Sex change and effective population size: implications for population genetic studies in marine fish

Coscia I[1*], Chopelet J[2*], Waples RS[3], Mann BQ[4] & Mariani S[5$]

Running title

Sex-change and genetic estimates of N_e

Keywords

Sequential hermaphroditism, protogyny, N_e, genetic diversity, reproductive success, sea bream, Sparidae.

Addresses

[1] Laboratory of Biodiversity and Evolutionary Genomics, University of Leuven (KU Leuven), Ch. Deberiotstraat 32, Leuven 3000, Belgium

[2] School of Biology and Environmental Science, University College Dublin, Belfield, Dublin 4, Ireland

[3] NOAA Fisheries, Northwest Fisheries Science Center, Seattle WA 98112 USA

[4] Oceanographic Research Institute, PO Box 10712, Marine Parade, Durban, 4056, South Africa

[5] Ecosystems and Environment Research Centre, School of Environment and Life Sciences, University of Salford, M5 4WT, UK.

[$]**Corresponding author:** s.mariani@salford.ac.uk

[*] these authors contributed equally.

Word count: 4645 (excluding references)

Abstract

Large variance in reproductive success is the primary factor that reduces effective population size (N_e) in natural populations. In sequentially hermaphroditic ('sex-changing') fish, the sex ratio is typically skewed and biased toward the 'first' sex, while reproductive success increases considerably after sex change. Therefore, sex-changing fish populations are theoretically expected to have lower N_e than gonochorists (separate sexes), assuming all other parameters are essentially equal. In this study, we estimate N_e from genetic data collected from two ecologically similar species living along the eastern coast of South Africa: one gonochoristic, the 'santer' sea bream *Cheimerius nufar*, and one protogynous (female-first) sex-changer, the 'slinger' sea bream *Chrysoblephus puniceus*. For both species, no evidence of genetic structuring, nor significant variation in genetic diversity, were found in the study area. Estimates of contemporary N_e were significantly lower in the protogynous species, but the same pattern was not apparent over historical timescales. Overall, our results show that sequential hermaphroditism may affect N_e differently over varying time frames, and that demographic signatures inferred from genetic markers with different inheritance modes also need to be interpreted cautiously, in relation to sex-changing life-histories.

Introduction

The amount of genetic diversity in a species reflects the effective population size (N_e), which in practical terms informs on the number of breeders that contribute to the offspring, generation after generation; this number, especially in some marine animals, has been estimated to be several orders of magnitude lower than census size (N_c) (Hauser and Carvalho, 2008). Census size fluctuations and ecological perturbations are known to reduce N_e, especially in fragmented populations. However, life history traits also play a fundamental role in determining the effective size of a population (reviewed in Caballero, 1994). Lee et al. (2011) suggest that delayed age-at-maturity and lowered juvenile survival reduce N_e/N_c. Recently, age-at-maturity and adult lifespan were shown to explain half of the variance in N_e/N_c among 63 animal and plant species (Waples et al., 2013). Thus, variation in key life history traits related to mating success and survival rates, through their effect on individual lifetime reproductive success, appear to shape N_e differences among populations.

In teleosts, older and larger females generally produce more and larger eggs (Chambers and Leggett, 1996; Palumbi, 2004). Larger size also improves the mating capacity of males through behaviours such as dominance and protection of territories (Warner, 1988). Hence, growth rate is also likely a key factor determining reproductive success. However, additional reproductive traits in marine fish might account for further components of lifetime variance in reproductive success. Sequential hermaphrodites first mature as one sex and, after changing, reproduce as the opposite sex. Since younger and smaller individuals of the first sex are generally more abundant than older and larger individuals of the second sex, sequential hermaphrodite species typically present skewed sex ratios compared to gonochoristic (separate sexes) species (Allsop and West, 2004). According to the size advantage model (Ghiselin, 1969; Warner, 1975), reproductive success in sequential hermaphrodites tends to increase considerably after sex change, with individuals of the 'second sex' expected to have a greater contribution to the next generation. Therefore, the age at which individuals change sex – which has been shown to fluctuate in response to environmental factors (Hamilton et al., 2007; Mariani et al., 2013) – will have a significant impact on the lifetime reproductive success in sequential hermaphrodites. As a result, intrinsically biased sex ratios (Wright, 1931) and variance in reproductive success are theoretically expected to result in lower N_e in sequentially hermaphroditic species, compared to gonochoristic ones, assuming that other characteristics are somewhat equal. This is inevitably complicated by the flexible

nature of age/size-at-sex change in natural populations (see Avise and Mank, 2008; Mariani et al., 2013; Ross, 1990, for discussion).

In the present study, by directly examining empirical data, we investigate whether sex-changing life history may indeed determine a reduction in N_e as a result of increased lifetime variance of reproductive success and skewed sex ratio. We compared genetic data in two closely related and sympatric species (family: Sparidae) with largely comparable habitat, ecology, abundance, behaviour and life-history traits, with the exception of their reproductive modes: one species being protogynous (the slinger sea bream, *Chrysoblephus puniceus*, which first matures as female, later turning to male) and the other being gonochoristic (santer sea bream, *Cheimerius nufar*, maturing either as male or female) (Garratt, 1985a, b, 1986, 1993). *Chrysoblephus puniceus* is endemic and restricted to the south-east coast of southern Africa, while *C. nufar* is distributed over a wider area of the Western Indian Ocean. Both species are targeted by the same local commercial and recreational line fisheries, and together represent a large proportion of catches from this region (Mann and Fennessy, 2013a, b). Both species are opportunistic predators found in shoals around coastal reefs feeding on crustaceans, mollusks and small fish (Garratt, 1986). Sex-ratios were found to fluctuate greatly in *C. puniceus*, both spatially, between southern Mozambique and the KwaZulu-Natal region (Garratt, 1985a), and temporally (Mariani et al., 2013), with changes to some extent influenced by the degree of fishing pressure. In contrast, even sex ratios were found in *C. nufar* (Garratt, 1985a). Results obtained using nuclear and mitochondrial molecular markers are interpreted as a function of inheritance mode and reproductive strategy. The findings enhance our understanding of the role of life-history in population genetics and may have implications for the management of exploited populations.

Materials and methods

Sampling

Chrysoblephus puniceus and *Cheimerius nufar* specimens were collected from commercial ski-boat line fishermen between May and July 2007 at three locations along the KwaZulu-Natal coast: Port Edward (PE), Park Rynie (PR) and Richards Bay (RB) (Fig. 1). A total of 138 *C. puniceus* (122 females and 16 males) and 139 *C. nufar* (69 females and 67 males) were collected.

Fork length and weight were measured and sex was assessed by macroscopic gonad identification. We used previously published length-age relationships to derive age from fork length for both *C. puniceus* (Garratt et al., 1993) and *C. nufar* (Druzhinin, 1975; Coetzee and Baird, 1981). Fin-clips were taken from the pectoral fins and stored in absolute ethanol for later DNA extraction.

