamental importance.

This article is protected by copyright. All rights reserved.
Smallmouth bass, *Micropterus dolomieu* (Lacepède, 1802), presents a suitable model system to investigate variation in genetic diversity through space and time, as the species’ exploitation and subsequent stocking events within the native range are well documented (Long, Allen, Porak, & Suski, 2015), and its formal introduction history and subsequent spread into and throughout South Africa well recorded (de Moor & Bruton, 1988). Twenty-nine *M. dolomieu* specimens originating from broodstock collected in the Wheeling River, West Virginia, USA, were shipped from the Lewistown hatchery in Maryland, USA, to the Jonkershoek hatchery in South Africa in 1937 (Powell, 1967; de Moor & Burton, 1988; Loppnow, Vascotto, & Venturelli, 2013). Here, they were reared and bred before being released into multiple water bodies across the country to provide opportunities for angling (de Moor & Bruton, 1988). Most of the documented stockings (de Moor & Bruton, 1988) occurred prior to the cessation of government support to stocking programs in the early 1990s (Ellender, Woodford, Weyl, & Cowx, 2014).

Considering that both the historical record and contemporary distributions of *M. dolomieu* in South Africa are well documented, the present study aims to: (1) assess the genetic differentiation and diversity within *M. dolomieu* populations in South Africa, (2) investigate how genetic diversity changed over time in both native and invasive ranges, and (3) assess the introduction history of *M. dolomieu* into South Africa. Given the small *M. dolomieu* founding population, we predict that the invasive South African range will have a lower genetic diversity when compared to the native (historic and contemporary) North American range due to a loss of alleles, as suggested by Dlugosch and Parker (2008). Furthermore, as heavily exploited species often experience genetic bottlenecks, leaving traces in the species’ genetic diversity (Pinsky & Palumbi, 2014), we predict that the genetic diversity will be lower in contemporary time when compared to historical samples in the native range.
MATERIALS AND METHODS

DNA collection and extraction from historical specimens

Specimens representing the historical native range (Figure 1), corresponding to the approximate time of introduction into South Africa (1930 – 1941), were obtained from a host of collections housed at the Smithsonian National Museum of Natural History (NMNH), The Academy of Natural Sciences of Drexel University (ANSP), University of Michigan Museum of Zoology (UMMZ), and the Ohio State University Museum (OSUM) (Table 1; Appendix 1). In total, 53 formalin-fixed specimens representing 11 drainage systems were obtained for genetic analyses (Table 1). These specimens represent a subset of the M. dolomieu genetic diversity that was present in the native range 20 – 25 generations ago (Barthel et al., 2008).

Genomic DNA was extracted from preserved muscle tissue (20-50 mg) in a room previously unexposed to fish DNA using sterilised equipment. Prior to each extraction, all equipment and surfaces were treated with 10 % bleach to remove any potential contaminants. Pikor, Enfield, Cameron, and Lam (2011) showed that high quality DNA can be extracted from formalin-fixed tissue if the samples are rehydrated with a series of ethanol washes prior to extraction. Thus, 500 μL of 100 % ethanol was added to each tissue sample and vortexed vigorously for 30 seconds. The liquid was removed and the process was repeated with 500 μL 70 % ethanol, followed by 1000 μL distilled water. Lastly, 1000 μL distilled water was added to each sample and left to soak at 55 °C for five days, vortexing the sample every 24 hours. Once rehydrated, the sample was moved to a dry Eppendorf tube before DNA extraction, using the QIAamp DNA FFPE tissue extraction kit (QIAGEN). In a recent review, Paireder et al. (2013) demonstrated that this kit consistently outcompeted other extraction methods when working with old (1820 - 1950), formalin fixed tissue. Apart from doubling the amount of Proteinase K added to each sample (60 μL), extraction followed the manufacturers’
DNA collection and extraction from contemporary specimens

Fresh tissue samples (muscle, liver, fin clippings) were derived from specimens collected by angling in both the native United States of America (USA) and Canada and the invasive South African (SA) ranges during the summer months of 2014 and 2015 (Figure 1). Collections in North America were led by a host of individuals and organisations based in the USA and Canada (see Acknowledgements). Nine localities, rendering a total of 213 specimens were sampled from the same ‘broad’ area represented by the historical samples to allow for direct genetic diversity comparisons (Table 1). Additional specimens collected in 2014 (n = 7; formalin fixed), representing the Detroit River, were obtained from the Royal Ontario Museum (ROM), Canada.

All SA specimens were euthanised with clove oil (CapeNature permit number 0056-AAA043-00004; Eastern Cape permit numbers CRO 165/14CR and CRO 166/14CR; Mpumalanga permit number MPB.5498/2; Ethical clearance reference number SU-ACUM14-00011, University of Stellenbosch) before sampling a piece of tissue. Tissue samples were stored in 70 % ethanol for subsequent DNA extraction. Additional specimens (n = 63) were obtained from the South African Institute for Aquatic Biodiversity (SAIAB), Grahamstown, South Africa, rendering a total sample size of 306 specimens representing eight river systems (Table 1; Appendix 1). DNA was extracted from each contemporary specimen (USA & SA) using the NucleoSpin Tissue extraction (gDNA) kit (MACHERY-NAGEL,

This article is protected by copyright. All rights reserved.
Separations, Cape Town, South Africa) following the manufacturers protocol. All DNA extractions were stored at -20 °C.

**Historical and contemporary DNA amplification**

To corroborate the morphological identification of the contemporary collected specimens and assess genetic diversity and demographic history of both native and invasive populations, two partial mitochondrial DNA (mtDNA) gene regions, namely cytochrome b (cytb) and control region (CR), were amplified for all the contemporary samples (n = 519). This was not possible for the historical samples due to the limited availability of tissue and the degraded nature of the DNA. A standard 25 μL mastermix was prepared for both mtDNA polymerase chain reactions (PCR). The internal cytb primers, basscytbf1 (5´-CAC CCC TAC TTC TCC TAC AAA GA- 3´) and basscytbr1 (5´-AAG GCR AAG CGG GTG AGG G- 3´; Near, Kassler, Koppelman, Dillman, & Philipp, 2003) were used to amplify the cytb fragment. The primer set CB3R-L (5´-CATATTAAACCGAATGATATT- 3´; Palumbi, 1996) and HN20-R (5´-GTGCTTATGCTTTAGTTAAGC- 3´; Bernatchez & Danzmann, 1993) was used to amplify the CR.

Both PCR reactions followed the authors’ protocols. All PCR products were visualised through gel electrophoresis before being sequenced (ABI 3730 XL DNA Analyzer, Applied Biosystems, CAF, Stellenbosch, South Africa). Chromatographs were visually inspected and aligned in Geneious® 10.0.2 (Biomatters, Auckland, New Zealand).

Fifteen microsatellite loci, designed for both species- and genus-level amplification, were selected from published literature (Table S1). Of these, only 11 loci (eight species-specific: Mdo3, Mdo4, Mdo5, Mdo7, Mdo8, Mdo9, Mdo10, Mdo11 – Malloy, Van Den Bussche, Coughlin, & Echelle, 2000; and three genus-specific: Lma21 – Colbourne, Neff, Wright, & Gross, 1996; Lma102, Lma117 – Neff, Fu, & Gross, 1999) were successfully amplified. As Lma102 and Lma117 were not polymorphic for a

This article is protected by copyright. All rights reserved.
subset of specimens they were excluded; therefore nine polymorphic loci were used in the present study (Table S1). Three multiplex reactions were performed using the KAPA2G™ Fast Multiplex PCR Kit (KapaBiosystems, Cape Town, South Africa).

The same nine microsatellite loci were amplified for the historic samples, following the amplification procedure used for the contemporary DNA, but due to the degraded nature of the DNA this did not yield results. Thus, the resulting PCR products for each multiplex were diluted with distilled water to obtain a 1/10 PCR product which, in turn, served as template in the subsequent PCR. To ensure amplification and to avoid the overestimation of genetic diversity often associated with the amplification of ancient- and formalin-fixed DNA (Buchan, Archie, Van Horn, Moss, & Alberts, 2005; Sefc et al., 2007), historical samples were amplified twice for each microsatellite locus. All microsatellite genotyping was performed on an ABI 3730 XL DNA Analyzer (Applied Biosystems, CAF, Stellenbosch, South Africa), using LiZ as an internal size marker, and scoring was conducted in Geneious® 10.0.2 (Biomatters, Auckland, New Zealand). To ensure accurate scoring, reference individuals previously scored were used as positive controls. Historical specimens were scored blindly (i.e. specimen name removed) and repeated three times to ensure accuracy and consistency. Where scoring inconsistencies were observed (historical specimens) and more than three loci could not be scored (for both historical and contemporary specimens), the entire specimen was removed from the dataset and excluded from the study. Similarly, since one microsatellite locus, Mdo8, did not amplify for the majority of historical samples, it was removed from the historical dataset entirely. Thus, nine microsatellite loci were analysed for the contemporary dataset, but only eight microsatellite loci were analysed for the historical dataset.
Contemporary mtDNA analyses

To assess genetic diversity levels in both the contemporary native (USA - CN) and invasive (SA - CI) ranges, the number of haplotypes ($H$), haplotype diversity ($h$) and nucleotide diversity ($\pi$) were calculated for each sample site. The population history for *M. dolomieu* in both ranges was examined using Fu’s $F_s$ (Fu, 1997) and Tajima’s $D$ (Tajima, 1989). Assessment of genetic population structure was conducted combining both native and invasive contemporary datasets for each gene fragment. Pairwise $F_{ST}$ values were calculated and a hierarchical Analysis of Molecular Variance (AMOVA) conducted to determine the amount of population subdivision among sampled localities. All analyses were conducted in ARLEQUIN 3.5.2.2 (Excoffier & Lischer, 2010), with statistical significance assessed with 10,000 permutations.

Contemporary and historical microsatellite analyses

All microsatellite loci were assessed for linkage disequilibrium and deviations from Hardy-Weinberg equilibrium (HWE) in Genepop 4.2.1 (Rousset, 2008), with statistical significance assessed after 10,000 iterations. The Bonferroni method was used to correct for multiple comparisons (Rice, 1989). Amplification errors associated with large allele drop-out and stuttering was assessed with MICROCHECKER 2.2.3 (Van Oosterhout, Weetman, & Hutchinson, 2006). As most of the populations were found to not comply with HWE assumptions, FreeNA 1.2 (Chapuis & Estoup, 2007) was used to check for the presence of null alleles using the EM algorithm (Dempster, Laird, & Rubin, 1977). Intraspecific and within-population genetic diversity levels were assessed as number of alleles (Na), allelic richness (AR), observed ($H_o$) and expected heterozygosity ($H_e$), and Wright’s inbreeding coefficient ($F_{IS}$), as implemented in FSTAT 2.9.3.2 (Goudet, 1995), Genepop 4.2 (Rousset, 2008), HP-Rare 1.1 (Kalinowski, 2005) and ARLEQUIN 3.5.2.2 (Excoffier & Lischer, 2010). Statistical significance of $F_{IS}$ was assessed after 1000 permutations in FSTAT 2.9.3.2 (Goudet, 1995). Allelic richness (AR)
was calculated using HP-Rare 1.1 (Kalinowski, 2005), correcting for sample size disparity through rarefaction analysis. Analyses were conducted per population for the two contemporary datasets, but due to the small sample size for most of the historical localities (Table 1) these were grouped (= MUS) to obtain the genetic diversity indices.

