Biological Journal of the Linnean Society, 2018, **XX**, 1–9. With 2 figures.

Phylogeography and population genetics of the Australian small carpenter bee, *Ceratina australensis*

ROBERT L. OPPENHEIMER, WYATT A. SHELL and SANDRA M. REHAN*

Department of Biological Sciences, University of New Hampshire, Durham, NH 03824, USA

Received 23 March 2018; revised 10 May 2018; accepted for publication 11 May 2018

The Australian small carpenter bee, *Ceratina australensis*, is the sole member of the small carpenter bees, genus *Ceratina*, in Australia. *Ceratina australensis* is found throughout eastern Australia in dead, broken stems of weedy plant species, where it makes its nests. Here we developed eight microsatellite loci and used these markers to characterize the genetic structure of three populations across three distinct ecoregions that are connected by the Murray–Darling River basin. We genotyped 57 female bees and found significant and geographically consistent variation in allelic diversity and heterozygosity between populations. Through comparisons with the results of mitochondrial DNA screening from a previous study, we infer the possibilities of male-biased dispersal and limited interpopulation migration. Based on available distribution data, the expansion of *C. australensis* into arid regions appears to follow the Murray–Darling River basin and associated waterways. This basin has undergone severe anthropogenic disturbance since European settlement of Australia, with large-scale changes to native vegetation communities. Although it is unknown how these changes have affected local insect communities, *C. australensis* does not rely on a single plant species for nesting habitat. We provide evidence for the hypothesis that expansion of *C. australensis* into Australia might have been aided by the introduction of non-native pithy stemmed plants.

ADDITIONAL KEYWORDS: biogeography – dispersal – microsatellite markers – native bee – population genetics – range expansion.

INTRODUCTION

Species distributions are highly dependent on the relative fitness of individuals in different environmental contexts (Hoffman & Blows, 1994). As a small founding population attempts to expand its range, it is likely to face new environmental challenges and reduced fitness. Repeated cycles of attempted founding events and subsequent extirpation can leave evidence of genetic bottlenecks in populations at range edges (Eckert, Samis & Lougheed, 2008; Cortázar-Chinarro *et al.*, 2017). Recent or particularly severe bottlenecks are expected to decrease overall genetic diversity, removing rare alleles, decreasing the effective population size (N_e) and increasing the relatedness between individuals in successive generations (Cornuet & Luikart, 1996). If immigration into a bottlenecked population is low, the effects of genetic drift will be amplified, leading either to the deleterious effects of drift and possible extirpation or to genetic segregation from source populations and evolution.

*Corresponding author. E-mail: sandra.rehan@unh.edu

As has been observed in many invasive species, some taxa are able to overcome the deleterious effects of inbreeding despite initially small effective population sizes (Tsutsui & Suarez, 2003; Arca *et al.*, 2015). Such species may avoid inbreeding through behavioural adaptions, such as kin recognition (Kukuk & Decelles, 1986; Jongepier & Foitzik, 2016) or sex-biased dispersal (Pusey, 1987; Baines, Ferzoco & McCauley, 2017), either of which limits the probability of mating with relatives. Those organisms that do not strictly conform to patterns of Mendelian inheritance may be particularly well equipped to reproduce successfully with limited outbreeding. Haplodiploid organisms, for instance, may be resistant to the effects of inbreeding owing to the increased selective pressures on the haploid sex (Lin & Michener, 1972).

Studies across haplodiploid Hymenoptera have repeatedly shown that certain members of this group demonstrate a strong resilience to negative founder effects otherwise expected during iterative range expansions (Schmid-Hempel *et al.*, 2007; Arca *et al.*, 2015; López-Uribe *et al.*, 2016; Soro *et al.*, 2017). However, Hymenoptera may be susceptible to the

increased deleterious effects of reduced allelic diversity in small populations owing to the complementary sex determination system (Asplen *et al.*, 2009), which is expected to produce sterile diploid males during increased inbreeding and homozygosity, greatly increasing the risk of extinction (Zayed & Packer, 2005). Male-mediated sex-biased dispersal can prevent inbreeding and has been shown to structure populations of species in which females nest philopatrically (Boomsma, Baer & Heinze, 2005; Ulrich, Perrin & Chapuisat, 2009; López-Uribe *et al.*, 2014, 2015).