Molecular analyses

DNA was extracted using a modified Phenol-Chloroform protocol (Sambrook and Russell, 2001). DNA concentration and quality were estimated on a NanoDropTM ND-1000 spectrophotometer. Samples of both species were screened at 11 microsatellite loci, some specifically developed for this study (Chopelet et al., 2009a). Of the eleven microsatellites designed for *C. puniceus*, five cross-amplified in *C. nufar*. Six supplementary microsatellites were specifically developed for *C. nufar*, using the same protocol as in Chopelet et al. (2009a) (see Table 1 for details). Microsatellites were amplified using fluorescence-labelled forward primers (Applied Biosystems, Waltham, Massachusetts) and 2X Multiplex PCR Master Mix (QIAGEN, Hilden, Germany) in a final volume of 10µl. Depending on size and dye, fragments were amplified into two multiplexed reactions for *C. puniceus* (Chopelet et al., 2009a), while another microsatellite (SL3) was amplified separately. For *C. nufar*, one reaction contained SL25 and SA2, and the other included the nine remaining loci (Table 1). All amplifications were carried out using the same conditions. An initial step of 15 min at 95°C was followed by 30 cycles of 45s at 94°C, 45s at 60°C, and 45s at 72°C and a final extension step at 72°C for 45 min. PCR products were sized on an ABI 3130xl alongside a GS600 ladder. Genemapper v 4 (Applied Biosystems) was used for allele scoring.

Universal primers Hsp1 and Lsp1 were also used to amplify the first hypervariable region of the mitochondrial DNA control region (Ostellari et al., 1996). Each reaction was carried out using 300 ng of genomic DNA in Ready Mix (Applied Biosystems) in a final volume of 25 µl. PCR cycles were as follows: a) 95°C (5min), b) 30 cycles at 95°C (50sec), 56°C 1min and 72°C (2min), and c) a final 10 min elongation at 72°C (10 min). Amplified products were purified with exonuclease I and shrimp alkaline phosphatase (Hanke & Wink 1994) and sequenced at GATC-Biotech (Konstanz, Germany).

Statistical analyses

Genetic diversity

Genetic analyses were first performed to detect patterns of spatial structure and estimate diversity within and between locations. For microsatellite data, frequencies of null alleles were estimated using FREENA (Chapuis and Estoup, 2007). For each location sample, number of alleles (N_a), observed (H_o) and expected (H_e) heterozygozity and inbreeding coefficient (F_{IS}) were assessed using GENEPOP ON THE WEB (http://genepop.curtin.edu.au/; Raymond and Rousset, 1995; Rousset, 2008). Marker neutrality was tested in LOSITAN (Antao et al., 2008). To correct for variance in the sample size among populations, we further estimated the allelic richness (A_R) based on the minimum sample size (see Table 1) using FSTAT 2.9.3 (Goudet, 1995) and the number of private alleles using the rarefaction method (Kalinowski, 2005) implemented in ADZE 1.0 (Szpiech et al., 2008). Population differentiation was estimated using the θ estimator of F_{ST} (Weir and Cockerham, 1984) and relative confidence intervals using 10,000 permutations on the individuals in GENETIX v.4.05.2 (Belkhir et al., 1996-2004). Although the mutation rate was in our instance likely to be orders of magnitude smaller than migration rates, the corrected F'_{ST} (Hedrick, 2005) was calculated, and Jost's D_{est} (Jost, 2008) estimated using SMOGD (Crawford, 2010), and reported as additional information on genetic substructure. POWSIM was employed to evaluate the power of the dataset to detect genetic differentiation (Ryman et al., 2006). Five-hundred replications and different N_e/t ratios (500/0; 2000/2; 1000/5; 1000/10; 500/10) were used to obtain the expected F_{ST} according to this equation: $F_{ST} = 1 - (1 - 1/2N_e)^t$, with t being the number of generations of isolation (Ryman et al., 2006). Bayesian assignment was performed in STRUCTURE 2.3 (Pritchard et al., 2000; Falush et al., 2003, 2007) to infer the most likely number of genetic clusters (K) present in the datasets using the admixture model, and 500,000 iterations, after 50,000 burn-in. The number of clusters was calculated by averaging the mean posterior probability of the data L(K) over 10 independent runs.

Nucleotide and haplotype diversities were estimated from mitochondrial DNA sequences using DNASP v4.5 (Rozas et al., 2003). Median-Joining networks were constructed for both species using POPART (http://popart.otago.ac.nz).

Estimating effective population size

We used the linkage disequilibrium (LD) method implemented in LDNe (Waples and Do, 2008) to estimate contemporary N_e for each location from the microsatellite data, both including and excluding the markers that were not under HWE according to the exact test performed in Genepop. The LD method is

based on the theoretical relationship (Hill, 1981) between a measure of LD (r^2 = squared correlation of alleles at pairs of unlinked gene loci), sample size (N), and N_e. LDNe implements a modification of Hill's method that accounts for bias from ignoring second order terms in N and N_e. LDNe allows one to screen out rare alleles, which tend to upwardly bias N_e estimates, by selecting a minimum allowable allele frequency (P_{Crit}). We focused especially on P_{Crit} = 0.02, which (given the minimum sample sizes of N = 39-45; Table 2) ensured that any alleles that occurred in a single copy were not used (Waples and Do, 2010). Another estimate of effective population size was obtained using the Approximate Bayesian Computation method implemented in DIYABC 2.01 (Cornuet et al., 2014). Calculations were performed for each species, pooling samples from all locations in order to reflect the lack of genetic substructure detected in our data. Three simple scenarios were simulated. Each one represented one single population whose N_e remained constant (scenario 1), one where N_e increased after a time t_1 (scenario 2), and the third where N_e decreased after t_1. Priors were as follows: effective population size was between 10 and 10^6, and t_1 between 10 and 10^4 generations,

A longer-term view of effective population size was also obtained through estimates of historical female N_e from mtDNA data. We first used the Watterson estimator of the mutation parameter theta (θ) obtained from the number of polymorphic sites (S) (Watterson, 1975). In DNAsp v4.5, θ is defined as $2N_e\mu$ for mitochondrial DNA, where N_e is the effective population size and μ is the mutation rate per DNA sequence per generation (Tajima 1996). We estimated the female effective population size (N_ef) from the haplotype mutation rate and generation time (T) according to this equation:

$$Nef = \frac{\theta}{2\mu T}$$

We assumed a widely accepted rate μ=11% per site per million year for the Sparid mtDNA control region (Bargelloni et al., 2003; Sala-Bozano et al., 2009; Coscia et al., 2012), equal to 0.055 substitutions/site/million years. The age at maximum egg production was estimated with L_{inf} = 47 cm for *C. puniceus* and L_{inf} = 75 cm for *C. nufar* (where $L_{inf,}$, a paramenter of the von Bertalanffy growth equation, is defined as the length that an individual would reach if it grew indefinitely). This, according to Beverton (1992), can be used as an approximation of generation time (T = 5 for *C. puniceus* and 7 for *C. nufar*). Therefore, to account for life-history plasticity, we estimated N_e in both species with generation time encompassing these values: T = 3, 5 and 8.