Multiple approaches were employed to investigate the population structuring and genetic connectivity among (contemporary and historical) populations. As only eight loci were successfully amplified for the historical native (HN) specimens, all comparative analyses incorporating the historical samples only compared the eight loci, while contemporary SA – USA comparisons encompassed nine loci. Firstly, to determine whether there was a difference in observed heterozygosity ($H_0$) between the three groups (CI, CN, HN), an Analysis of Variance (ANOVA) was conducted in SPSS STATISTICS 20.0.0 (SPSS Inc., Chicago, IL, USA), with loci selected as random factors. Subsequently, a Bonferroni post-hoc test was used to further assess the differences between groups.

In addition, a stacked bar graph was constructed to visualise the variation among localities and loci.

Secondly, Weir’s (1986) $F_{ST}$ was employed to assess the genetic differentiation among sampled localities using FreeNA 1.2 (Chapuis & Estoup, 2007). FreeNA, employing the ENA correction method (Chapuis & Estoup, 2007), which has been shown to correctly estimate $F_{ST}$ values in the presence of null alleles (detected in the previous analysis - Chapuis & Estoup, 2007). A jackknife approach with 1000 bootstrap replicates was employed to assess statistical significance (Chapuis & Estoup, 2007).

Next, BOTTLENECK 1.2.02 (Piry, Luikart, & Cornuet, 1999) was used to test the prediction that both contemporary populations (CI and CN) experienced a recent genetic bottleneck. Populations that have undergone a genetic bottleneck are often associated with a loss of (rare) alleles and display elevated levels of heterozygosity when compared to stable populations (Piry et al., 1999). Thus, significant heterozygote excess was evaluated for each of the three groups using a Wilcoxon Rank
test (10,000 iterations) for two mutational models often associated with microsatellite evolution; the Two-Phase Mutation model (TPM) and the Infinite Alleles model (IAM).

To investigate the genetic associations within each of the three groups as well as among them, without being influenced by the lack of HWE or the presence of null alleles, a Principal Component Analysis (PCA) using microsatellite allelic frequencies was conducted in the R package Adegenet 1.3.1 (Jombart & Ahmed, 2011). Next, we used STRUCTURE 2.3.4 (Pritchard, Stephens, & Donnelly, 2000) to (a) identify and visualise the population structure within each of the three groups (CI, CN, HN), (b) compare overlapping populations from the historical and contemporary native range, and (c) search for a potential source population from where the invasive South African stocks originated. Four STRUCTURE analyses (each group independently followed by an analysis combining CI, CN and HN) were conducted using the admixture model with correlated allele frequencies, allowing each individual to be allocated to multiple clusters as determined by its genotype frequency. Five replicate runs were conducted for each $K$ ($1 < K < 15$). Runs were conducted using an initial burn-in of 75,000 Markov Chain Monte Carlo (MCMC) generations, followed by 350,000 MCMC steps. STRUCTURE HARVESTER 0.6.94 (Earl & vonHoldt, 2012) was used to determine the most probable $K$ following the Evanno method (Evanno, Regnaut, & Goudet, 2005), before using CLUMPP 1.1.2 (Jakobsson & Rosenberg, 2007) to compile the five replicate runs for the most likely $K$. DISTRUCT 1.1 (Rosenberg, 2004) was used to visualise the composite assignments.

Lastly, we performed an Approximate Bayesian Computation (ABC) on the microsatellite dataset to determine whether the invasive South African *M. dolomieu* populations originated from a single introduction event from the USA as stated by the historical records, using DIYABC 2.1.0 (Cornuet et al., 2014). As null alleles were only observed in one locus (Mdo9) of the HN dataset, all loci and
populations were included. Sampled localities were pooled into three groups (CI, CN and HN), and six simple, yet competing, introduction scenarios were generated under a coalescent framework (Figure 5: 1 - 6), in order to focus the computational efforts on probable introduction scenarios rather than an exhaustive list of possibilities (see Appendix 2 for detailed introduction scenarios). As the STRUCTURE results revealed that a subsample of the invasive South African *M. dolomieu* individuals (CI\textsubscript{S}) were more closely related to the historic native samples than to the remaining SA individuals (CI) (predominantly individuals from populations BE and OL; Figure 4: B), we simulated nine additional scenarios to test the theory of multiple introductions (Figure 5: A – I; Appendix 2). Lastly, as suggested by Guillemaud, Beaumont, Coisi, Cornuet, & Estoup (2010), three supplementary scenarios were simulated to determine whether the two SA groupings (CI and CI\textsubscript{S}) originated from (1) a single serial introduction from the source population (CN + HN), (2) two independent introduction events from the same source or (3), an unsampled source population (Figure 5: i – iii; Appendix 2). To prevent over-parameterization, parameters were specified according to the program guidelines (Cornuet et al., 2014). Firstly, we performed a pre-evaluation of the dataset to ensure that at least one scenario and its associated priors could generate simulated data sets similar to that of the observed. This was accomplished by simulating 100,000 data sets and comparing summary statistics for both simulated single-sample (i.e. mean number of alleles, genetic diversity and allele size variance across loci) and two-sample statistics (i.e. mean genetic diversity, number of alleles, allele size variance, mean index of classification, shared allele distance, distance between samples and *F*\textsubscript{ST}) to the observed data (Cornuet et al., 2014). As the mean *M* index across loci (Garza & Williamson, 2001) was initially developed with conservation planning in mind, this statistic does not perform well with small, unequal sampling sizes and small starting population sizes (Garza & Williamson, 2001). Hence, it was excluded from the summary statistics used in the current analyses. Next, we simulated $10^6$ data sets per scenario before calculating the posterior probability (PP) for each. Scenarios were subsequently compared through a logistic regression, which was conducted on the linear discriminant analysis components (Cornuet et al., 2014). Each scenarios error rate was
evaluated by generating 100 pseudo-observed data sets, using parameter values obtained from one of the scenarios (e.g. scenario 1). The type I error rate was determined by counting the number of times the PPs were higher for any scenario other than the chosen scenario, divided by the number of pseudo-observed data sets (i.e. 100), while the type II error rate was calculated by counting the number of pseudo-observed data sets that unrightfully received the highest PP support (Cornuet, Ravigne, & Estoup, 2010).

RESULTS

Contemporary mtDNA analyses

A total of 292 M. dolomieu specimens collected from eight river systems in the invasive SA range (CI) were successfully sequenced for 306bp of cytb and 979bp of CR, while the nine native USA (CN) localities yielded a total of 209 and 174 successfully sequenced M. dolomieu specimens for cytb and CR, respectively. Both cytb and CR rendered fewer haplotypes for the CN range when compared to the CI range, but similar haplotype and nucleotide diversity levels were observed (Table 2). Overall, high haplotype and low nucleotide diversity levels were observed for both native (cytb: $h = 0.976 \pm 0.005$, $\pi = 0.051 \pm 0.025$; CR: $h = 0.977 \pm 0.007$, $\pi = 0.044 \pm 0.021$) and invasive (cytb: $h = 0.967 \pm 0.007$, $\pi = 0.087 \pm 0.043$; CR: $h = 0.985 \pm 0.003$, $\pi = 0.039 \pm 0.019$) populations, but differed between sampling localities and gene fragment (Table 2). In particular, overall nucleotide diversity was higher for cytb in the CI populations (Table 2). Significant deviations from neutrality were observed for Tajima’s $D$ and Fu’s $F_S$ in both native and invasive range and both gene fragments (Table 2). Pairwise $F_{ST}$ measures revealed two significantly differentiated groupings: CI and CN (Table S2), with comparisons between localities from the two groups ranging from $F_{ST} = 0.013$ to $F_{ST} = 0.172$ (both $P < 0.05$) for cytb (DO - SAR and KO - VES) and $F_{ST} = 0.013$ to $F_{ST} = 0.125$ (both $P < 0.05$) for CR (KR - NIA and BE - LOL; Table S2). With regards to the cytb gene fragment, the CN DET population was not
significantly different from any of the CI populations, except KO. Similarly, for the CR, the CN populations ONEO and SAR was not significantly different from the majority of CI populations (Table S2). Significant within grouping differentiation (though markedly less so for the USA cyt b) was also observed in both cyt b and CR (Table S2). The AMOVA results revealed that the largest proportion of genetic variation (cyt b: 94.79 %; CR: 95.79 %) was distributed within each population, with very little variation observed between the groups (cyt b: 2.15 %; CR: 1.58 %), as well as among populations within groups (cyt b: 3.06 %; CR: 2.26 %). All variance components were significantly different from 0 (P < 0.001).

Contemporary and historical microsatellite analyses

A total of 519 contemporary sampled specimens, representing both invasive (n = 306; eight localities) and native (n = 213; nine localities) ranges, were successfully genotyped for nine microsatellite loci, while 53 museum samples, representing 11 localities within the historical native range, were successfully genotyped for eight microsatellite loci. Neither of the three groups (CI, CN, HN) displayed amplification errors (i.e. large allele dropout, stuttering), nor did any loci exhibit linkage disequilibrium. FreeNA (Chapuis & Estoup, 2007) revealed the presence of null alleles in microsatellite Mdo9 within the HN samples, but this was not the case for either of the contemporary groups. Deviations from HWE were observed in two CI populations (BE and OL) as well as the HN population (FIS: BE = 0.26, OL = 0.17, MUS = 0.43; Table S3). Further inspection revealed that this deviation was due to a heterozygote deficit within each of the three populations, suggesting the presence of a Wahlund effect (Wahlund, 1928; Waples, 2014), albeit negligible (Lye, Lepais, & Goulson, 2011; Guillemaud et al., 2015). Hence, all further analyses were conducted on the complete dataset. The number of alleles (Na) and allelic richness (AR) were consistently higher in the HN dataset, and similar between the two contemporary datasets: HN AR = 4.25, CI AR = 1.79-3.15, CN AR = 2.17-2.69 (Table S3). Multi-locus genetic diversity (observed heterozygosity, H0) ranged
from 0.39 (ONEI) to 0.59 (DET), while levels of expected heterozygosity ($H_e$) ranged from 0.35 (MP) to 0.73 (MUS) across all loci.