Ceratina small carpenter bees have a cosmopolitan distribution and are represented by ~200 species of generalist pollinators (Michener, 2007; Rehan & Schwarz, 2015). Although speciose on most continents, the genus is represented by a single species, *Ceratina australensis* (Perkins), in Australia. This species inhabits a wide variety of climates across its range, including subtropical and temperate forests, persistently dry grasslands and temperate coastal dunes (Dew, Rehan & Schwarz, 2016). *Ceratina australensis* is largely subsocial; most nests contain a single female, which provides extended maternal care to her brood (Michener, 1974; Rehan, Richards & Schwarz, 2010). However, this species also demonstrates a facultative capacity for social nesting; a low frequency of nests within a given population (3–20%) will contain two females, which operate under a reproductive division of labour (Rehan *et al.*, 2010; Rehan, Schwarz & Richards, 2011; Dew, Shell & Rehan, 2018). Social nesting in this species is thought to be reinforced by the limited dispersal of females from natal nests (Rehan *et al.*, 2014). Dispersal patterns have been implicated in structuring social systems in halictid sweat bees by increasing the densities of philopatric females and varying levels of drifting between nests (Kukuk & Decelles, 1986; Yanega, 1990; Ulrich *et al.*, 2009; Johnstone, Cant & Field, 2012; Friedel, Paxton & Soro, 2017). As such, assessing the phylogeography and population genetics of *C. australensis* can improve our understanding of both the historical biogeography and the sociobiology of facultatively social insects.

A previous assessment of mitochondrial *CO1* haplotype variation across the *C. australensis* range supported an ancestral dispersal out of Asia, coinciding with the last major glaciation event ~18 kya, followed by expansion south and west (Dew *et al.*, 2016). The most southern population sampled during that study was represented by a single haplotype, which indicated either a recent genetic bottleneck or a founding event followed by little to no additional immigration. Comparisons between mitochondrial DNA (mtDNA) and nuclear DNA (nucDNA) shed light on maternal and paternal lineages and potential population structuring by sex-biased dispersal (Ross & Shoemaker, 1997; Jorde, Bamshad & Rogers, 1998; Goudet, Perrin & Waser, 2002; Ulrich *et al.*, 2009; Ambrose *et al.*, 2014; Tamang *et al.*, 2018). Here we develop and use a suite of microsatellite markers for *C. australensis* to gain insights into the species' population genetics and evolutionary history through analysis of the phylogeographical structure of three focal populations.

MATERIAL AND METHODS

Field collections and demographic assessments

Bees were collected from nests representing three geographically distinct *C. australensis* populations in January of 2015, 2016, and 2017 (Fig. 1). Our Queensland (QLD) population (28.24°S, 152.09°E) was drawn from subtropical temperate forests characterized by warm summers and cold winters; Victoria (VIC; 34.15°S, 142.16°E) from a semi-arid riverine area with hot, dry summers and cold winters; and our South Australia (SA) population (34.94°S, 138.50°E) from coastal dunes with warm summers and cold winters (Dew *et al.*, 2016, 2018). Victoria is a drier region of Australia than either QLD or SA, but is connected to both regions by a series of waterways which form the Murray–Darling River basin (MDRB; Fig. 1). Lake Alexandrina, which lies ~100 km southwest of our SA collection site, is not a part of the MDRB, but is fed by the Murray River. The area between Lake Alexandrina and the collection site is not restricted by arid habitat. During sampling, nests were found primarily within stems of giant fennel (*Ferula communis*) in QLD, dark sago-weed (*Plantago drummondii*) in VIC, and European searocket (*Cakile maritima*) in SA (Dew *et al.*, 2018). Collections were made during dawn and dusk, outside of *C. australensis* foraging hours, which greatly increases the chances of collecting all nest inhabitants. Nests were first refrigerated to sedate occupants and then opened by cutting the stems in half lengthwise with a pocket knife. The contents of each brood cell were then carefully recorded and bees immediately frozen in liquid nitrogen and stored at –80 °C.