185 Furthermore, we applied the Bayesian Skyline Plot (BSP) approach implemented in BEAST v
186 1.7 (Drummond et al., 2012) to estimate trends in past effective population size. Firstly, jModelTest 0.1.1
187 (Posada, 2008; Guindon and Gascuel, 2003) was used to select the best model of substitution for each
188 dataset via the AIC (Akaike Information Criterion): GTR (Generalised Time Reversible described in
189 Tavare (1986)) was selected for *C. nufar* and HKY (Hasegawa, Kishino and Yano 1985) for *C. puniceus*.
190 To avoid convergence issues, several independent runs (each 10^6 generations and 10% burn-in) were used
191 for each species until each effective sample size value (ESS) reached ~200 as per the user's manual.
192
193 **Results**
194 *Species and population characteristics*
195 Males (375-1404 g) were larger than females (257-952 g) in *C. puniceus*, while male (420-1436 g)
196 and female sizes (440-2820 g) overlapped in *C. nufar* (Figures S1 and S2). Individuals from the most
197 northern location, Richards Bay, had slightly larger sizes in both species (Fig. S2). Fifteen male *C.*
198 *puniceus* were found in Richards Bay and one male in Park Rynie, while only females were collected in
199 Port Edward (southernmost location). A recent study has shown that the likely age-at-sex-change for *C.*
200 *puniceus* is around 301mm, significantly lower than three decades ago (Mariani et al., 2013), and
201 potentially decreased as a result of increased fishing pressure.
202
203 *Genetic variation*
204 No evidence of null alleles was detected within the dataset. At least one marker in each of the two
205 species deviated from equilibrium expectations across all locations with strongly positive associated F_{IS}
206 values (SA3 in *C. nufar*, and SL35 in *C. puniceus*, Table S3). Since LOSITAN did not detect any signature
207 of selection (data not shown) and no significant F_{IS} was recorded for any species at each location (Table
208 2), all markers were retained in the subsequent analyses.
209 Expected and observed heterozygosities were 0.83 in all locations for *C. puniceus*, and varied
210 between 0.72 (PE and RB) and 0.74 (PR) for *C. nufar*. Allelic richness and the number of alleles were
211 higher in *C. puniceus* (overall N_A=18.4 and A_R=18.1) than *C. nufar* (overall N_A=12 and A_R=12.1) (Table
212 2). No significant F_{ST} values were found within either species (Table S1; *C. puniceus* overall F_{ST} =
213 0.0011, p=0.19; *C. nufar* overall F_{ST} = 0.0004, p=0.37). No significant genetic differentiation was found

between the three samples, for each species, irrespective of the method employed (overall D_{est} was 0.0036 for *C. puniceus* and 0.0010 for *C. nufar*) or corrections used (Hedrick's corrected G'_{ST} was 0.0047 for *C. puniceus* and 0.0019 for *C. nufar*). Hence, as expected, STRUCTURE detected one genetic cluster in each species (K=1, Fig S3). According to the power estimations implemented in POWSIM, the probability that our datasets can detect low genetic differentiations up to an F_{ST} of 0.005 is 100% (Fig. S4), with high probabilities (70-80%) also for values around 0.0025. This suggests a lack of genetic structuring among the three locations sampled at microsatellites (see also Table S1).

A total of 75 individual *C. puniceus* and 77 *C. nufar* were sequenced (179 bp and 223 bp long fragments of the control region respectively; Genbank Accession Nos.: XXXX–XXXX). Overall, nucleotide diversity π was 0.03 for *C. puniceus* and 0.04 for *C. nufar*, while haplotype diversity H_d was 0.996 for *C. puniceus*, with 65 haplotypes, and 0.984 for *C. nufar*, with 55 haplotypes. Within species, both π and H_d did not vary, with the former being 0.03 and 0.04 for *C. puniceus* and *C. nufar* respectively, across all sampling locations (Table 2). H_d ranged between 1 (Richards Bay) and 0.996 (Port Edward) for *C. puniceus*, and between 0.994 (Richards Bay) and 0.969 (Park Rynie) for *C. nufar*. No significant Φ_{ST} was detected for any of the two species between any location (Table S1), and visual inspection of median joining networks also showed a lack of geographical structure (Figures S5 and S6).

Effective population size

N_e estimates using LDNe tended to increase with extreme values of allelic frequencies (P_{crit}) (Fig. 2). However, *C. puniceus* consistently had much smaller N_e and smaller variances than *C. nufar*, irrespective of the allelic frequencies included (Fig. 2). Estimates of effective population size (N_e^) at P_{crit} = 0.02 are reported in Table 2. In *C. nufar*, negative N_e were interpreted as infinite (Waples and Do, 2010); the lower boundaries of the confidence intervals for *C. nufar* did not overlap with the highest boundaries of the *C. puniceus* estimates, using all loci, and only marginally overlapped when we excluded the locus not in HWE (Table 2). The pattern was found to be robust and consistent, whether samples were pooled or treated as separate geographical collections. Historical N_e was estimated from microsatellites using DIYABC. Of the three scenarios simulated, the second one was by far the most likely for *C. puniceus* (N_e increased in time), while scenarios 1 and 2 were equally likely for *C. nufar*. For consistency with *C. puniceus* data – and in line with population expansion results from mtDNA data (see below) – we

used scenario 2 for *C. nufar* too. Estimates of effective population size (t_0) were slightly higher for *C. puniceus* than *C. nufar*, but the confidence limits were highly overlapped (Table 2).

Historical female population size inferred from mtDNA offered further insight into demographic aspects, but the obtained estimates proved to be of different orders of magnitude depending on the method used. The Watterson method returned estimates in the order of the tens of thousands (Fig 3), slightly larger in *C. nufar* than in *C. puniceus*, all else being equal; however, considering the longer generation time in the latter, estimates for the protogynous *C. puniceus* are slightly greater than for *C. nufar* (Fig 3). The Bayesian Skyline method (Fig 4) also indicates larger effective size in *C. puniceus*, although the confidence limits largely overlap.

Discussion

No significant genetic divergence was detected among populations of either species, using either microsatellites or the variable mitochondrial control region, suggesting that the three sampled populations for each species do not belong to independent evolutionarily significant units (ESUs). This lack of genetic differentiation is also mirrored by similar levels of genetic diversity among locations. Similar findings for *C. puniceus* were recently reported by Duncan et al. (2015) who reported a single panmictic stock. The life histories and behaviour of both species are consistent with the lack of genetic differentiation within this area: they live in large shoals in the vicinity of rocky and coral reefs and are capable of migration across tens of kilometres (Garratt, 1984; Griffiths and Wilke, 2002). Based on studies of histological changes in the ovaries, it appears that spawning in *C. puniceus* occurs in large aggregations off the coasts of southern Mozambique and Northern KwaZulu-Natal (mostly to the north of Durban) and decreases to the south of Park Rynie (Garratt, 1985). Most *C. puniceus* males sampled in this study were from Richards Bay, in accordance with this pattern. The majority of eggs are fertilized and scattered above coral and rocky reefs from northern locations, and the developing larvae drift southward influenced by the strong Agulhas Current in this region (Garratt, 1985). Juveniles settling out to the south of Park Rynie, where no spawning takes place, are believed to move northward later in life to reproduce (Maggs et al., 2013). Accordingly, average sizes were found to be higher in the northernmost location of Richard's Bay. Such southward larval dispersal followed by northward adult migration could result in a constant redistribution of alleles across the area, ultimately ensuring population cohesion.

Cheimerius nufar has an extended distribution along the African east coast. The spawning area stretches from the Eastern Cape to the Gulf of Aden and individuals in spawning condition are found regularly in both Eastern Cape and KwaZulu-Natal waters (Coetzee, 1983; Garratt, 1985). Patterns of adult migration and settlement of juveniles are less well known for *C. nufar* (Griffiths and Wilke, 2002), but it is likely that other separate spawning aggregations exist in more northern, unsampled areas of the Western Indian Ocean.

This study examined population genetic inference of N_e in the context of sex-change in marine fish. Other sources of N_e variation were minimized by comparing two species with very similar taxonomical and biological characteristics and by obtaining replicate samples from the same biological units and dwelling in the same locations. No estimates of census size exist for these species, but these are also likely to be similar in the study area: *C. puniceus* and *C. nufar* are known to be the two most abundant commercial line-fishery targets in the South-African east coast, with annual landings around 300,000 and 100,000 individuals, respectively, but with the fishery that targets a habitats slightly more suited to the former (Winker et al., 2012).