There was substantial variation in observed heterozygosity ($H_o$) among populations and loci, with reservoirs (catchment size < 5000 km$^2$) consistently displaying lower levels of $H_o$ (Figure 2, S2). Moreover, the ANOVA revealed significant differences in $H_o$ between the three groups ($F_{2,214} = 22.90$, $P = < 0.001$), with $H_o$ being higher for HN compared to both contemporary groups (Bonferroni post hoc test $P < 0.001$). A significant marker effect ($F_{7,214} = 19.82$, $P < 0.001$) was, however, observed. Overall $F_{ST}$ among HN samples was significantly low ($F_{ST} = 0.013$; $P < 0.05$), but this was not so for the CI ($F_{ST} = 0.211$; $P < 0.05$) and CN ($F_{ST} = 0.091$; $P < 0.05$) populations. Likewise, pairwise $F_{ST}$ values revealed significant population differentiation among CI populations, ranging from $F_{ST} = 0.066 – 0.469$ (DO – KO and BE - MP), with similar results being observed when comparing populations across all three groups, i.e. CI, CN and HN ($F_{ST} = 0.123 – 0.537$; MP – SAR and OL – MUS; Table S4). In contrast, CN populations displayed significantly less population differentiation among sampled localities within this group ($F_{ST} = 0.072 – 0.129$; LOL – NIA and SAR – STL; Table S4). As predicted, the Wilcoxon Rank test revealed a significant excess of heterozygotes for both CI and CN under the IAM model ($P = 0.002$ and $P = 0.010$, respectively), but this was not the case under the TPM model (CI: $P = 0.230$; CN: $P = 0.473$). Similarly, no significant excess of heterozygotes was detected for the HN population (IAM: $P = 0.473$; TPM: $P = 0.998$).

The principal component analysis (PCA), based on allelic frequencies, revealed two distinct groups along the first two axes; the first comprising both CN and CI populations and the second comprising the HN populations (Figure 3). Limited genetic associations between the two groups were observed. The Bayesian clustering analyses conducted in STRUCTURE revealed population sub-structuring...
within the CI localities, with Delta K (Evanno et al., 2005) retrieving $K = 5$ as the most probable number of clusters (Figure 4A). Both CI reservoirs (BU and MP) were represented by their own cluster and showed very little population variation, corroborating the genetic diversity results (Figure 2; Table S3). The remaining six CI populations, however, displayed substantial levels of admixture, in particular localities BE and OL (Figure 4A). The CN populations exhibited high levels of population admixture indicative of shallow population differentiation, with Delta K revealing the most probable $K = 4$ (Figure 4A). Similar levels of admixture and Delta K ($K = 4$) were obtained for the HN populations (Figure 4A). To determine the most probable source population of the CI populations, all 28 localities were combined (Figure 4B). Delta K revealed the most probable number of clusters to be $K = 3$, with each cluster representing a group, though admixture between the two contemporary groups was observed. Interestingly, a subset of individuals within the CI localities BE and OL (and to a lesser extent DO and KO) shared a cluster with HN, but this was not the case for any of the CN populations, despite overlapping sampling localities (DET, HUD, Susquehanna River: LOL, ONEO, VES, SU; Table 1; Figure 4B).

The ABC analysis consistently supported the notion of a more recent introduction. The first set of scenarios tested (Scenarios 1 – 6; Figure 5), revealed that Scenario 2 had the highest posterior probability (Table S5). The second set of analyses (Scenario A - I; Figure S1) supported both Scenarios C and F (Table S5). The third set of simulations (Scenarios i – iii; Figure S1), where we tested for a single versus multiple introductions from a single source or an unsampled source population, was inconclusive. Scenario iii (unsampled source population) did, however, marginally receive the most support (Table S5). Type I and Type II error rates were marginally low for the first two sets of simulations conducted (Table S5), but not for the third simulation (Table S5).
DISCUSSION

Numerous studies have compared genetic diversity levels across native and invasive ranges in an attempt to reconstruct the invasion history of invasive species (reviewed in Lee, Patel, Conlan, Wainwright, & Hipkin, 2004; Novak & Mack, 2005; Roman & Darling, 2007; Dlugosh & Parker, 2008; Rius & Darling, 2014), yet most of these studies only utilise contemporary genetic specimens. This, however, does not account for allele frequency shifts and assumes that the contemporary population structure within the native range would correspond to that of the historically native population. Using *M. dolomieu* as a study organism and incorporating both historical and contemporary native and invaded range samples, our results reveal that genetic diversity and population dynamics can indeed differ across both spatial and temporal scales.

*Genetic diversity through space and time*

Elevated levels of genetic diversity were observed in the contemporary invasive (CI) range when compared to the contemporary native (CN) range, contradicting the general assumption that genetic diversity is lower in recently invaded ranges than in long-established native populations. However, when comparing all three groups, the historical native (HN) range displayed the highest levels of heterozygosity, number of alleles (Na) and allelic richness (AR). Although this might have resulted from a statistical artefact due to the smaller sample size for the HN range, similar findings were previously reported for Atlantic salmon (*Salmo salar*; Nielsen, Hansen, & Loeschcke, 1997). The authors observed a significant decrease in Na for the contemporary population when compared to samples collected 60 years before, likely due to a recent genetic bottleneck. Our results support this proposition, as the CN population exhibited high haplotype, but low nucleotide genetic diversity, as well as significantly negative Tajima’s *D* and Fu’s *Fs* levels, all of which are commonly observed in a population that had undergone a genetic bottleneck before experiencing expansion (Grant & Bowen,
Moreover, the lack of population structure in the CN range, as well as low AR and Na, further support this notion. Strong and sustained declines in population size, such as the ones experienced by commercially exploited species, are known to leave signatures in the genetic diversity of species, in particular by reducing Na and AR (Pinsky & Palumbi, 2014). Thus, the observed contemporary population dynamics of *M. dolomieu* in its native range might have resulted from the interaction between overfishing and re-stocking events during the last two centuries (Long et al., 2015).

*Micropterus dolomieu* has been harvested both commercially and recreationally since the 1800s and has experienced several population declines and even extirpations in some localities (Marsh, 1867). This led the USA government to start breeding programmes and enforce stricter regulations on fishing in the 1870s (Long et al., 2015). In 1903 alone, ~500,000 juvenile black bass were released into waterbodies across the USA (Bowers, 1905; Loppnow et al., 2013; Long et al., 2015).

Concomitant fluctuations in population sizes are likely to have left genomic signatures and contributed to the observed elevated admixture in CN populations, as re-introductions were conducted without consideration for genetic population structure in *M. dolomieu*. Similar findings have been reported for another exploited freshwater species, the brook charr (*Silvanus fontinalis*), with individual admixture levels increasing with stocking intensity (Marie, Bernatchez, & Garant, 2010; Lamaze, Sauvage, Marie, Garant, & Bernatchez, 2012).

*Population sub-structuring in an invaded range*

Elevated levels of genetic diversity are, however, not uncommon in invasive species in a novel invaded range and are often attributed to multiple introductions and/or population mixture (see Rius & Darling, 2014 for a comprehensive review). The results from the STRUCTURE analyses appear to contradict the historical records stating that invasive South African *M. dolomieu* populations originate from a single introductory event from the USA in 1937. A genetic cluster encompassing samples from the Berg (BE: n = 14), Olifants (OL: n = 7), Doring (DO: n = 2), and Kouga (KO: n = 1)
Rivers, suggests shared ancestry with the HN samples, but the remainder of the invasive South African populations belong to four additional clusters, hinting at the idea of multiple introductions. The ABC results support this notion, as the best-fit scenario suggested a second, more recent, introduction from North America (Scenario 2). Furthermore, when considering the invasive South African individuals associated with the HN STRUCTURE cluster as a separate South African population (Cl2), the ABC analyses supported the STRUCTURE results and suggested at least two introductions: one coinciding with the recorded historic introduction and at least one more recent introduction. Indeed, the observed admixture between CI and CN suggests that the more recent introduction also originated from the USA. Unexpectedly, no support was obtained for either scenario examining single versus multiple introductions from a single source (Scenarios i, ii), nor any scenario postulating admixture (Scenarios 4, 5, 6). This may be due to several factors, such as the unequal sample sizes between HN and CI/ CN range, the simplicity of the ABC models, or perhaps it could be ascribed to the fact that the HN population was not in HWE. Furthermore, the presence of a temporal Wahlund effect within the HN range, likely due to the pooling of multiple populations, may too have decreased the accuracy of the ABC results. Although our results support the notion of multiple introductions, this should be interpreted with caution as several factors may be responsible for this pattern, including an unsampled source population, post-invasion genetic drift, insufficient marker resolution and admixture in the source population (Chown et al., 2015; Gray et al., 2014).

Given that hatcheries make use of artificial selection techniques to enhance species production and abundance (e.g. Aprahamian, Smith, McGinnity, McKelvey, & Taylor, 2003; Lamaze et al., 2012), it is possible that the introduced *M. dolomieu* were of admixed or hybrid origin, as has been reported for stockings of *S. fontinalis* (Sloss, Jennings, Franckowiak, & Pratt, 2008; Cooper, Miller, & Kapuscinski, 2010; Lamaze et al., 2012).

Invasive species capable of harbouring large, genetically diverse source populations are thought to make better invaders (Gaither, Bowen, & Toonen, 2013), as they are equipped with higher adaptive potential (Dlugosch, 2006; Lavergne & Molofsky, 2007; Wellband & Heath, 2017). Within the

This article is protected by copyright. All rights reserved.
invasive South African range, *M. dolomieu* experiences an array of climatic conditions with fluctuating rainfall and temperature regimes (Rutherford, Mucina, & Powrie, 2006). However, despite this, *M. dolomieu* has not only survived, but also established viable populations and spread throughout the systems into which it was introduced (Van Der Walt, Weyl, Woodford, & Radloff, 2016). Although the initial introduced individuals may have been of admixed stock, the substantial admixture observed among *M. dolomieu* populations in the invaded range may also have resulted from hybridisation post introduction (Diedericks et al., in review) as has been observed for *M. dolomieu* introductions elsewhere (Whitmore & Butler, 1982; Whitmore & Hellier, 1988; Avise et al., 1997; Pipas & Bulow, 1998; Bagley, Mayden, Roe, Holznagel, & Harris, 2011). Further, although sampling was conducted away from known angling ‘hotspots’, *M. dolomieu* are popular angling species and human-mediated, long-distance dispersal via intentional stocking cannot be excluded as a mechanism. Such long-distance (human-mediated) dispersal events are known to increase population mixing, ultimately increasing the species’ genetic diversity and hence, adaptability to the novel environment (Berthouly-Salazar et al., 2013).

**The influence of sampling design on genetic diversity**

Molecular techniques are indispensable tools in invasion biology (Muirhead et al., 2008; Blanchet, 2012), particularly for reconstructing species invasion histories and routes (Wilson, Dormont, Prentis, Lowe, & Richardson, 2009; Estoup & Guillemaud, 2010; Guillemaud, Beaumont, Ciosi, Cornuet, & Estoup, 2010; Guillemaud et al., 2015). However, sampling problems such as the number of native versus invasive populations sampled and the number of individuals sampled per population, may hinder the accuracy of the molecular markers to identify the source population (Guillemaud, Beaumont, Ciosi, Cornuet, & Estoup, 2010). To date, however, no study has looked at the effect that ‘sampling locality’ may have on each populations’ genetic composition, and hence, genetic diversity. For example, aquatic freshwater species, particularly fish, are often collected from
natural lakes or man-made reservoirs due to the ease of collection and the large number of individuals present. These specific sampling sites, however, often display much lower levels of genetic variability when compared to rivers, as suggested by our results (localities BU and MP in the invasive range). Similarly, a recent study reconstructing the invasion history of the largemouth bass, *Micropterus salmoides*, identified extremely low levels of neutral genetic diversity within invasive populations in lentic environments with limited connectivity (Hargrove, Weyl, & Austin, 2017). Their results revealed that all lentic populations had allele frequencies dominated by a single allele, but that a population sampled from Kowie Weir, located at the end of a 580 km² catchment, was more diverse, suggesting multiple introduction events or hybridisation between co-occurring *Micropterus* species (Hargrove et al., 2017). Thus, choice of sampling locality and in particular the degree of isolation, are important considerations when assessing the demographic or invasion history of a species.