Microsatellite development

Eight microsatellite primers were designed using the *C. australensis* genome (S.M.R., unpubublished data), following Shell & Rehan (2016). Forward primers were designed with an M13 oligo extension, following Schuelke (2000), to incorporate a universal fluorescently labelled dye (VIC or FAM from the DS-33 set). DNA was extracted from the abdominal tissues of collected adult females using a modified phenol–chloroform extraction (Kirby, 1965). Polymerase chain reaction (PCR) was then carried out using a total reaction volume of 10 µL [4.975 µL deionized water; 1 µL 10× buffer, 1 µL

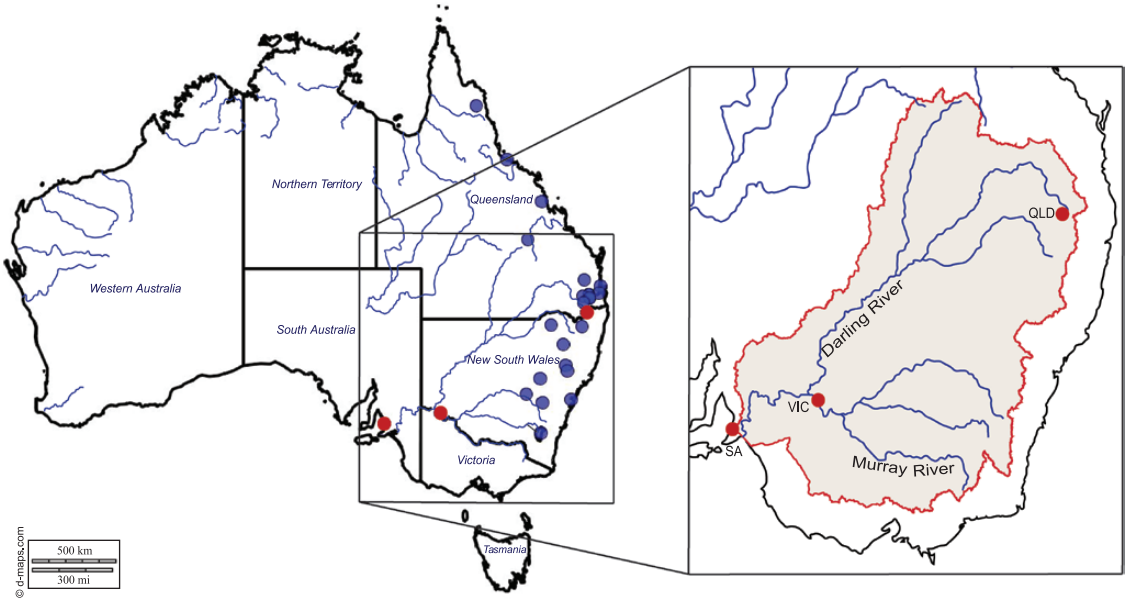

Figure 1. Map of Australia, indicating all known *Ceratina australensis* collection sites (circles) and the Murray–Darling River basin (red outline). Red circles denote populations sampled for use in this study; blue circles denote other collection sites, recorded by the Atlas of Living Australia (ala.org.au) and Michener (1962).