All methods to calculate effective population size will assume discrete generations; therefore, underestimation of N_e may occur here, but will likely affect both species similarly (Waples et al., 2014). Predictions from population genetics theory led to the expectation that a sex changing species would present smaller N_e than a gonochoristic one for the following reasons: a) the sex ratio of the sex-changing species is biased toward the 'first sex', which is known to reduce N_e (Wright, 1931; Chopelet et al., 2009b); and b) the dynamic balance of age-at-sex change, which affects individual variance in reproductive success (V_k). Using the LD method, we find agreement with these expectations, with substantially lower effective sizes in the protogynous *C. puniceus*; while N_e estimates in *C. nufar* were one order or magnitude larger, with infinite upper confidence boundaries, indicating that more precise estimates in this species would ideally require greater sample sizes (Palstra and Ruzzante, 2008). Obviously, the N_e/N_c ratios of these species may be different, but unpublished surveys indicate that local abundances are in the same order of magnitude (Winker et al., 2012). Protogynous populations can still therefore sustain large numbers through a relatively small number of breeders; over short time scales, a population may be dominated by a relatively small number of strong, large breeding males, belonging to

two or three year classes, and able to pass their gene combinations to most of the new cohorts. On the other hand, ABC computation provides no evidence that N_e differs significantly between breeding strategies. Point estimates and confidence limits are very similar in the two species, suggesting that over evolutionary time scales, some factors are at play in reducing the expected high variance in individual reproductive success of sex-changers. One such scenario could be the inherently 'bet-hedging' nature of a sex-changing life history, whereby successful genotypes with optimal timing of first maturation and sex-change are able to maximize reproductive output as both females and males, hence reducing long-term stochasticity.

Effective population size estimates based on mtDNA do not reveal substantial differences between *C. nufar* and *C. puniceus*, except for the magnitude of the point estimates (i.e. millions with BEAST and tens of thousands with the Watterson method). Although the Bayesian Skyline Plot shows overlapping confidence limits, N_e appears larger in *C. puniceus*; with the Watterson method – once taken into account that the generation time for *C. puniceus* is lower (Garratt, 1985, Mann et al., 2014) – there also seems to be a trend towards greater N_e in *C. puniceus*. Interestingly, in a protogynous species like *C. puniceus*, virtually every individual has the potential to reproduce as a female, with only a fraction becoming males. This is likely to maintain a large mtDNA diversity over evolutionary timescales. In gonochoristic species like *C. nufar*, only about half of individuals (the females) from every new cohort will transmit their mtDNA, which may explain the patterns detected here using maternally inherited loci. Based on this, protogynous species can be expected to have mtDNA diversity similar or higher, while protandrous species would likely be lower, than that estimated for gonochoristic species.

This is potentially an important fact to be considered in conservation biology, and presumably similar (yet opposite) considerations should be applied to protandrous species (where only a smaller proportion of individuals will reproduce as females). This hypothesis can potentially be tested already using the available literature. However, expectations will always require adjustment according to the degree of plasticity of sex-changing patterns; for instance, a recent stock assessment of the *C. puniceus* population in South Africa has revealed strong evidence of recovery since the introduction of a cut in commercial fishing effort in 2003-2006 (Winker et al., 2012). This may have been facilitated by earlier age-at-sex-change inferred by Mariani et al. (2013).

Much remains to be understood about the interaction among life history traits in determining lifetime variance in reproductive success and, by reflection, the effective size of a population, especially when the added complexities of sex-change, its timing and its extent are added to the picture. Beyond fecundity, longevity and age at maturation, sex change is a particularly labile trait: individuals of the first sex might sometimes be larger than individuals of the second sex, it is not always the largest individuals of a group that change sex (Munoz and Warner, 2003), and some individuals in the population might not change sex at all (Mann and Buxton, 1998). In the case of *C. puniceus* it is likely that the large females that do not change sex make a significant contribution to future generations (Berkeley et al., 2004; Palumbi, 2004). Furthermore, the social status of individuals within the group will affect the timing of sex change, as will behavioural and/or biochemical (pheromones) interactions (Munday et al., 2006). Our initial empirical analyses show that theoretical predictions may be supported over contemporary time scales, but other mechanisms may play bigger roles over evolutionary periods.

In a population, it is generally possible to distinguish different groups such as individuals of the same age or with the same sex. The reproductive success (the average number of gametes transmitted to the next generations) of these groups can differ significantly (males can produce and transmit more gametes than females, or vice versa, depending on operational sex-ratios). In sex changing species the partitioning of reproductive success depends on the age at sex change. Some individuals change sex earlier and might present higher overall lifetime reproductive success than those changing sex at older stages, although shifts in age-at-sex-change will likely be linked with trade-offs with growth, maturity and longevity (Allsop and West, 2003). The next step towards offering generality to this indication will require a detailed understanding of how age-at-sex and other life-history traits shape the reproductive success and N_e. Furthermore, the remarkably diverse magnitude range of N_e estimates obtained through different methods, using the same data sets, suggests that, while these can be useful in a within-method comparative approach, much remains to be done to match method estimation with the appropriate time scales (Waples, 2005). We expect that the analysis of genetic data in larger sets of sex-changing and gonochoristic species pairs will be required to shed more light on the significance of sex change in affecting N_e and the evolutionary trajectories of populations and species.

Acknowledgements

This study is part of a project funded by Science Foundation Ireland. We are grateful to all the fishermen of the KwaZulu-Natal coast and the staff at the Oceanographic Research Institute, who helped with sampling. The paper was improved through the suggestions of the associate editor and three reviewers.

Data Archiving

Microsatellite and mitochondrial DNA data:

DRYAD entry XXXX → upon acceptance of manuscript.

Conflict of interests

The authors declare no conflict of interest.

References

Allsop D, West S (2003). Changing sex at the same relative body size - Similar forces may select for gender switching across taxa in all animals with this facility. *Nature* **425:** 783-784.

Allsop D, West S (2004). Sex-ratio evolution in sex changing animals. *Evolution* **58:** 1019-1027.

Antao T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G (2008). LOSITAN: a workbench to detect molecular adaptation based on a F_{ST}-outlier method. *BMC Bioinformatics* **9**.

Avise JC, Mank J (2009). Evolutionary perspectives on hermaphroditism in fishes. *Sexual Development* **3:** 152-163.

Bargelloni L, Alarcón JA, Alvarez MC, Penzo E, Magoulas A, Reis C *et al* (2003). Discord in the family Sparidae (Teleostei): divergent phylogeographical patterns across the Atlantic-Mediterranean divide. *J Evol Biol* **16**(6)**:** 1149-1158.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F. (1996-2004). Laboratoire Génome, Populations, Interactions, CNRS UMR 5171, Université de Montpellier II, Montpellier (France).

Berkley S, Chapman C, Sogard S (2004). Maternal age as a determinant of larval growth and survival in a marine fish, *Sebastes melanops*. *Ecology* **85:** 1258-1264.

Beverton RJH (1992). Patterns of reproductive strategy parameters in some marine teleost fishes. *J Fish Biol* **41:** 137-160.

Caballero A (1994). Developments in the prediction of effective population size. *Heredity* **73:** 657-679.

Chambers RC, Leggett WC (1996). Maternal influences on variation in egg sizes in temperate marine fishes. *Am Zool* **36**(2)**:** 180-196.

Chapuis MP, Estoup A (2007). Microsatellite null alleles and estimation of population differentiation. *Mol Biol Evol* **24**(3)**:** 621-631.

Chopelet J, Helyar S, Mann B, Mariani S (2009a). Novel polymorphic microsatellite loci for the protogynous hermaphrodite slinger sea bream (*Chrysoblephus puniceus*, Sparidae). *Molecular Ecology Resources* **9**(4)**:** 1223-1226.

Chopelet J, Waples RS, Mariani S (2009b). Sex change and the genetic structure of marine fish populations. *Fish Fish* **10**(3)**:** 329-343.