**Management implications**

Understanding the introduction history of an invasive species is crucial when wanting to decide on a management strategy for the species in question (Prentis et al., 2009). Our results reveal a complex demographic history for *M. dolomieu*, both within its native USA and invasive SA range. With regard to management in the native range, our data support the management recommendations by Brewer and Orth (2015) that stocking should be guided by a range-wide analysis of genetic variation. In South Africa, eradication of *M. dolomieu* is no longer a feasible option due to the magnitude of the invasion, and the current management strategy is to prevent spread into previously unininvaded catchments by restricting stocking (see Woodford, Ivey, Jordaan, Kimberg, Zengeya & Weyl, 2017). This is a prudent strategy as the facilitation of strategies that might further increase genetic diversity, thought to assist population establishment, persistence and ultimately local adaptation to novel environments, may increase the fitness of this already highly successful invader. As our study
demonstrates the possibility of undocumented *M. dolomieu* introductions into the country, it is imperative that South Africa strictly enforces its current legislation with regard to avoiding new introductions of this already invasive species. In addition, introductions even in river systems that have already been invaded may aid in increasing the genetic fitness of these already highly successful invaders, and could facilitate further spread and exacerbate the already considerable impacts on native biota (van der Walt et al., 2016).

In conclusion, while studies comparing contemporary genetic variation among native and invasive ranges are valuable (Lozier & Cameron, 2009), incorporating historical DNA is essential for monitoring temporal changes in genetic diversity that are often overlooked in comparisons using only contemporary data (Hansen, 2002; Lozier & Cameron, 2009). Using the smallmouth bass, *M. dolomieu*, as study organism, our results corroborate the idea that genetic variation can indeed change over spatio-temporal scales. Both CI and CN range displayed high levels of admixture and limited population structuring. Although this pattern is not uncommon for invasive species that have been introduced multiple times, our results suggest that various factors may have played a role in shaping the genetic diversity of the CI range.

Our study highlights the importance of including historical DNA, however caution should be taken when working with historical specimens as the degraded nature of the DNA not only hampers the successful amplification of the specimens (Sefc et al., 2003; 2007), but also renders it susceptible to genotyping discrepancies. Despite this, we recommend that future studies attempting to infer the demographic history of invasive species should incorporate native historical samples.
ACKNOWLEDGEMENTS

The authors would like to thank Louis Bernatchez and two anonymous reviewers for valuable comments and suggestions on earlier versions of the manuscript. We would like to thank the following people and institutions for their sample contributions: USA - Erling Holm (Royal Ontario Museum), Douglas Nelson (University of Michigan Museum of Zoology), Jeff Williams (National Museum of Natural History), Mark Sabaj Pérez (Academy of Natural Sciences), Mark Kibbey (Ohio State University Museum), Jason Barnucz (Fisheries and Oceans Canada), Wil Wegman (Ontario Ministry of Natural Resources and Forestry, Aurora District), Gene Gilliland (B.A.S.S conservation director), Rich Carter (Ohio Department of Natural Resources), Randy Jackson (Department of Natural Resources, Cornell University), and Jeff Loukmas (New York State Department of Environmental Conservation); SA – Chris Broeckhoven, Riaan, Beverley & Adriaan Diedericks, Jacques Aproskie and Craig Fraser (South African Bass Angling Association). A special word of thanks to Francois Roux (Mpumalanga Parks Board) for assistance in obtaining sampling permits, anglers, and coordinating the sampling trips to the Blyde River Canyon, a world heritage site. We acknowledge use of specimens, infrastructure and equipment provided by the NRF-SAIAB Research and Collections Platforms.

This work is based on research supported partly by the Department of Science and Technology (DST) and National Research Foundation (NRF) of South Africa (Grant numbers, OLFW: 110507, 109015; CH: 89967, 109244), and the DST-NRF Centre of Excellence for Invasion Biology.

DATA ARCHIVING STATEMENT

Data available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.5jf41k5

This article is protected by copyright. All rights reserved.
REFERENCES

Aprahamian, M. W., Smith, K. M., McGinnity, P., McKelvey, S., & Taylor, J. (2003). Restocking of salmonids - opportunities and limitations. *Fisheries Research, 62*, 211-227.

Avise, J. C., Pierce, P. C., Van Den Avyle, M. J., Smith, M. H., Nelson, W. S., & Asmussen, M. A. (1997). Cytonuclear introgressive swamping and species turnover of bass after an introduction. *Journal of Heredity, 88*, 14-20.

Bagley, J. C., Mayden, R. L., Roe, K. J., Holznagel, W., & Harris, P. M. (2011). Congeneric phylogeographical sampling reveals polyphyly and novel biodiversity within black basses (Centrarchidae: *Micropterus*). *Biological Journal of the Linnean Society, 104*, 346-363.

Baker, H. G., & Stebbins, G. L. (eds.) (1965). The genetics of colonizing species. Academic Press, New York, New York.

Barthel, B. L., Cooke, S. J., Svec, J. H., Suski, C. D., Bunt, C. M., Phelan, F. J. S., ... Philipp, D. P. (2008). Divergent life histories among smallmouth bass *Micropterus dolomieu* inhabiting a connected river-lake system. *Journal of Fish Biology, 73*, 829-852.

Beneteau, C. L., Walter, R. P., Mandrak, N. E., & Heath, D. D. (2012). Range expansion by invasion: genetic characterization of invasion of the greenside darted (*Etheostoma blennioides*) at the northern edge of its distribution. *Biological Invasions, 14*, 191-201.

Bernatchez, L., & Danzmann, R. G. (1993). Congruence in control-region sequence and restriction-site variation in mitochondrial DNA of brook charr (*Salvelinus fontinalis* Mitchell). *Molecular Biology and Evolution, 10*, 1002-1002.

Berthouly-Salazar, C., Hui, C., Blackburn, T. M., Gaboriaud, C., van Rensburg, B. J., van Vuuren, B. J., ... Le Roux, J. J. (2013). Long-distance dispersal maximizes evolutionary potential during rapid geographic range expansion. *Molecular Ecology, 22*, 5793-5804.

This article is protected by copyright. All rights reserved.
Blanchet, S. (2012). The use of molecular tools in invasion biology: an emphasis on freshwater ecosystems. *Fisheries Management and Ecology, 19*, 120-132.

Bouzat, J. L. (2000). The importance of control populations for the identification and management of genetic diversity. *Genetica, 110*, 109-115.

Bowers, G. M. (1905). Report of the commissioner for the year ending June 30, 1903, U.S. Commission of Fish and Fisheries. Government Printing Office, Washington, D.C.

Brewer, S. K., & Orth, D. J. (2015). Smallmouth bass *Micropterus dolomieu* lacepede, 1802. *American Fisheries Society Symposium, 82*, 9-26.

Buchan, J. C., Archie, E. A., Van Horn, R. C., Moss, C. J., & Alberts, S. C. (2005). Locus effects and sources of error in non-invasive genotyping. *Molecular Ecology Resources, 5*, 680-683.

Chapuis, M. P., & Estoup, A. (2007). Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution, 24*, 621-631.

Chown, S. L., Hodgins, K. A., Griffin, P. C., Oakeshott, J. G., Byrne, M., & Hoffmann, A. A. (2015). Biological invasions, climate change and genomics. *Evolutionary Applications, 8*, 23-46.

Colbourne, J. K., Neff, B. D., Wright, J. M., & Gross, M. R. (1996). DNA fingerprinting of bluegill sunfish (*Lepomis macrochirus*) using (GT) n microsatellites and its potential for assessment of mating success. *Canadian Journal of Fisheries and Aquatic Sciences, 53*, 342-349.

Cooper, A. M., Miller, L. M., & Kapuscinski, A. R. (2010). Conservation of population structure and genetic diversity under captive breeding of remnant coaster brook trout (*Salvelinus fontinalis*) populations. *Conservation Genetics, 11*, 1087-1093.

Cornuet, J. M., Ravigne, V., & Estoup, A. (2010). Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). *BMC Bioinformatics, 11*, 401.

This article is protected by copyright. All rights reserved.
Cornuet, J. M., Pudlo, P., Veyssier, J., Deheu-Deheue-Garcia, A., Gautier, M., Leblois, R., ... Estoup, A. (2014). DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics*, 30, 1187-1189.

De Moor, I. J., & Bruton, M. N. (1988). Atlas of alien and translocated indigenous aquatic animals in southern Africa. South African National Scientific Programmes Report 144. Foundation for Research Development and Council for Scientific and Industrial Research, Pretoria, South Africa.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (methodological)*, 1-38.

Dlugosch, K. M. (2006). Adaptation and colonization in *Hypericum canariense*: past and present invasions of an island endemic. PhD Thesis, University of California, Santa Cruz.

Dlugosch, K. M., & Parker, I. M. (2008). Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology*, 17, 431-449.

Dlugosch, K. M., Anderson, S. R., Braasch, J., Cang, F. A., & Gillette, H. D. (2015). The devil is in the details: genetic variation in introduced populations and its contributions to invasion. *Molecular Ecology*, 24, 2095-2111.

Dormontt, E. E., Gardner, M. G., Breed, M. F., Rodger, J. G., Prentis, P. J., & Lowe, A. J. (2014). Genetic bottlenecks in time and space: reconstructing invasions from contemporary and historical collections. *PloS One*, 9, 106874.

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4, 359-361.

Edelaar, P., Roques, S., Hobson, E. A., Gonçalves da Silva, A., Avery, M. L., Russello, M. A., ... Tella, J. L. (2015). Shared genetic diversity across the global invasive range of the monk parakeet suggests a
common restricted geographic origin and the possibility of convergent selection. *Molecular Ecology*, 24, 2164-2176.

Ellender, B. R., Woodford, D. J., Weyl, O. L. F., & Cowx, I. G. (2014). Managing conflicts arising from fisheries enhancements based on non-native fishes in southern Africa. *Journal of Fish Biology*, 85, 1890-1906.

Estoup, A., & Guillemaud, T. (2010). Reconstructing routes of invasion using genetic data: why, how and so what? *Molecular Ecology*, 19, 4113-4130.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, 14, 2611-2620.

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

Ficetola, G. F., Bonin, A., & Miaud, C. (2008). Population genetics reveals origin and number of founders in a biological invasion. *Molecular Ecology*, 17, 773-782.

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

Funk, W. C., Garcia, T. S., Cortina, G. A., & Hill, R. H. (2011). Population genetics of introduced bullfrogs, *Rana (Lithobates) catesbeianus*, in the Willamette Valley, Oregon, USA. *Biological Invasions*, 13, 651-658.

Gaither, M. R., Bowen, B. W., & Toonen, R. J. (2013). Population structure in the native range predicts the spread of introduced marine species. *Proceedings of the Royal Society B: Biological Sciences*, 280, 20130409.