$MgCl_2$ (25 mM), 0.125 µL forward primer with M13 tail (10 mM), 0.5 µL reverse primer (10 mM), 0.5 µL M13 oligo (10 mM); 0.2 µL dNTPs (10 mM); 0.2 µL recombinant Taq DNA polymerase (Fisher Scientific, Pittsburgh, PA, USA); and 1.5 µL of DNA template] in an Eppendorf Mastercycler gradient thermocycler (Eppendorf North America, Hauppauge, NY, USA). The PCR runs involved the following five stages: (1) an initial denaturing at 94 °C for 1 min; (2) a touchdown series of ten cycles starting at 94 °C for 10 s, 60 °C for 15 s, then cooling incrementally to primer-specific annealing temperature, and a 72 °C extension for 15 s; (3) 20–35 cycles at 94 °C for 15 s, followed by primer-specific annealing temperature for 15 s; a second extension at 72 °C for 25 s; (4) eight cycles at 94 °C for 10 s, 53 °C for 15 s, 72 °C for 20 s; and (5) a final extension at 72 °C for 10 min. The PCR products were sent to the DNA Analysis Facility at Yale University for fragment analysis on a 3730xl Analyzer. Allele data were then scored using Peak Scanner 2 (Applied Biosystems, Foster City, CA, USA). We assessed the genetic profiles of 57 females from separate nests across the range of *C. australensis* (QLD, 15; SA, 31; VIC, 11). These eight microsatellites were uploaded to GenBank under accession numbers MH061300–MH061307 (Supporting Information, Table S1).

Locus characteristics and population structure

Tests for deviation from Hardy–Weinberg equilibrium (HWE) and for loci displaying linkage disequilibrium were conducted in GenePop 4.2 (Rousset, 2008) using one female per nest. Allele number (N_a), effective allele number (A_e) and observed and expected heterozygosity (H_o and H_e) were calculated using GenAlEx 6.502 (Peakall & Smouse, 2006). Allele number is a measure of allelic richness, and A_e is a measure of allelic diversity. Effective allele number accounts for the frequency of alleles, such that alleles with low frequency contribute less to A_e than alleles that occur frequently. The values of N_a, A_e, H_o and H_e were compared between populations using the Conover–Iman test for multiple comparisons (Conover & Iman, 1979) in the package conover.test in R v1.1.383 (R Core Team, 2017). An analysis of molecular variance (AMOVA) was carried out in ARLEQUIN v.3.5.2.2 (Excoffier, Laval & Schneider, 2005) to assess within-, among- and between-population variances during a 1000 permutation run. Genetic fixation (F_{ST}) and genetic differentiation (D_{est}; Jost, 2008) between each population were compared with 100 random permutations and 1000 resamplings, respectively, to assess statistical significance. Values of D_{est} were calculated using DEMEtics (Gerlach *et al.*, 2010).

The extent of genetic admixture between populations was inferred using the program STRUCTURE v.2.3.4 (Pritchard, Stephens & Donnelly, 2000). STRUCTURE groups individuals into genetic clusters based on allele frequency and then calculates the probability of membership of each cluster for each individual. If an individual's genotype appears to belong to more than one cluster, this indicates admixture between populations (Pritchard, Stephens & Donnelly, 2000). We provided

a population count estimate of K = 2–4, where K is the potential number of genetic clusters, and then ran a 100 000 step Markov chain Monte Carlo, following an initial burn-in of 50 000. STRUCTURE can be calibrated with known population information to assign individuals more accurately (Pritchard, Stephens & Donnelly, 2000). STRUCTURE uses prior population information sets only if location is correlated with genetic information (Pritchard, Wen & Falush, 2010). We ran two tests: one in which we estimated admixture (α) using prior population information, and another with no set priors (the basic model). α values range from 0–10, with higher values indicating greater degrees of shared ancestry between populations. We ran ten simulations for each test and used the outputs to generate an average for each test type, which we compiled and visualized using CLUMPAK (Kopelman *et al.*, 2015). CLUMPAK presents a summary estimate of the vector Q, which represents the proportion of each individual's genome assigned to each cluster across runs (Pritchard, Stephens & Donnelly, 2000).