Coetzee P (1983). Seasonal histological and microscopical changes in the gonads of *Cheimerius nufar* (Ehrenberg, 1820) (Sparidae: Pisces). *S Afr J Zool* **18:** 76-88.

Coetzee P, Baird D (1981). Age, growth and food of *Cheimerius nufar* (Ehrenberg, 1820) (Sparidae), collected off St Croix Island, Algoa Bay. *S Afr J Zool* **18:** 76-88.

Coscia I, Vogiatzi E, Kotoulas G, Tsigenopoulos C, Mariani S (2012a). Exploring neutral and adaptive processes in expanding populations of gilthead sea bream, Sparus aurata L., in the North-East Atlantic. *Heredity* **108**(5)**:** 537-546.

Crawford NG (2010). SMOGD: software for the measurement of genetic diversity. *Molecular Ecology Resources* **10**(3)**:** 556-557.

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol*.

Druzhinin A (1975). Some data on sparid fishes (Fam. Sparidae) of the Gulf of Aden region. *J Ichthyol* **15:** 531-541.

Duncan M, James N, Fennessy S, Mutombene R, Mwale M (2015). Genetic structure and consequences of stock exploitation of Chrysoblephus puniceus, a commercially important sparid in the South West Indian Ocean. *Fish Res* **164:** 64-72.

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

Falush D, Stephens M, Pritchard JK (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164**(4)**:** 1567-1587.

Falush D, Stephens M, Pritchard JK (2007). Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Mol Ecol Notes* **7**(4)**:** 574-578.

Garratt P (1985a). The Offshore Linefishery of Natal: I. Exploited Population Structures of the Sparids: *Chrysoblephus puniceus* and *Cheimerius nufar*. *Investigational Report Oceanographic Research Institute* **62:** 1-18.

Garratt P (1985b). The offshore linefishery of Natal: II. Reproductive biology of the sparids *Chrysoblephus puniceus* and *Cheimerius nufar*. *Investigational Report Oceanographic Research Institute* **62:** 1-21.

Garratt P (1986). Protogynous hermaphroditism in the slinger *Chrysoblephus puniceus* (Gilchrist & Thompson 1908) (Teleostei: Sparidae). *J Fish Biol* **28:** 297-306.

Garratt P (1993). Comparative aspects of the reproductive biology of seabreams (Pisces: Sparidae). PhD thesis, Rhodes University, Grahamstown, South Africa.

Garratt P, Govender A, Punt A (1993). Growth acceleration at sex change in the protogynous hermaphrodite *Chrysoblephus puniceus* (Pisces: Sparidae). *S Afr J Mar Sci* **13:** 187-193.

Ghiselin MT (1969). Evolution of hermaphroditims among animals. *Q Rev Biol* **44**(2)**:** 189-&.

Goudet J (1995). FSTAT (Version 1.2): A computer program to calculate F-statistics. *J Hered* **86**(6)**:** 485-486.

Griffiths M, Wilke C (2002). Long-term movement patterns of five temperate reef-fishes (Pisces: Sparidae): implications for marine reserves. *Mar Freshw Res* **53:** 233-244.

Guindon S, Gascuel O (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* **52:** 696-704.

Hamilton S, Caselle J, Standish J, Schroeder D, Love M, Rosales-Casian J *et al* (2007). Size-selective harvesting alters life histories of a temperate sex-changing fish. *Ecol Appl* **17:** 2268-2280.

Hanke M, Wink M (1994). Direct DNA sequencing of PCR-amplified vector inserts following enzymatic degradation of primers and dNTPs. *BioTechniques* **17**(5)**:** 858-860.

Hauser L, Carvalho GR (2008). Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. *Fish Fish* **9**(4)**:** 333-362.

Hedrick PW (2005). A standardized genetic differentiation measure. *Evolution* **59**(8)**:** 1633-1638.

Hill W (1981). Estimation of effective population size from data on linkage disequilibrium. *Genet Res* **38:** 209-216.

Jost L (2008). G(ST) and its relatives do not measure differentiation. *Mol Ecol* **17**(18)**:** 4015-4026.

Kalinowski ST (2005). HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. *Mol Ecol Notes* **5**(1)**:** 187-189.

Lee AM, Engen S, Saether B-E (2011). The influence of persistent individual differences and age at maturity on effective population size. *Proceedings of the Royal Society B-Biological Sciences* **278**(1722)**:** 3303-3312.

Maggs J, Mann B, Cowley P (2013). Reef fish display station-keeping and ranging behaviour in the Pondoland Marine Protected Area on the east coast of South Africa. *Afr J Mar Sci* **35:** 183-193.

Mann B, Buxton C, Russell B, Pollard D, Carpenter K. (2014). *The IUCN Red List of Threatened Species, v 2014.3*.

Mann B, Buxton CD (1998). The reproductive biology of *Diplodus sargus capensis* and *D. cervinus hottentotus* (Sparidae) off the south-east Cape coast, South Africa. *Cybium* **22**(1)**:** 31-47.

Mann B, Fennessy S. (2013a). *Southern African Marine Linefish Species Profiles, Vol. 9*. Mann B (ed.). Oceanographic Research Institute: Durban, pp 212-214.

Mann B, Fennessy S. (2013b). *Southern African Marine Linefish Species Profiles, Vol. 9*. Mann B (ed.). Oceanographic Research Institute: Durban, pp 225-227.

Mariani S, Sala-Bozano M, Chopelet J, Benvenuto C (2013). Spatial and temporal patterns of size-at-sex-change in two exploited coastal fish. *Environ Biol Fishes* **96**(4)**:** 535-541.

Munday PL, White JW, Warner RR (2006). A social basis for the development of primary males in a sex-changing fish. *Proceedings of the Royal Society B-Biological Sciences* **273:** 2845-2851.

Munoz RC, Warner RR (2003). A New Version of the Size-Advantage Hypothesis for Sex Change: Incorporating Sperm Competition and Size-Fecundity Skew. *The American Naturalist* **161**(5)**:** 749-761.

Ostellari L, Bargelloni L, Penzo E, Patarnello P, Patarnello T (1996). Optimization of single-strand conformation polymorphism and sequence analysis of the mitochondrial control region in Pagellus