Garza, J. C., & Williamson, E. G. (2001). Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology*, 10, 305-318.
Gillis, N. K., Walters, L. J., Fernandes, F. C., & Hoffman, E. A. (2009). Higher genetic diversity in introduced than in native populations of the mussel Mytella charruana: evidence of population admixture at introduction sites. Diversity and Distributions, 15, 784-795.

Goudet, J. (1995). FSTAT (Version 1.2): a computer program to calculate F-statistics. Journal of Heredity, 86, 485-486.

Grant, W. A. S., & Bowen, B. W. (1998). Shallow population histories in deep evolutionary lineages of marine fishes: insights from sardines and anchovies and lessons for conservation. Journal of Heredity, 89, 415-426.

Gray, M. M., Wegmann, D., Haasl, R. J., White, M. A., Gabriel, S. I., Searle, J. B., ... & Payseur, B. A. (2014). Demographic history of a recent invasion of house mice on the isolated Island of Gough. Molecular Ecology, 23, 1923-1939.

Guillemaud, T., Beaumont, M. A., Ciosi, M., Cornuet, J. M., & Estoup, A. (2010). Inferring introduction routes of invasive species using approximate Bayesian computation on microsatellite data. Heredity, 104, 88-99.

Guillemaud, T., Blin, A., Le Goff, I., Desneux, N., Reyes, M., Tabone, E., ... & Lombaert, E. (2015). The tomato borer, Tuta absoluta, invading the Mediterranean Basin, originates from a single introduction from Central Chile. Scientific Reports, 5, 8371-8376.

Guinand, B., Scribner, K. T., & Page, K. S. (2003). Genetic variation over space and time: analyses of extinct and remnant lake trout populations in the Upper Great Lakes. Proceedings of the Royal Society of London B: Biological Sciences, 270, 425-433.

Hansen, M. M. (2002). Estimating the long-term effects of stocking domesticated trout into wild brown trout (Salmo trutta) populations: An approach using microsatellite DNA analysis of historical and contemporary samples. Molecular Ecology, 11, 1003-1015.

This article is protected by copyright. All rights reserved.
Hargrove, J. S., Weyl, O. L., & Austin, J. D. (2017). Reconstructing the introduction history of an invasive fish predator in South Africa. *Biological Invasions, 19*, 2261-2276.

Hui, C., & Richardson, D. M. (2017). Invasion dynamics. Oxford University Press.

Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics, 23*, 1801-1806.

Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics, 27*, 3070-3071.

Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics, 27*, 3070-3071.

Kalinowski, S. T. (2005). hp-rare 1.0: a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes, 5*, 187-189.

Kelly, D. W., Muirhead, J. R., Heath, D. D., & Macisaac, H. J. (2006). Contrasting patterns in genetic diversity following multiple invasions of fresh and brackish waters. *Molecular Ecology, 15*, 3641-3653.

Kolbe, J. J., Glor, R. E., Schettino, L. R., Lara, A. C., Larson, A., & Losos, J. B. (2004). Genetic variation increases during biological invasion by a Cuban lizard. *Nature, 431*, 177-180.

Lacepède, B. G. E. (1802). Histoire naturelle des poissons. [Natural history of fish.] Chez Plasson, Paris.

Lamaze, F. C., Sauvage, C., Marie, A., Garant, D., & Bernatchez, L. (2012). Dynamics of introgressive hybridization assessed by SNP population genomics of coding genes in stocked brook charr (*Salvelinus fontinalis*). *Molecular Ecology, 21*, 2877-2895.

Lavergne, S., & Molofsky, J. (2007). Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proceedings of the National Academy of Sciences, 104*, 3883-3888.

This article is protected by copyright. All rights reserved.
Lee, C. E. (2002). Evolutionary genetics of invasive species. *Trends in Ecology and Evolution, 17*, 386-391.

Lee, P. L., Patel, R. M., Conlan, R. S., Wainwright, S. J., & Hipkin, C. R. (2004). Comparison of genetic diversities in native and alien populations of hoary mustard (*Hirschfeldia incana* [L.] Lagreze-Fossat). *International Journal of Plant Sciences, 165*, 833-843.

Long, J., Allen, M., Porak, W., & Suski, C. D. (2015). A historical perspective of black bass management in the United States. *American Fisheries Society Symposium, 82*, 99-122.

Loppnow, G. L., Vascotto, K., & Venturelli, P. A. (2013). Invasive smallmouth bass (*Micropterus dolomieu*): history, impacts, and control. *Management of Biological Invasions, 4*, 191-206.

Lozier, J. D., & Cameron, S. A. (2009). Comparative genetic analyses of historical and contemporary collections highlight contrasting demographic histories for the bumble bees *Bombus pensylvanicus* and *B. impatiens* in Illinois. *Molecular Ecology, 18*, 1875-1886.

Lye, G. C., Lepais, O., & Goulson, D. (2011). Reconstructing demographic events from population genetic data: the introduction of bumblebees to New Zealand. *Molecular Ecology, 20*, 2888-2900.

Malloy, T. P., Van Den Bussche, Jr. R. A., Coughlin, W. D., & Echelle, A. A. (2000). Isolation and characterization of microsatellite loci in smallmouth bass, *Micropterus dolomieu* (Teleostei: Centrarchidae), and cross-specific amplification in spotted bass, *M. punctulatus*. *Molecular Ecology, 9*, 191-195.

Marie, A. D., Bernatchez, L., & Garant, D. (2010). Loss of genetic integrity correlates with stocking intensity in brook charr (*Salvelinus fontinalis*). *Molecular Ecology, 19*, 2025-2037.

Marsh, G. P. (1867). Man and nature; or, physical geography as modified by human action. Charles Scribner & Company, New York.

Mayr, E. (1963). *Animal species and evolution*. Cambridge, Mass: Belknap Press, Harvard.
Meyer, L., Causse, R., Pernin, F., Scalone, R., Bailly, G., Chauvel, B., ... Le Corre, V. (2017). New gSSR and EST-SSR markers reveal high genetic diversity in the invasive plant Ambrosia artemisiifolia L. and can be transferred to other invasive Ambrosia species. *PloS One, 12*, p.e0176197.

Muirhead, J. R., Gray, D. K., Kelly, D. W., Ellis, S. M., Heath, D. D., & Macisaac, H. J. (2008). Identifying the source of species invasions: sampling intensity vs. genetic diversity. *Molecular Ecology, 17*, 1020-1035.

Naccarato, A. M., Dejarnette, J. B., & Allman, P. (2015). Successful establishment of a non-native species after an apparent single introduction event: investigating ND4 variability in introduced black spiny-tailed iguanas (*Ctenosaura similis*) in southwestern Florida. *Journal of Herpetology, 49*, 230-236.

Near, T. J., Kassler, T. W., Koppelman, J. B., Dillman, C. B., & Philipp, D. P. (2003). Speciation in North American black basses, *Micropterus* (Actinopterygii: Centrarchidae). *Evolution, 57*, 1610-1621.

Neff, B. D., Fu, P., & Gross, M. R. (1999). Microsatellite evolution in sunfish (Centrarchidae). *Canadian Journal of Fisheries and Aquatic Sciences, 56*, 1198-1205.

Neilson, M. E., & Stepień, C. A. (2011). Historic speciation and recent colonization of Eurasian monkey gobies (*Neogobius fluviatilis* and *N. pallasi*) revealed by DNA sequences, microsatellites, and morphology. *Diversity and Distributions, 17*, 1-15.

Nielsen, E. E., Hansen, M. M., & Loeschcke, V. (1997). Analysis of microsatellite DNA from old scale samples of Atlantic salmon *Salmo salar*: a comparison of genetic composition over 60 years. *Molecular Ecology, 6*, 487-492.

Novak, S. J., & Mack, R. N. (2005). Genetic bottlenecks in alien plant species: influences of mating systems and introduction dynamics. In: *Species Invasions: Insights into Ecology, Evolution, and Biogeography*. Eds D. F. Sax, J. J. Stachowicz & S. D. Gaines. Sinauer Associates, Sunderland, Massachusetts.

This article is protected by copyright. All rights reserved.
Paireder, S., Werner, B., Bailer, J., Werther, W., Schmid, E., Patzak, B., ... Cichna-Markl, M. (2013). Comparison of protocols for DNA extraction from long-term preserved formalin fixed tissues. *Analytical Chemistry, 439*, 152-160.

Palumbi, S. R. (1996). Nucleic Acids II: The Polymerase Chain Reaction. In Molecular Systematics, 2nd ed., D.M. Hillis, C. Moritz, and B.K. Mable, eds., pp. 205–247. Sunderland, MA: Sinauer Associates, Inc.

Pikor, L. A., Enfield, K. S., Cameron, H., & Lam, W. L. (2011). DNA extraction from paraffin embedded material for genetic and epigenetic analyses. *Journal of Visualized Experiments, 49*, e2763-e2763.

Pinsky, M. L., & Palumbi, S. R. (2014). Meta-analysis reveals lower genetic diversity in overfished populations. *Molecular Ecology, 23*, 29-39.

Pipas, J. C., & Bulow, F. J. (1998). Hybridization between redeye bass and smallmouth bass in Tennessee streams. *Transactions of the American Fisheries Society, 127*, 141-146.

Piry, S., Luikart, G., & Cornuet, J. M. (1999). BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity, 90*, 502-503.

Prentis, P. J., Wilson, J. R., Dormontt, E. E., Richardson, D. M., & Lowe, A. J. (2008). Adaptive evolution in invasive species. *Trends in Plant Science, 13*, 288-294.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics, 155*, 945-959.

Rice, W. R. (1989). Analyzing tables of statistical tests. *Evolution, 43*, 223-225.

Rius, M., & Darling, J. A. (2014). How important is intraspecific genetic admixture to the success of colonising populations? *Trends in Ecology and Evolution, 29*, 233-242.

This article is protected by copyright. All rights reserved.
Rollins, L. A., Woolnough, A. P., Wilton, A. N., Sinclair, R., & Sherwin, W. B. (2009). Invasive species can't cover their tracks: using microsatellites to assist management of starling (Sturnus vulgaris) populations in Western Australia. *Molecular Ecology, 18*, 1560-1573.

Roman, J., & Darling, J. A. (2007). Paradox lost: genetic diversity and the success of aquatic invasions. *Trends in Ecology and Evolution, 22*, 454-464.

Rosenberg, N. A. (2004). DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes, 4*, 137-138.

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

Rutherford, M. C., Mucina, L., & Powrie, L. W. (2006). Biomes and bioregions of southern Africa. In *The vegetation of South Africa, Lesotho and Swaziland*, eds L. Mucina & M.C. Rutherford, pp. 30-51. South African National Biodiversity Institute, Pretoria.

Sefc, K. M., Payne, R. B., & Sorenson, M. D. (2003). Microsatellite amplification from museum feather samples: effects of fragment size and template concentration on genotyping errors. *The Auk, 120*, 982-989.

Sefc, K. M., Payne, R. B., & Sorenson, M. D. (2007). Single base errors in PCR products from avian museum specimens and their effect on estimates of historical genetic diversity. *Conservation Genetics, 8*, 879-884.

Sloss, B. L., Jennings, M. J., Franckowiak, R., & Pratt, D. M. (2008). Genetic identity of brook trout in Lake Superior south shore streams: potential for genetic monitoring of stocking and rehabilitation efforts. *Transactions of the American Fisheries Society, 137*, 1244-1251.