We then used the program BOTTLENECK v.1.2.02 (Luikart & Cornuet, 1999) to detect any evidence for a recent bottleneck within each of our three supported populations. Recent bottlenecks (within $2N_e$–$4N_e$ generations) exhibit reduced N_a and heterozygosity. Heterozygosity can refer to either the expected level of heterozygosity based on the relative frequency of alleles (H_e) or the level of heterozygosity at mutation-drift equilibrium (H_{eq}), which is calculated using N_a. During a population bottleneck, rare alleles are more readily removed from a population on average, causing a relatively rapid loss of N_a compared with H_e. It is thus expected that $H_e > H_{eq}$ if a population bottleneck is to have reduced N_e within the last two to four reproductive generations (Luikart & Cornuet, 1999). This can be visualized using the Luikart *et al.* (1998) method, which bins allelic frequencies into ten frequency classes (0–0.1, 0.1–0.2 … 0.9–1.0). As is recommended for microsatellite loci, we used the two-phase model (TPM) in BOTTLENECK under an assumption of 95% single-step mutations, 5% multi-step mutations, and a variance of 12 among multi-step mutations. We then assessed the degree to which $H_e = H_{eq}$ using Wilcoxon's test with 1000 replications.

RESULTS

Microsatellite descriptive statistics

Eight microsatellite loci were successfully amplified across individuals from 57 nests collected across *C. australensis*' range (Table 1). The number of total alleles per locus (N_a) ranged from three to ten (mean = 5.13 ± 0.811), and the number of effective alleles per locus (A_e) ranged from 1.26 to 2.72 (mean = 1.8 ± 0.168; Table 1). Assessing across populations, three loci were found to deviate from HWE. However, within populations no loci deviated significantly from HWE after Bonferroni correction (Supporting Information, Table S2), nor were any loci in linkage disequilibrium.

Total heterozygosity, N_a and A_e were highest in QLD and lowest in SA (Supporting Information, Table S2, Fig. S1). The value of N_a was not statistically different between populations (Kruskal–Wallis: $\chi^2 = 3.75$, d.f. = 2, $P = 0.15$). Allelic diversity in SA was significantly lower than in both QLD (Kruskal–Wallis: $\chi^2 = 9.93$, d.f. = 2, $P = 0.01$; Conover–Iman $T = 3.68$, $P = 0.002$) and VIC (Conover–Iman $T = -3.18$, $P = 0.007$) but did not differ between QLD and VIC (Conover–Iman $T = 0.493$, $P = 0.94$). The value of H_o in SA was significantly lower than in both QLD (Kruskal–Wallis: $\chi^2 = 10.7$, d.f. = 2, $P < 0.0001$; Conover–Iman $T = 4.17$, $P < 0.001$) and VIC (Conover–Iman $T = -2.92$, $P = 0.012$), but did not differ between QLD and VIC (Conover–Iman $T = 1.25$, $P = 0.34$). The same pattern was observed for H_e

Table 1. Microsatellite loci descriptions across populations

Locus	N	N_a	A_e	H_o	H_e	HWE P	Bonferroni α
Caust35	56	3.000	1.658	0.375	0.397	0.350	0.006
Caust50	57	4.000	1.261	0.158	0.207	0.099	0.007
Caust42	57	4.000	1.610	0.333	0.379	0.239	0.008
Caust43	56	10.000	2.715	0.518	0.632	0.001	**0.010**
Caust44	57	7.000	2.317	0.561	0.568	0.036	0.013
Caust11	55	4.000	1.517	0.164	0.341	0.000	**0.017**
Caust21	56	4.000	1.622	0.232	0.383	0.002	**0.025**
Caust29	57	5.000	1.709	0.404	0.415	0.422	0.050
Mean	56.375	5.125	1.801	0.343	0.415	**< 0.0001**	–
SEM	0.263	0.811	0.168	0.054	0.047	–	–

A_e, number of effective alleles considering allele frequency; Bonferroni α, level of significance after Bonferroni correction for multiple comparisons; H_e, expected heterozygosity; H_o, observed heterozygosity; HWE P, probability of deviation from Hardy–Weinberg equilibrium; N, number of individuals; N_a, number of alleles. Values significant after Bonferroni correction are indicated in bold.