519 bogaraveo (Sparidae, Teleostei): Rationalized tools in fish population biology. *Anim Genet* **27**(6)**:** 423-
520 427.
521
522 Palstra FP, Ruzzante DE (2008). Genetic estimates of contemporary effective population size: what can
523 they tell us about the importance of genetic stochasticity for wild population persistence? *Mol. Ecol.* **17**:
524 3428–3447.
525
526 Palumbi S (2004). Why mothers matter. *Nature* **430:** 621-622.
527
528 Posada D (2008). jModelTest: Phylogenetic Model Averaging. *Mol Biol Evol* **25:** 1253-1256.
529
530 Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus
531 genotype data. *Genetics* **155**(2)**:** 945-959.
532
533 Raymond M, Rousset F (1995). GENEPOP (version 1.2): population genetics software for exact tests and
534 ecumenicism. *J Hered* **86:** 248-249.
535
536 Ross R (1990). The evolution of sex-change mechanisms in fishes. *Environ Biol Fishes* **29:** 81-93.
537
538 Rousset F (2008). GENEPOP ' 007: a complete re-implementation of the GENEPOP software for
539 Windows and Linux. *Molecular Ecology Resources* **8**(1)**:** 103-106.
540
541 Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003). DnaSP, DNA polymorphism analyses by
542 the coalescent and other methods. *Bioinformatics* **19**(18)**:** 2496-2497.
543
544 Ryman N, Palm S (2006). POWSIM: a computer program for assessing statistical power when testing for
545 genetic differentiation. *Mol Ecol Notes* **6**(3)**:** 600-602.
546
547 Sala-Bozano M, Ketmaier V, Mariani S (2009). Contrasting signals for multiple markers illuminate
548 population connectivity in a marine fish. *Mol Ecol* **18**(23)**:** 4811-4826.
549
550 Sambrook J, Russell D (2001). *Molecular Cloning*, 3rd edn. Cold Spring Harbor Laboratory Press: New
551 York.
552
553 Szpiech ZA, Jakobsson M, Rosenberg NA (2008). ADZE: a rarefaction approach for counting alleles
554 private to combinations of populations. *Bioinformatics* **24**(21)**:** 2498-2504.
555
556 Waples RS (2005). Genetic estimates of contemporary effective population size: To what time periods do
557 the estimates apply? *Molecular Ecology* **14**: 3335-3352.
558
559 Waples RS, Antao T, Luikart G (2014). Effects of overlapping generations on linkage disequilibrium
560 estimates of effective population size. *Genetics* **197:** 769-780.
561
562 Waples RS, Do C (2008). LDNE: a program for estimating effective population size from data on linkage
563 disequilibrium. *Molecular Ecology Resources* **8**(4)**:** 753-756.
564
565 Waples RS, Do C (2010). Linkage disequilibrium estimates of contemporary N-e using highly variable
566 genetic markers: a largely untapped resource for applied conservation and evolution. *Evolutionary*
567 *Applications* **3**(3)**:** 244-262.
568
569 Waples RS, Luikart G, Faulkner JR, Tallmon DA (2013). Simple life-history traits explain key effective
570 population size ratios across diverse taxa. *Proceedings of the Royal Society B-Biological Sciences*
571 **280**(1768).
572
573 Warner RR (1975). Adaptive significance of sequential hermaphroditism in animals. *The American*
574 *Naturalist* **109:** 61-82.
575
576 Warner RR (1988). Sex change and the size-advantage model. *Trends Ecol Evol* **3:** 133-136.

Watterson G (1975). On the number of segregating sites in genetical models without recombination. *Theor Popul Biol* **7:** 256-276.

Weir B, Cockerham C (1984). Estimating F-statistics for the analysis of population structure. *Evolution* **38:** 1358 - 1370.

Winker H, Kerwath S, Attwood C. (2012). South African Department of Agriculture, Forestry and Fisheries, Linefish Scientific Working Group Report No. 3, pp 65.

Wright S (1931). Evolution in Mendelian Populations. *Genetics* **16:** 97-159.

FIGURE TITLES AND LEGENDS

Figure 1 – MAP: Sampling locations off the KwaZulu Natal Coast of South Africa.

Figure 2 – EFFECTIVE POPULATION SIZE ESTIMATES: Effective population size of *Cheimerius nufar* (santer, in red) and *Chrysoblephus puniceus* (slinger, in green) estimated with different allelic frequencies (or P_{crit}). The three sampled populations were pooled.

Figure 3 – HISTORICAL FEMALE EFFECTIVE POPULATION SIZE: Historical female effective population size (N_ef) of slinger, *Chrysoblephus puniceus* (continuous line) and santer, *Cheimerius nufar* (dotted line) based on θ for different mutation rates (μ) and three generation lengths (T = 3, 5 and 8 years). Red values correspond to N_ef estimations for μ = 11% and T = 5.

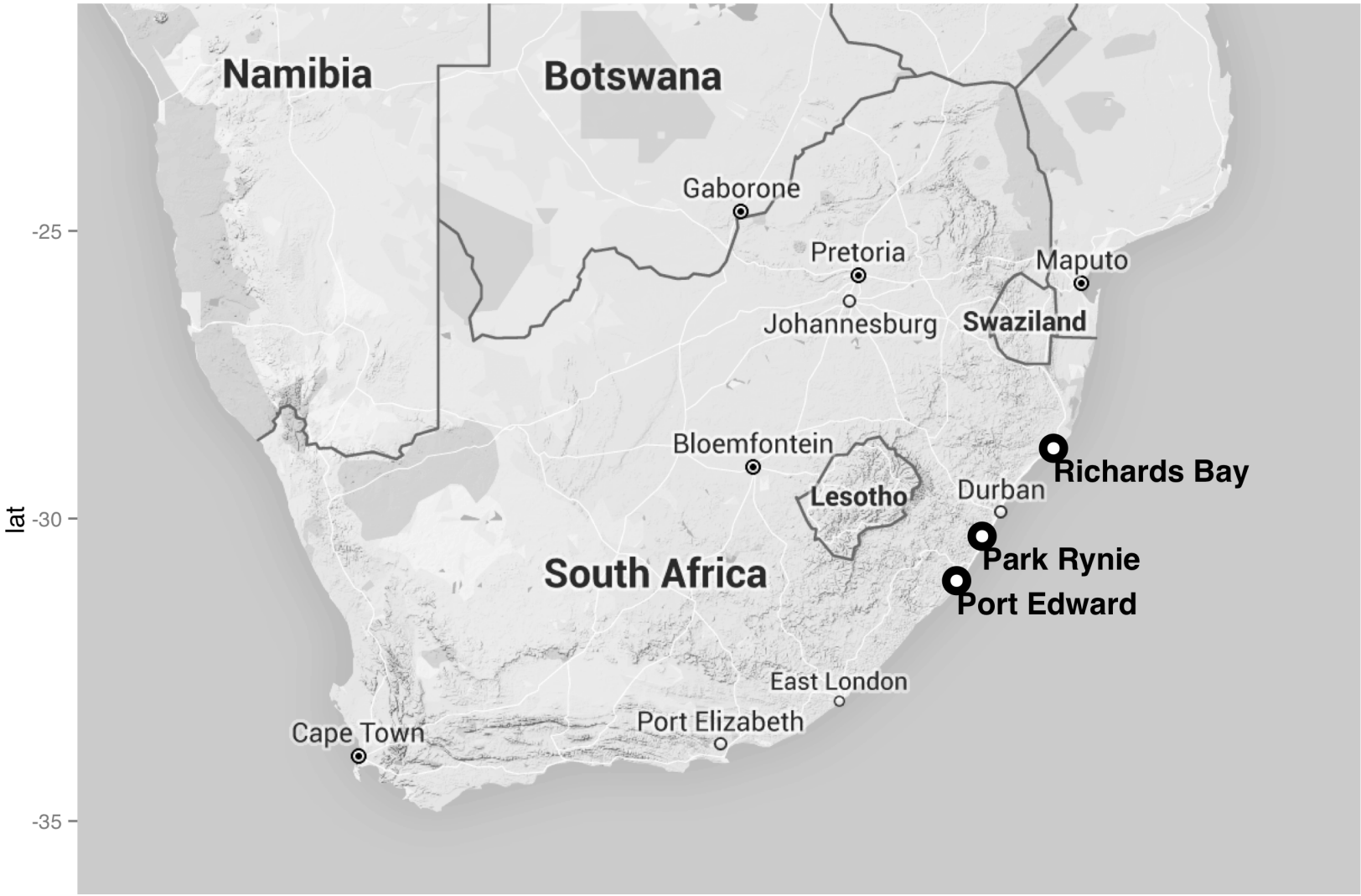
Figure 4: BAYESIAN SKYLINE PLOTS: Bayesian Skyline plots for santer, *Cheimerius nufar* and slinger, *Chrysoblephus puniceus*. The continuous and dashed lines represent the median and mean values, respectively.