Snyder, M. R., & Stepien, C. A. (2017). Genetic patterns across an invasion's history: a test of change versus stasis for the Eurasian round goby in North America. *Molecular Ecology, 26*, 1075-1090.

This article is protected by copyright. All rights reserved.
Stapley, J., Santure, A. W., & Dennis, S. R. (2015). Transposable elements as agents of rapid adaptation may explain the genetic paradox of invasive species. *Molecular Ecology, 24*, 2241-2252.

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

Thompson, G. D., Robertson, M. P., Webber, B. L., Richardson, D. M., Le Roux, J. J., & Wilson, J. R. U. (2011). Predicting the subspecific identity of invasive species using distribution models: *Acacia saligna* as an example. *Diversity and Distributions, 17*, 1001-1014.

Van Der Walt, J. A., Weyl, O. L., Woodford, D. J., & Radloff, F. G. (2016). Spatial extent and consequences of black bass (Micropterus spp.) invasion in a Cape Floristic Region river basin. *Aquatic Conservation: Marine and Freshwater Ecosystems, 26*, 736-748.

Van Kleunen, M., Weber, E., & Fischer, M. (2010). A meta-analysis of trait differences between invasive and non-invasive plant species. *Ecology Letters, 13*, 235-245.

Van Oosterhout, C., Weetman, D., & Hutchinson, W. F. (2006). Estimation and adjustment of microsatellite null alleles in nonequilibrium populations. *Molecular Ecology Notes, 6*, 255-256.

Wahlund, S. (1928). Zusammensetzung von population und korrelationserscheinung vom standpunkt der vererbungslehre aus betrachtet. Hereditas. 11:65–106 [English translation. In: Weiss, K. M., & Ballonoff, P. A., editors. (1975). Demographic Genetics. Dowden, Hutchinson and Ross, Stroudsburg, 224–263.

Waples, R. S. (2014). Testing for Hardy–Weinberg proportions: have we lost the plot? *Journal of Heredity, 106*, 1-19.

Webster, M. S., & Reichart, L. (2005). Use of microsatellites for parentage and kinship analyses in animals. *Methods in Enzymology, 395*, 222-238.

This article is protected by copyright. All rights reserved.
Wellband, K. W., & Heath, D. D. (2017). Plasticity in gene transcription explains the differential performance of two invasive fish species. *Evolutionary Applications, 10*, 563-576.

Whitmore, D. H., & Butler, W. (1982). Interspecific hybridization of smallmouth and Guadalupe bass (*Micropterus*): Evidence based on biochemical genetic and morphological analyses. *The Southwestern Naturalist, 27*, 99-106.

Whitmore, D. H., & Hellier, T. R. (1988). Natural hybridization between largemouth and smallmouth bass (*Micropterus*). *Copeia*, 493-496.

Woodford, D. J., Ivey, P., Jordaan, M. S., Kimberg, P. K., Zengeya, T., & Weyl, O. L. (2017). Optimising invasive fish management in the context of invasive species legislation in South Africa. *Bothalia-African Biodiversity & Conservation, 47*, 1-9.

Zenni, R. D., Bailey, J. K., & Simberloff, D. (2014). Rapid evolution and range expansion of an invasive plant are driven by provenance–environment interactions. *Ecology Letters, 17*, 727-735.
Table 1. An overview of the sampled populations from the contemporary invasive (CI), contemporary native (CN) and historic native (HN) ranges. Abbreviations correspond to those used in subsequent tables, text and Appendix 1.

| Native/ Invasive | State/ Province | Collection Date | Sampled Locality | Drainage System | Abbr. in Tables | N | Formaldehyde Exposure | Material Supplied By | Symbol on Sampling Map (Fig. 1) |
|------------------|-----------------|-----------------|------------------|-----------------|----------------|---|---------------------|---------------------|-------------------------------|
| Native Ohio      | 1930            | White Oak Creek  | Ohio River       | OH              | 3              | yes | OSUM                | A                   |
| Native Ohio      | 1940; 1941      | Auglaize River   | Auglaize River   | AU              | 5              | yes | OSUM                | B                   |
| Native Michigan; Ontario | 1934; 1935; 1940 | Detroit River    | Detroit River    | DET             | 18             | yes | UMMZ                | C                   |
| Native Ohio      | 1941            | Lake Erie        | Lake Erie        | LE              | 3              | yes | OSUM                | D                   |
| Native Ohio      | 1938            | Mosquito Creek Lake | Mosquito Creek | MO              | 2              | yes | OSUM                | E                   |
| Native New York  | 1937            | Allegheny River  | Allegheny River  | AL              | 3              | yes | UMMZ                | F                   |
| Native New York  | 1931            | Fall Creek       | Cayuga Lake, Etna | FC              | 2              | yes | UMMZ                | G                   |
| Native New York  | 1935            | Otselic River; Susquehanna River | Susquehanna River | SU              | 5              | yes | UMMZ                | H                   |
| Native New York  | 1936            | Rondout River    | Hudson River     | HUD             | 4              | yes | UMMZ                | I                   |
| Native Maryland  | 1941            | Monocacy River   | Potomac River    | PO              | 4              | no  | ANSP                | J                   |
| Native Virginia; West Virginia | 1933 - 1936 | Shenandoah River | Shenandoah River | SH              | 4              | yes | NMNH                | K                   |

This article is protected by copyright. All rights reserved.
| Native/Invasive | Location | Year(s) | River | Collector Location | DET/River | DET/USA | ROM/Collectors |
|----------------|----------|---------|-------|--------------------|-----------|---------|----------------|
| Native         | Ontario  | 2013; 2014 | Detroit River | Detroit River | DET | 7 | yes | ROM 1 |
| Native         | New York | 2014    | Niagra River | Niagra River | NIA | 49 | no | USA collectors 2 |
| Native         | New York | 2014    | St Lawrence River | St Lawrence River | STL | 55 | no | USA collectors 3 |
| Native         | New York | 2015    | Oneida Lake | Oneida River | ONEI | 27 | no | USA collectors 4 |
| Native         | New York | 2015    | Saratoga Lake | Hudson River | SAR | 10 | no | USA collectors 5 |
| Native         | New York | 2015    | Vestal; Susquehanna River | Susquehanna River | VES | 14 | no | USA collectors 6 |
| Native         | New York | 2015    | Oneonta; Susquehanna River | Susquehanna River | ONEO | 10 | no | USA collectors 7 |
| Native         | New York | 2015    | Lolliersville | Susquehanna River | LOL | 20 | no | USA collectors 8 |
| Native         | New York | 2014    | Hudson River | Hudson River | HUD | 21 | no | USA collectors 9 |
| Invasive       | Western Cape | 2014 | Doring River | Doring River | DO | 38 | no | Self-Collected 1 |
| Invasive       | Western Cape | 2014 | Olifants River; Jan Dissels River | Olifants River | OL | 44 | no | Self-Collected 2 |
| Invasive       | Western Cape | 2014 | Berg River | Berg River | BE | 22 | no | Self-Collected 3 |

21 3
| Invasive Area | Year | Site | River | Code | Number | Collector | Code |
|---------------|------|------|-------|------|--------|-----------|------|
| Western Cape  | 2014 | Breede River | Breede River | BR | 43 | Self- Collected | 4 |
| Eastern Cape  | 2014 | Kouga River | Kouga River | KO | 46 | Self- Collected | 5 |
| Eastern Cape  | 2012 | Krom River | Krom River | KR | 15 | SAIAB | 6 |
| Eastern Cape  | 2014 | Rooikranz Dam | Buffalo River | BU | 48 | SAIAB | 7 |
| Mpumalanga    | 2014 | Blyde Dam | Blyde River | MP | 50 | MPB | 8 |
|               |      |       |       |     |       |           | 30 |
|               |      |       |       |     |       |           | 6  |
Table 2. Genetic diversity indices (haplotype ($h$) and nucleotide ($\pi$)) and neutrality tests (Tajima’s $D$ and Fu’s $F_S$) for the contemporary invasive (CI) and contemporary native (CN) ranges, based on mtDNA cytb and CR. Sample size is denoted by $n$, while $H$ refers to the number of haplotypes. Statistically significant results ($P < 0.05$) are indicated in bold.

| Contemporay Invasive SA Localities | Cytochrome b (cytb) | Control Region (CR) |
|-----------------------------------|---------------------|---------------------|
|                                   | n      | $H$    | $h$      | $\pi$   | $D$     | $Fs$       | $n$ | $H$    | $h$      | $\pi$   | $D$     | $Fs$       |
| BE                                | 20     | 16     | 0.963 ± 0.033 | 0.066 ± 0.034 | -1.682 | -1.758     | 21  | 14     | 0.867 ± 0.074 | 0.088 ± 0.044 | -2.277 | 6.160     |
| BR                                | 42     | 33     | 0.976 ± 0.014 | 0.061 ± 0.031 | -1.295 | -9.88      | 43  | 33     | 0.981 ± 0.011 | 0.036 ± 0.018 | -2.011 | -4.340     |
| BU                                | 47     | 30     | 0.965 ± 0.013 | 0.061 ± 0.031 | -2.004 | -4.574     | 47  | 33     | 0.984 ± 0.008 | 0.020 ± 0.010 | -2.594 | -10.918    |
| DO                                | 35     | 30     | 0.987 ± 0.012 | 0.263 ± 0.129 | 0.314  | -1.295     | 36  | 30     | 0.979 ± 0.016 | 0.084 ± 0.041 | -2.537 | 0.321      |
| KO                                | 46     | 24     | 0.756 ± 0.071 | 0.044 ± 0.022 | -2.310 | -2.777     | 45  | 36     | 0.984 ± 0.010 | 0.013 ± 0.007 | -1.71  | -21.924     |
| KR                                | 14     | 9      | 0.835 ± 0.101 | 0.050 ± 0.027 | -1.768 | 0.833      | 15  | 15     | 1.000 ± 0.024 | 0.046 ± 0.024 | -2.047 | -2.642     |
| MP                                | 45     | 37     | 0.987 ± 0.009 | 0.071 ± 0.036 | -0.257 | -11.881    | 45  | 31     | 0.942 ± 0.024 | 0.063 ± 0.031 | -2.646 | 0.974      |
| OL                                | 43     | 24     | 0.947 ± 0.020 | 0.033 ± 0.017 | -2.071 | -5.458     | 40  | 17     | 0.906 ± 0.029 | 0.045 ± 0.022 | -1.603 | 8.417      |
| Overall                           | 292    | 176    | 0.967 ± 0.007 | 0.087 ± 0.043 | -1.899 | -23.547    | 292 | 179    | 0.985 ± 0.003 | 0.039 ± 0.019 | -2.717 | -23.604    |
| Contemporary Native USA Localities|        |        |            |            |     |            |        |        |            |            |      |            |
| DET                               | 7      | 7      | 1.000 ± 0.076 | 0.144 ± 0.083 | 0.767  | -0.226     | -  | -      | -          | -          | -     | -          |
| HUD                               | 20     | 15     | 0.968 ± 0.025 | 0.050 ± 0.026 | -2.140 | -1.675     | 17  | 17     | 1.000 ± 0.020 | 0.134 ± 0.068 | 0.692  | -1.145     |
| LOL                               | 20     | 16     | 0.974 ± 0.025 | 0.040 ± 0.021 | -1.940 | -3.662     | 20  | 13     | 0.884 ± 0.067 | 0.001 ± 0.001 | -1.174 | -15.968    |
| NIA                               | 48     | 31     | 0.957 ± 0.018 | 0.032 ± 0.017 | -2.445 | -12.403    | 38  | 28     | 0.976 ± 0.014 | 0.011 ± 0.006 | -2.157 | -13.583    |
| ONEI                              | 30     | 26     | 0.989 ± 0.013 | 0.022 ± 0.012 | -1.545 | -20.166    | 18  | 17     | 0.994 ± 0.021 | 0.082 ± 0.042 | -2.389 | -0.867     |
|       | Mean |   Std    | Min | Max | Lower  | Upper  |
|-------|------|---------|-----|-----|--------|--------|
| ONEO  | 0.956 ± 0.059 | 0.156 ± 0.084 | -0.689 | 2.782 |
| SAR   | 0.987 ± 0.035 | 0.030 ± 0.017 | -0.615 | -4.471 |
| STL   | 0.966 ± 0.017 | 0.032 ± 0.017 | -0.829 | -18.178 |
| VES   | 0.923 ± 0.060 | 0.022 ± 0.012 | -1.950 | -2.114 |
| Overall | 0.976 ± 0.005 | 0.051 ± 0.025 | -2.191 | -23.870 |
|       | 1.000 ± 0.045 | 0.012 ± 0.007 | -1.575 | -4.188 |
|       | 1.000 ± 0.076 | 0.301 ± 0.169 | -1.806 | 2.179 |
|       | 0.942 ± 0.023 | 0.002 ± 0.001 | -1.960 | -28.464 |
|       | 0.962 ± 0.041 | 0.059 ± 0.031 | -1.418 | 2.703 |
|       | 0.977 ± 0.007 | 0.044 ± 0.021 | -1.829 | -23.756 |
Appendix 1. A detailed description of specimens obtained from various museums, including the specimen origin, collection date, specimen abbreviation corresponding to that used in Table 1, museum responsible for the specimen and its corresponding accession number.