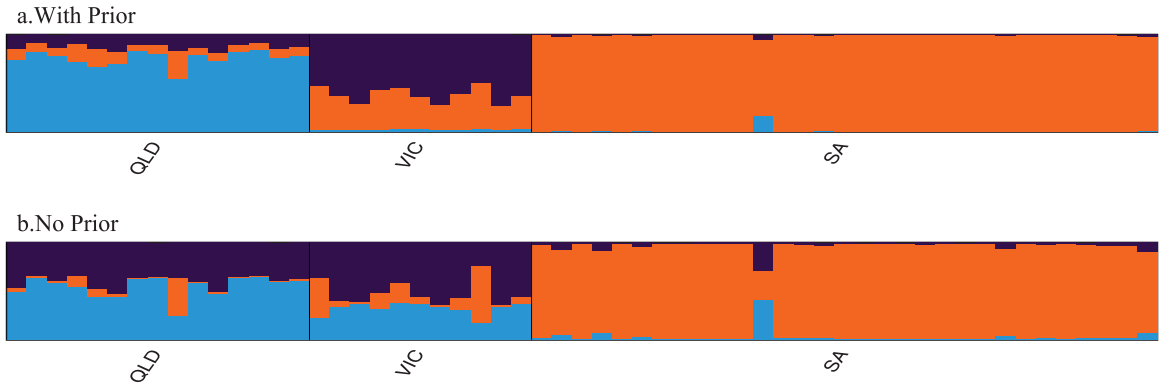

Figure 2. STRUCTURE output of $K = 3$ clusters represented by three colours (blue, QLD; purple, VIC; orange, SA) and frequency distribution: A, using a priori population information; and B, with no a priori information. Individual bands represent individual genotypes, and the proportion of colour within each band is proportional to the likelihood of belonging to each cluster.

(Kruskal–Wallis: $\chi^2 = 9.93$, d.f. = 2, $P = 0.01$; SA–QLD, Conover–Iman $T = 3.68$, $P = 0.002$; SA–VIC, Conover–Iman $T = -3.18$, $P = 0.007$; QLD–VIC, Conover–Iman $T = 0.493$, $P = 0.94$).

Population structure

AMOVA revealed significant population substructure between all three sampled populations (Supporting Information, Table S3). Global F_{ST} (0.237) was significant ($P < 0.0001$), whereas the global inbreeding coefficient ($F_{IS} = 0.016$) was not significant ($P = 0.36$). Although not significantly different from expected, local inbreeding coefficients increased from QLD ($F_{IS} = -0.076$, $P = 0.85$) to VIC ($F_{IS} = 0.059$, $P = 0.27$) and was highest in SA ($F_{IS} = 0.105$, $P = 0.09$). All pairwise F_{ST} comparisons between each population were significant ($P < 0.0001$), as were all pairwise D_{est} values ($P = 0.003$; Table 2).

The results of STRUCTURE analyses are summarized in Fig. 2. Regardless of the test model specified, the most likely cluster count (K) was three. Individuals from SA formed a distinct cluster from QLD and VIC. The probability of shared ancestry between individuals from VIC and individuals from QLD or SA differed based on model specification. Calibration with population information resulted in a lower degree of admixture ($\alpha = 0.065$) and higher shared ancestry between VIC and SA. Calibration without prior population information resulted in more admixture ($\alpha = 2.83$) and higher shared ancestry between VIC and QLD.