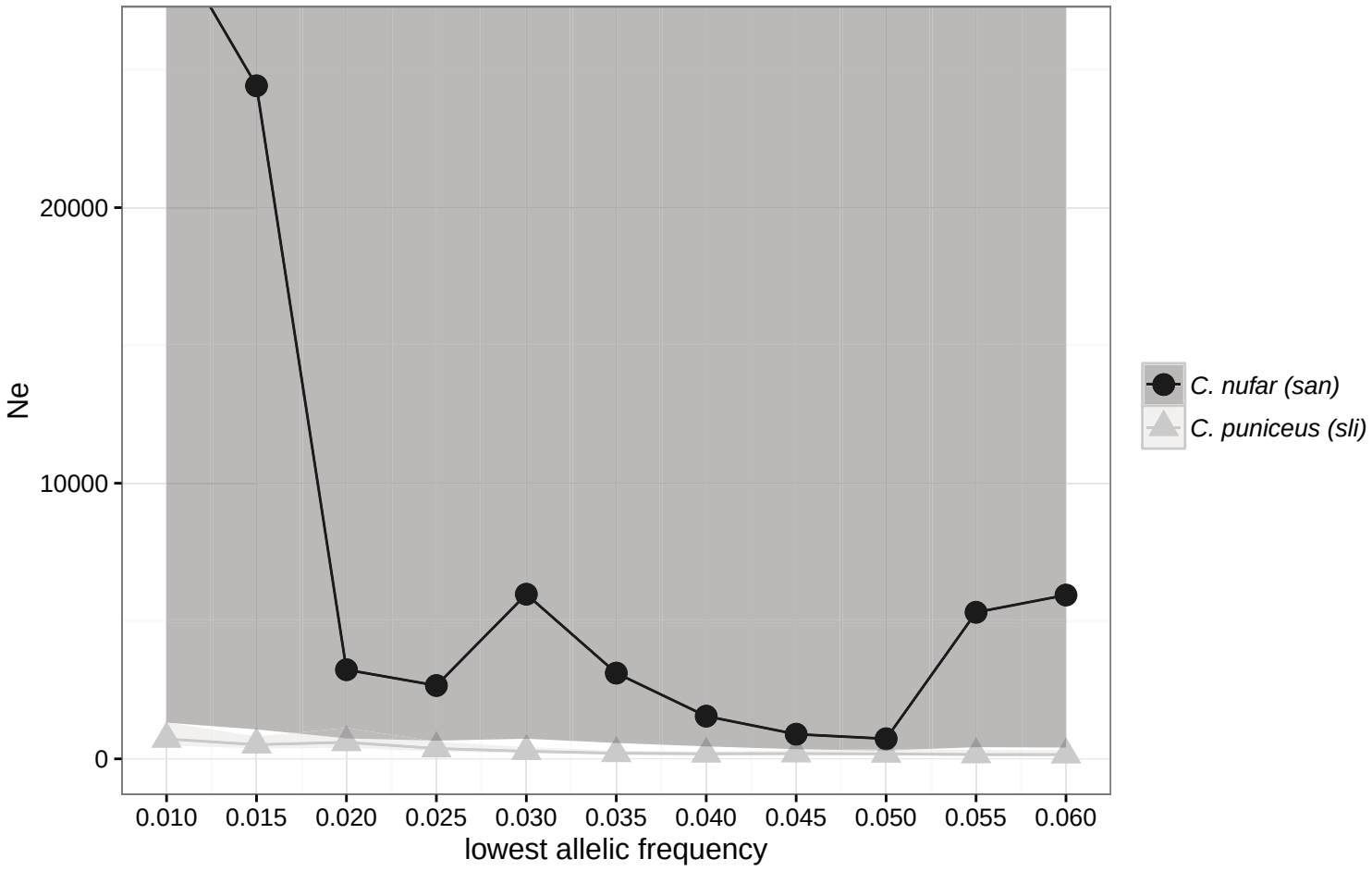
Table 1: Variability of ten polymorphic microsatellite loci in santer, *Cheimerius nufar*, from South Africa (n = 131) grouped into two multiplex reactions (I and II). Ta, annealing temperature (°C); Na, number of alleles observed; H_o, observed heterozygosity; H_e, expected heterozygosity. For slinger, *Chrysoblephus puniceus*, see Chopelet et al. (2009a).

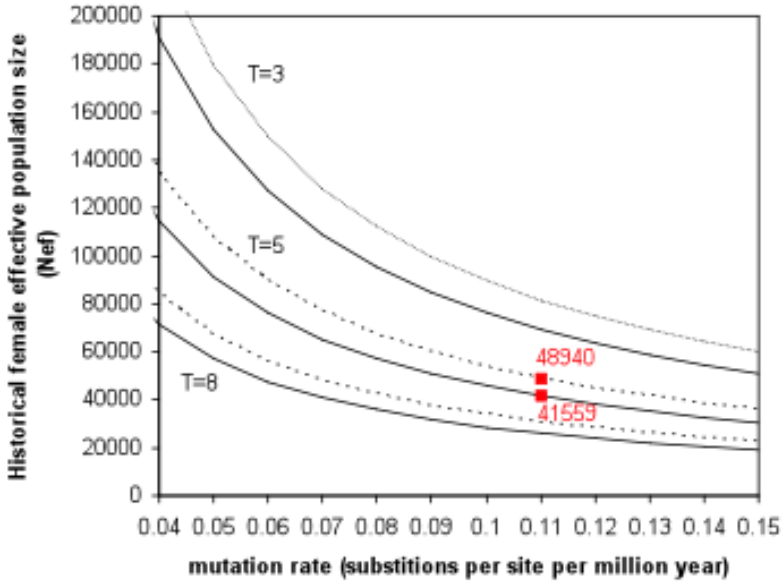
Locus name	GenBank Accession Number	Dye label	Primer sequences (5'-3')	Repeat Motif	T_a	Size Range (bp)	N_A	H_e	H_o
SA1 (I)		6FAM	F:CAGCGATGCACAGTAAAGTACC R:AGCATACAGAGGCCTTCAGC	$(TG)_{33}$	58	253-293	19	0.93	0.94
SA2 (II)		NED	F:GAGCCAGACTCCAGACATCC R:CCGGACAGGAGTATTGAAGC	$(GTCT)_{11}$	58	190-234	8	0.75	0.73
SA3 (I)		PET	F:CCAGAGTCTGTGCTGTGAGTGC R:TCCTTGTGGTCCACTTTACG	$(CA)_{17}$	58	386-420	13	0.89	0.77
SA6 (I)		6FAM	F:AGCTGCTGCTCATCTCACG R:GCAGTGTTAACATCTTCGAATGC	$(TG)_{12}$	58	187-209	9	0.75	0.79
SA10 (I)		6FAM	F:GAAGCCAAACGAGGACAGC R:GTGAGGAGCATGCTAATACCG	$(GT)_{15}GA(GT)_{26}$ $GAGTGAGTGA(GT)_{18}$	58	428-532	42	0.96	0.92
SA25 (I)		VIC	F:GGAGGAAATGAACCGATGG R:GCAGCTGGTCAATAGTGTGG	$(TG)_7CA(TG)_6$	58	152-220	19	0.61	0.70
SL25 (II)	FJ526983	NED	F:GGTACTGTTTGGCCCTTGC R:GCCTGGTAATATGCCTGAGC	$(GA)_{11}TG(GA)_3TA(GA)_4GTCA$ $(GA)_2AA(GA)_9CA(GA)_2AA(GA)_{12}$	58	207-271	7	0.61	0.66
SL26 (I)	FJ526984	PET	F:TGAAGGTGCTGATGACTTTCC R:CAGTCCTGCCTCTGACTGG	$(TC)_2TT(TC)_9GC(TC)_4TGCCTT$ $(TC)_5GC(TC)_7$	58	239-247	5	0.38	0.32
SL27 (I)	FJ526985	VIC	F:CAGCCTCAGCTCATTTCTCC R:CCTGCCCTCCTGTAGATGC	$(TG)_{55}$	58	183-201	6	0.35	0.32
SL34 (I)	FJ526988	VIC	F:GCGTGCACACTCTTACAGTACC R:TCGGATGTGCATCTCATAGG	$(CA)_{17}$	58	321-361	19	0.84	0.85