| Country | State               | Sampled Locality               | Drainage System | Collection Date | Specimen Abbrev. | Material Supplied By | Accession #     | Notes                                                                 |
|---------|---------------------|--------------------------------|-----------------|-----------------|------------------|----------------------|------------------|------------------------------------------------------------------------|
| USA     | Maryland            | Monocacy River                 | Potomac River   | 1941            | PO_1             | ANSP                 | ANSP 95683       | fry                                                                     |
| USA     | Maryland            | Monocacy River                 | Potomac River   | 1941            | PO_2             | ANSP                 | ANSP 95683       | fry                                                                     |
| USA     | Maryland            | Monocacy River                 | Potomac River   | 1941            | PO_3             | ANSP                 | ANSP 95683       | fry                                                                     |
| USA     | Maryland            | Plummer Is., Maryland.         | Potomac River   | 1930            | PO_4             | NMNH                 | USNM 284083       | fin snip & bits of gillraker; might have been exposed to arsenic (As), mercury (Hg), lead (Pb) |
| USA     | Virginia            | Shenandoah River               | Shenandoah River| 1934            | SH_1             | NMNH                 | USNM 102132       | Muscle tissue                                                     |
| USA     | Virginia            | Shenandoah River               | Shenandoah River| 1935            | SH_2             | NMNH                 | USNM 93780        | Muscle tissue                                                     |
| USA     | West Virginia       | Shenandoah River               | Shenandoah River| 1936            | SH_3             | NMNH                 | USNM 100694       | Muscle tissue                                                     |
| USA     | Virginia            | Shenandoah River               | Shenandoah River| 1933            | SH_4             | NMNH                 | USNM 104928       | Muscle tissue                                                     |
| USA     | Ohio                | Mosquito Creek                 | Mosquito Creek  | 1938            | MO_1             | OSUM                 | OSUM 3568         | Muscle tissue                                                     |
| USA     | Ohio                | Mosquito Creek                 | Mosquito Creek  | 1938            | MO_2             | OSUM                 | OSUM 3568         | Muscle tissue                                                     |
| USA     | Ohio                | Auglaize River                 | Auglaize River  | 1940            | AU_1             | OSUM                 | OSUM 3814         | Muscle tissue                                                     |
| USA     | Ohio                | Auglaize River                 | Auglaize River  | 1940            | AU_2             | OSUM                 | OSUM 3814         | Muscle tissue                                                     |
| USA     | Ohio                | Auglaize River                 | Auglaize River  | 1940            | AU_3             | OSUM                 | OSUM 3942         | Muscle tissue                                                     |
| USA     | Ohio                | Pusheta Creek                  | Auglaize River  | 1941            | AU_4             | OSUM                 | OSUM 4343         | Muscle tissue                                                     |
USA Ohio Pusheta Creek Auglaize River 1941 AU_5 OSUM OSUM 4343 Muscle tissue
USA Ohio Lake Erie Lake Erie 1941 LE_1 OSUM OSUM 4272 Muscle tissue
USA Ohio Lake Erie Lake Erie 1941 LE_2 OSUM OSUM 4272 Muscle tissue
USA Ohio Lake Erie Lake Erie 1941 LE_3 OSUM OSUM 4272 Muscle tissue
USA Ohio White Oak Creek Ohio River 1930 OH_1 OSUM OSUM 10834 Muscle tissue
USA Ohio White Oak Creek Ohio River 1930 OH_2 OSUM OSUM 10834 Muscle tissue
USA Ohio White Oak Creek Ohio River 1930 OH_3 OSUM OSUM 10834 Muscle tissue
USA Michigan Grosse Isle shore, Detroit river Detroit River 1935 DE_1 UMMZ UMMZ 243459 Muscle tissue
USA Michigan Grosse Isle shore, Detroit river Detroit River 1935 DE_2 UMMZ UMMZ 243459 Muscle tissue
USA Michigan Grosse Isle shore, Detroit river Detroit River 1935 DE_3 UMMZ UMMZ 243459 Muscle tissue
USA Michigan Grosse Isle shore, Detroit river Detroit River 1935 DE_4 UMMZ UMMZ 243459 Muscle tissue
USA Michigan Detroit River Detroit River 1935 DE_5 UMMZ UMMZ 243226 Muscle tissue
USA Michigan Detroit River Detroit River 1935 DE_6 UMMZ UMMZ 243226 Muscle tissue
USA Michigan Detroit River Detroit River 1935 DE_7 UMMZ UMMZ 243077 Muscle tissue
USA Michigan Detroit River Detroit River 1935 DE_8 UMMZ UMMZ 243077 Muscle tissue
USA Michigan Detroit River Detroit River 1935 DE_9 UMMZ UMMZ 243077 Muscle tissue
Canada Ontario Detroit River Detroit River 1940 DE_10 UMMZ UMMZ 130878 Muscle tissue
Canada Ontario Detroit River Detroit River 1940 DE_11 UMMZ UMMZ 130878 Muscle tissue
USA Michigan Detroit River Detroit River 1934 DE_12 UMMZ UMMZ 243009 Muscle tissue
USA Michigan Detroit River Detroit River 1934 DE_13 UMMZ UMMZ 243009 Muscle tissue
| Country | State | Location 1 | Location 2 | Year | Sample Code | Repository Code | Repository | Tissue Type |
|---------|-------|------------|------------|------|-------------|----------------|-----------|-------------|
| USA     | Michigan | Detroit River | Detroit River | 1934 | DE_14 | UMMZ 243009 | Muscle tissue |
| USA     | Michigan | Detroit River | Detroit River | 1934 | DE_15 | UMMZ 243009 | Muscle tissue |
| USA     | Ontario | Detroit River | Detroit River | 1940 | DE_16 | UMMZ 130896 | Muscle tissue |
| USA     | Ontario | Detroit River | Detroit River | 1940 | DE_17 | UMMZ 130896 | Muscle tissue |
| USA     | Ontario | Detroit River | Detroit River | 1940 | DE_18 | UMMZ 130896 | Muscle tissue |
| USA     | New York | Otselic River | Susquehanna River | 1935 | SU_1 | UMMZ 109652 | Muscle tissue |
| USA     | New York | Otselic River | Susquehanna River | 1935 | SU_2 | UMMZ 109652 | Muscle tissue |
| USA     | New York | Otselic River | Susquehanna River | 1935 | SU_3 | UMMZ 109652 | Muscle tissue |
| USA     | New York | Susquehanna River | Susquehanna River | 1935 | SU_4 | UMMZ 109759 | Muscle tissue |
| USA     | New York | Susquehanna River | Susquehanna River | 1935 | SU_5 | UMMZ 109759 | Muscle tissue |
| USA     | New York | trib Rondout River to Hudson River | Hudson River | 1936 | HU_1 | UMMZ 114240 | Muscle tissue |
| USA     | New York | trib Rondout River to Hudson River | Hudson River | 1936 | HU_2 | UMMZ 114240 | Muscle tissue |
| USA     | New York | trib Rondout River to Hudson River | Hudson River | 1936 | HU_3 | UMMZ 114240 | Muscle tissue |
| USA     | New York | trib Rondout River to Hudson River | Hudson River | 1936 | HU_4 | UMMZ 114240 | Muscle tissue |
| USA     | New York | Allegheny River | Alleghany River | 1937 | AL_1 | UMMZ 180878 | Muscle tissue |
| USA     | New York | Allegheny River | Alleghany River | 1937 | AL_2 | UMMZ 180878 | Muscle tissue |
| USA     | New York | Allegheny River | Alleghany River | 1937 | AL_3 | UMMZ 180878 | Muscle tissue |