We found no evidence for a population bottleneck based on excess heterozygosity for any of the three populations ($H_e > H_{eq}$; QLD, $P = 0.73$; VIC, $P = 0.99$; SA, $P = 0.98$). Both VIC and SA exhibited heterozygosity deficiencies ($H_e < H_{eq}$; QLD, $P = 0.32$; VIC, $P = 0.02$; SA, $P = 0.04$). Furthermore, BOTTLENECK revealed a population-specific shift in allele frequencies towards the 10–20% class in SA compared with QLD and VIC, as well as a loss of intermediately occurring alleles (i.e. 40–70% frequency classes; Supporting Information, Fig. S2). South Australia also showed a relative increase in the number of common alleles (i.e. 70–100% frequency classes).

DISCUSSION

Population genetics and marker comparisons

Our detection of a significant decrease in genetic diversity in SA corroborates previously determined mitochondrial homogeneity for this population (Dew *et al.*, 2016). Pairwise F_{ST} values are often much higher when calculated using mitochondrial marker data compared with microsatellite data (Carlsson *et al.*, 2004; Goropashnaya *et al.*, 2007) owing to both the high variability of microsatellites and the lower effective population size of mtDNA (Jorde *et al.*, 1998). It is thus not unexpected that the F_{ST} values we determined through targeting microsatellite loci were lower than those secured using mtDNA (Dew *et al.*, 2016; Table 2).

There was a greater degree of fixation between SA and VIC compared with QLD and VIC using both mtDNA and nucDNA (Table 2). It is likely that recent and relatively isolated founding of SA has led to increased genetic differences between this population relative to range-wide diversity. Microsatellite F_{ST} values mirrored mtDNA F_{ST} except for greater fixation of mitochondrial sequences between VIC and SA ($F_{ST} = 0.59$) compared with QLD and SA ($F_{ST} = 0.43$). If females were the primary dispersing sex, or dispersed at an equal rate with males, we would predict that variation in maternally inherited loci between populations (i.e. mtDNA F_{ST}) would be consistent with loci inherited from both parents (i.e. nucDNA F_{ST}), which it is not. We would also expect

Table 2. Pairwise F_{ST} and D_{est} values among studied populations

	mtDNA F_{ST}	MSAT F_{ST}	MSAT D_{est}
QLD–SA	0.43	0.33	0.26
QLD–VIC	0.36	0.13	0.15
VIC–SA	0.59	0.18	0.10

Genetic fixation (F_{ST}) values from a previous study using mitochondrial *CO1* variation and from the present study, with genetic differentiation (D_{est}) values from the present study. All F_{ST} values are significant at $P < 0.0001$, and D_{est} values are significant at $P = 0.003$. n mtDNA = (QLD, 30; VIC, 42; SA, 19); n MSAT = (QLD, 15; VIC, 11; SA, 31). The F_{ST} values measure fixation, whereas D_{est} values measure differentiation (Jost, 2008). Microsatellite and mtDNA fixation are both greater between VIC and SA compared with QLD and VIC; however, microsatellite fixation is greatest between QLD and SA, whereas mtDNA fixation is greatest between VIC and SA. Microsatellite differentiation is greatest between QLD and SA and lowest between VIC and SA. The contrast between F_{ST} and D_{est} demonstrates that although VIC and SA are characterized by varying levels of heterozygosity, these two populations share a high proportion of similar alleles.

consistent differences between populations, regardless of marker, if the differences between nucDNA and mtDNA were attributable solely to the more extreme effects of drift acting on mtDNA (Jorde *et al.*, 1998). In this scenario, mitochondrial alleles may not be moving between source and sink populations as quickly as are nuclear alleles. As females must begin nesting at the start of the reproductive period, their ability to disperse might be limited compared with males, which do not help in nest construction (Rehan *et al.*, 2010).