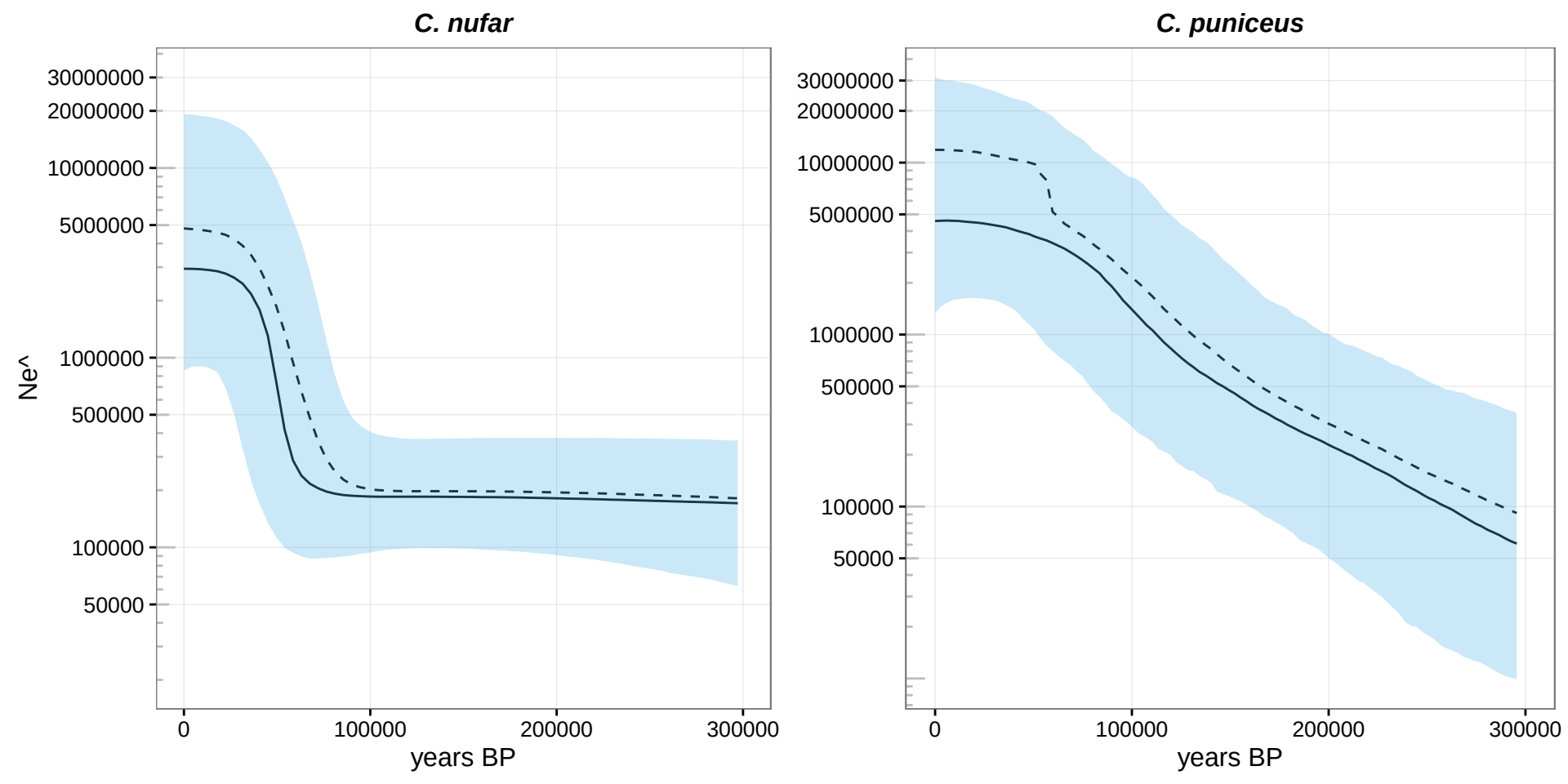
Table 2: Genetic diversity of slinger, *Chrysoblephus puniceus* and santer, *Cheimerius nufar* estimated with microsatellites and mitochondrial DNA. N, number of samples analysed; N_A, number of alleles; A_R, allelic richness; *H*e, expected heterozygosity; *H*o, observed heterozygosity; *F*is, inbreeding coefficient; N_{hap}, number of haplotypes; *H*d, haplotype diversity; π, nucleotide diversity. The three N_e estimates refer to, in order: i) estimates based on LDNe using all loci, ii) estimates based on LDNe excluding loci departing from HW equilibrium, iii) estimates based on DIYABC.

		microsatellites								mitochondrial			
	location	**N**	N_A	A_R	***H*e**	***H*o**	***F*is**	***N*e^**	***N*e^(HWE)LD**	***N*e^(HWE)ABC**	**N**	N_{hap}	***H*d**	**p**
C. puniceus *[slinger]*	PE	43	18.5	18	0.83	0.78	0.05	835(287-∞)	707(259-∞)		24	23	0.996	0.03
	PR	41	18.7	18.2	0.83	0.81	0.02	128(93-187)	535(228-∞)		25	24	0.997	0.03
	RB	38	18	18	0.83	0.79	0.04	166(114-292)	462(198-∞)		26	26	1	0.03
	ALL	***125***	***18.4***	***18.1***	***0.83***	***0.79***	***0.04***	***371(289-511)***	***601 (404-1124)***	***560,000 (52,000-980,000)***	***75***	***65***	***0.996***	***0.03***
C. nufar *[santer]*	PE	43	11.6	11.8	0.72	0.74	0.06	∞(-261-∞)	∞ (327-∞)		25	22	0.99	0.04
	PR	45	12.4	12.4	0.74	0.76	-0.04	∞(-340-∞)	∞ (733-∞)		26	20	0.969	0.04
	RB	42	12.1	12.1	0.72	0.70	0.02	∞(405-∞)	∞ (353-∞)		26	24	0.994	0.04
	ALL	***131***	***12***	***12.1***	***0.73***	***0.73***	***0.01***	***2236(654-∞)***	***3233 (743-∞)***	***370,000 (23,000-960,000)***	***77***	***55***	***0.984***	***0.04***


Namibia
Botswana
Gaborone
Pretoria
Maputo
Johannesburg
Swaziland
Bloemfontein
Richards Bay
Lesotho
Durban
Park Rynie
South Africa
Port Edward
East London
Port Elizabeth
Cape Town
-25
-30
-35
lat


Ne
20000
10000
0
0.010
0.015
0.020
0.025
0.030
0.035
0.040
0.045
0.050
0.055
0.060
lowest allelic frequency
C. nufar (san)
C. puniceus (sli)


Historical female effective population size (Nef)
200000
180000
160000
140000
120000
100000
80000
60000
40000
20000
0
T=3
T=5
T=8
48940
41559
0.04 0.05 0.06 0.07 0.08 0.09 0.1 0.11 0.12 0.13 0.14 0.15
mutation rate (substitions per site per million year)


C. nufar
C. puniceus
30000000
20000000
10000000
5000000
1000000
500000
100000
50000
Ne^
0
100000
200000
300000
years BP