This article is protected by copyright. All rights reserved.
| Country | State Province | Location | Date | Code | Repository | Number | Type |
|---------|----------------|----------|------|------|------------|--------|------|
| USA     | New York       | Fall Creek, trib. to Cayuga Lake, Etna | 1931 | FC_1 | UMMZ       | 94455  | Muscle tissue |
| USA     | New York       | Fall Creek, trib. to Cayuga Lake, Etna | 1931 | FC_2 | UMMZ       | 94455  | Muscle tissue |
| SA      | Eastern Cape   | Elandsjacht Dam | Krom | 2012 | KR2        | SAIAB   | AC09 B425 | Muscle tissue |
| SA      | Eastern Cape   | Elandsjacht Dam | Krom | 2012 | KR3        | SAIAB   | AC09 B955 | Muscle tissue |
| SA      | Eastern Cape   | Elandsjacht Dam | Krom | 2012 | KR4        | SAIAB   | AC09 B875 | Muscle tissue |
| SA      | Eastern Cape   | Elandsjacht Dam | Krom | 2012 | KR5        | SAIAB   | AC09 B992 | Muscle tissue |
| SA      | Eastern Cape   | Elandsjacht Dam | Krom | 2012 | KR6        | SAIAB   | AC09 B994 | Muscle tissue |
| SA      | Eastern Cape   | Elandsjacht Dam | Krom | 2012 | KR7        | SAIAB   | AC09 B977 | Muscle tissue |
| SA      | Eastern Cape   | Elandsjacht Dam | Krom | 2012 | KR8        | SAIAB   | AC09 B960 | Muscle tissue |
| SA      | Eastern Cape   | Elandsjacht Dam | Krom | 2012 | KR9        | SAIAB   | AC09 B964 | Muscle tissue |
| SA      | Eastern Cape   | Elandsjacht Dam | Krom | 2012 | KR10       | SAIAB   | AC09 B982 | Muscle tissue |
| SA      | Eastern Cape   | Elandsjacht Dam | Krom | 2012 | KR11       | SAIAB   | AC09 B978 | Muscle tissue |
| SA      | Eastern Cape   | Elandsjacht Dam | Krom | 2012 | KR12       | SAIAB   | AC09 B971 | Muscle tissue |
| SA      | Eastern Cape   | Elandsjacht Dam | Krom | 2012 | KR13       | SAIAB   | AC09 B997 | Muscle tissue |
| SA      | Eastern Cape   | Elandsjacht Dam | Krom | 2012 | KR14       | SAIAB   | AC09 B970 | Muscle tissue |
| SA      | Eastern Cape   | Elandsjacht Dam | Krom | 2012 | KR15       | SAIAB   | AC09 B984 | Muscle tissue |
| SA      | Eastern Cape   | Elandsjacht Dam | Krom | 2012 | KR16       | SAIAB   | AC09 B963 | Muscle tissue |
| SA      | Eastern Cape   | Rooikranz Dam  | Buffalo River | 2014 | BU1      | SAIAB   | OW14-965 | Muscle tissue |
| SA      | Eastern Cape   | Rooikranz Dam  | Buffalo River | 2014 | BU2      | SAIAB   | OW14-985 | Muscle tissue |
| Country | Province | Location | Year | Collector | Code | Code Type | Tissue Type |
|---------|----------|----------|------|-----------|------|-----------|-------------|
| SA      | Eastern Cape | Rooikranz Dam | 2014 | BU3       | SAIAB OW14-979 | Muscle tissue |
| SA      | Eastern Cape | Rooikranz Dam | 2014 | BU4       | SAIAB OW14-941 | Muscle tissue |
| SA      | Eastern Cape | Rooikranz Dam | 2014 | BU5       | SAIAB OW14-835 | Muscle tissue |
| SA      | Eastern Cape | Rooikranz Dam | 2014 | BU6       | SAIAB OW14-828 | Muscle tissue |
| SA      | Eastern Cape | Rooikranz Dam | 2014 | BU7       | SAIAB OW14-791 | Muscle tissue |
| SA      | Eastern Cape | Rooikranz Dam | 2014 | BU8       | SAIAB OW14-700 | Muscle tissue |
| SA      | Eastern Cape | Rooikranz Dam | 2014 | BU9       | SAIAB OW14-798 | Muscle tissue |
| SA      | Eastern Cape | Rooikranz Dam | 2014 | BU10      | SAIAB OW14-688 | Muscle tissue |
| SA      | Eastern Cape | Rooikranz Dam | 2014 | BU11      | SAIAB OW14-684 | Muscle tissue |
| SA      | Eastern Cape | Rooikranz Dam | 2014 | BU12      | SAIAB OW14-808 | Muscle tissue |
| SA      | Eastern Cape | Rooikranz Dam | 2015 | BU13      | SAIAB OW14-737 | Muscle tissue |
| SA      | Eastern Cape | Rooikranz Dam | 2015 | BU14      | SAIAB OW14-735 | Muscle tissue |
| SA      | Eastern Cape | Rooikranz Dam | 2015 | BU15      | SAIAB OW14-742 | Muscle tissue |
| SA      | Eastern Cape | Rooikranz Dam | 2015 | BU16      | SAIAB OW14-724 | Muscle tissue |
| SA      | Eastern Cape | Rooikranz Dam | 2015 | BU17      | SAIAB OW14-686 | Muscle tissue |
| SA      | Eastern Cape | Rooikranz Dam | 2015 | BU18      | SAIAB OW14-797 | Muscle tissue |
| SA      | Eastern Cape | Rooikranz Dam | 2015 | BU19      | SAIAB OW14-796 | Muscle tissue |
| SA      | Eastern Cape | Rooikranz Dam | 2015 | BU20      | SAIAB OW14-675 | Muscle tissue |
| SA      | Eastern Cape | Rooikranz Dam | 2015 | BU21      | SAIAB OW14-702 | Muscle tissue |
| SA      | Eastern Cape | Rooikranz Dam | 2015 | BU22      | SAIAB OW14-744 | Muscle tissue |

This article is protected by copyright. All rights reserved.
| SA | Eastern Cape | Rooikranz Dam | Buffalo River | 2015 | BU23 | SAIAB | OW14-705 | Muscle tissue |
| SA | Eastern Cape | Rooikranz Dam | Buffalo River | 2015 | BU24 | SAIAB | OW14-782 | Muscle tissue |
| SA | Eastern Cape | Rooikranz Dam | Buffalo River | 2015 | BU25 | SAIAB | OW14-732 | Muscle tissue |
| SA | Eastern Cape | Rooikranz Dam | Buffalo River | 2015 | BU26 | SAIAB | OW14-746 | Muscle tissue |
| SA | Eastern Cape | Rooikranz Dam | Buffalo River | 2015 | BU27 | SAIAB | OW14-756 | Muscle tissue |
| SA | Eastern Cape | Rooikranz Dam | Buffalo River | 2015 | BU28 | SAIAB | OW14-738 | Muscle tissue |
| SA | Eastern Cape | Rooikranz Dam | Buffalo River | 2015 | BU29 | SAIAB | OW14-733 | Muscle tissue |
| SA | Eastern Cape | Rooikranz Dam | Buffalo River | 2015 | BU30 | SAIAB | OW14-739 | Muscle tissue |
| SA | Eastern Cape | Rooikranz Dam | Buffalo River | 2015 | BU31 | SAIAB | OW14-799 | Muscle tissue |
| SA | Eastern Cape | Rooikranz Dam | Buffalo River | 2015 | BU32 | SAIAB | OW14-715 | Muscle tissue |
| SA | Eastern Cape | Rooikranz Dam | Buffalo River | 2015 | BU33 | SAIAB | OW14-704 | Muscle tissue |
| SA | Eastern Cape | Rooikranz Dam | Buffalo River | 2015 | BU34 | SAIAB | OW14-762 | Muscle tissue |
| SA | Eastern Cape | Rooikranz Dam | Buffalo River | 2015 | BU35 | SAIAB | OW14-727 | Muscle tissue |
| SA | Eastern Cape | Rooikranz Dam | Buffalo River | 2015 | BU36 | SAIAB | OW14-690 | Muscle tissue |
| SA | Eastern Cape | Rooikranz Dam | Buffalo River | 2015 | BU37 | SAIAB | OW14-690 | Muscle tissue |
| SA | Eastern Cape | Rooikranz Dam | Buffalo River | 2015 | BU38 | SAIAB | OW14-690 | Muscle tissue |
| SA | Eastern Cape | Rooikranz Dam | Buffalo River | 2015 | BU39 | SAIAB | OW14-690 | Muscle tissue |
| SA | Eastern Cape | Rooikranz Dam | Buffalo River | 2015 | BU40 | SAIAB | OW14-690 | Muscle tissue |
| SA | Eastern Cape | Rooikranz Dam | Buffalo River | 2015 | BU41 | SAIAB | OW14-690 | Muscle tissue |
| SA | Eastern Cape | Rooikranz Dam | Buffalo River | 2015 | BU42 | SAIAB | OW14-690 | Muscle tissue |
| SA | Eastern Cape | Rooikranz Dam | Buffalo River | 2015 | BU43 | SAIAB | Muscle tissue |
| SA | Eastern Cape | Rooikranz Dam | Buffalo River | 2015 | BU44 | SAIAB | Muscle tissue |
| SA | Eastern Cape | Rooikranz Dam | Buffalo River | 2015 | BU45 | SAIAB | Muscle tissue |
| SA | Eastern Cape | Rooikranz Dam | Buffalo River | 2015 | BU46 | SAIAB | Muscle tissue |
| SA | Eastern Cape | Rooikranz Dam | Buffalo River | 2015 | BU47 | SAIAB | Muscle tissue |
| SA | Eastern Cape | Rooikranz Dam | Buffalo River | 2015 | BU48 | SAIAB | Muscle tissue |
Appendix 2. The scenario information used in the Approximate Bayesian Computation (ABC) implemented in DIYABC.

**Scenario 1-6**

Scenario 1: CI originated from the HN stock, which represents a subsample of the CN populations; Scenario 2: CI originated from CN populations, with both populations being derived from HN (i.e. a more recent introduction event than the one on record); Scenario 3: CI did not originate from either CN or HN population, but rather from an unsampled population; Scenario 4: CI populations represent admixed populations from both CN and HN; Scenario 5: CI populations originate from an admixture event between the sampled HN and an unsampled ghost population; Scenario 6: CI populations originate from an admixture event between the sampled CN populations and an unsampled ghost population.

**Scenario A-I**

Scenario A: Most of the CI individuals and the subsample of SA individuals (CI$_S$) are more closely related to one another than to any other population, but originated from HN stock which came from the CN gene pool. Scenario B: Both CI and CI$_S$ individuals are closest related to one another, while CN and HN are more closely related to one another. Both invasive (CI and CI$_S$) and native (CN and HN) groupings stem from a communal source population. Scenario C, like scenario A, states that CI and CI$_S$ are most closely related, originating from the CN population. Both CN and CI + CI$_S$ populations, in turn, originating from the HN stock. Scenario D proposes a closer tie between HN and CI$_S$. This grouping (HN + CI$_S$) along with CI individuals originated from a CN population. In scenario E, the HN and CI$_S$ are once again closest related to one another, originating from CN. The Remaining CI
individuals along with the HN + CI \textsubscript{5} + CN grouping originate from an unsampled population. Scenario F supports the STRUCTURE results, and states that HN and CI \textsubscript{5} are most closely related, while CI and CN are more closely related. Both groupings (HN + CI \textsubscript{5} and CI + CN) share an unsampled ghost origin. Like scenario F, scenario G groups HN and CI together and CN and CI \textsubscript{5} together. Both groupings (HN + CI and CN + CI \textsubscript{5}) originate from an unsampled ghost population. Scenario H proposes a closer tie between HN and CI. This grouping (HN + CI) along with CI \textsubscript{5} individuals originated from a CN population. Lastly, like scenario H, scenario I suggests a closer tie between HN and CI. This grouping (HN + CI) as well as the CN population each originate from independent introductions from CI \textsubscript{5}.

Scenario i-iii

The following three scenarios were run to test if both introductions (CI and CI \textsubscript{5}) did in fact originate from one source population i.e. USA (CN + HN). Scenario G: Both CI and CI \textsubscript{5} originated independently from the source population (i.e. multiple introductions from single source). Contrastingly, scenario H suggests that only CI \textsubscript{5} originated from the source population, with CI originating from CI \textsubscript{5} (i.e. single introduction). Lastly, scenario I states that both CI and CI \textsubscript{5} were founded independently from an unsampled source population, which in turn originated from the source (i.e. multiple introductions, but only a single introduction from the source).