Furthermore, our results support SA as a likely range edge for *C. australensis*. This species is thought to have dispersed out of Asia, moved south and east along Australia's coast, and then headed south and west through VIC to SA (Dew *et al.*, 2016). Ongoing migration into SA is probably very limited; although no bottleneck was detected there (Supporting Information, Fig. S2) genetic diversity remains very low compared with the other populations (Supporting Information, Fig. S1). Genetic differentiation (D_{est}) values were lowest between VIC and SA, demonstrating that these populations have a similar allelic make-up despite the severe genetic homogeneity of SA. Although habitat fragmentation has probably restricted migration into both VIC and SA, the allelic diversity of VIC suggests that this population has been in place long enough to accumulate rare alleles, whereas the population in SA has not. It is unlikely that our SA population is restricted to the coastal dunes where they were collected, because historically *C. australensis* have been collected farther inland around Adelaide (South Australia Museum). There is a chance that our results might be attributable to sampling bias if *C. australensis* tend to nest in aggregations and the coastal dunes represent one such large aggregation of philopatric females. Additional sampling around the Adelaide area could help to address this uncertainty.

ECOLOGY AND DISPERSAL HISTORY

By 1987, an estimated 95% of the native vegetation within the MDRB had been replaced with crop and non-native plant species (Sivertsen & Metcalfe, 1995). Current estimates suggest that ~69% of the MDRB area remains devoted to grazing pasture land (ABS/ABARE/BRS, 2009). When Michener (1962) collected 52 *C. australensis* nests from Queensland, New South Wales and the island of Guinea, he found only one nest within a native plant species. He subsequently hypothesized that the introduction of non-native plants had helped to facilitate expansion of *C. australensis* into Australia. Accordingly, *C. australensis* are now frequently found nesting in giant fennel (*F. communis*) surrounding Warwick in QLD (Rehan *et al.*, 2010, 2011) and in European searocket (*C. maritima*) in SA (Dew *et al.*, 2018). *Ferula communis* is a non-native species originally from the Mediterranean and Africa, and *C. maritima* is a native of Europe and the Mediterranean. *Cakile maritima*, which was first recorded in Western Australia in 1897, spread to SA by 1918 (Cody & Cody, 2004) and is now a part of the climax community of the dunes there (Lock & Cordingley, 2008). Non-native plant materials have also been implicated in the spread of *Ceratina* invasions into Hawaii (Shell & Rehan, 2017), and congeners in North America nest in plants associated with habitat edges and agriculture (Rehan & Richards, 2010).

The results of our population structural analyses suggest that the MDRB could explain the current distribution of *C. australensis* from northern Queensland (Michener, 1962) south and west to Adelaide, SA (Dew *et al.*, 2016; Fig. 1). Specifically, the MDRB might act as a natural migration corridor through the more arid regions of inland Australia. Our STRUCTURE results and D_{est} values support this theory by indicating that bees are dispersing from VIC into SA. Although it is possible that *C. australensis* might be migrating along the coast of south-eastern Australia into SA, there are no published records of populations or individuals south of the MDRB. A thorough analysis of the nesting preference of *C. australensis* for native vs. non-native plants could better inform our understanding of the dispersal requirements for this species. Future investigations could also strive to detect and sample additional populations along and outside of the MDRB to test our river-facilitated dispersal hypothesis.

ACKNOWLEDGEMENTS

We thank Rebecca Dew, Sarah Lawson, Sean Lombard and Sean Rodriguez for help with collecting bee nests and Michael Steffen for assistance with developing microsatellite markers. We thank three anonymous

reviewers for their helpful comments. Funding was provided by National Science Foundation (IOS-1456296) and National Geographic (9659-15) grants to S.M.R.

REFERENCES

ABS/ABARE/BRS. 2009. *Socio-economic context for the Murray-Darling Basin – descriptive report*. Canberra: ABS/ABARE/BRS Report to the Murray-Darling Basin Authority, 158.

Ambrose L, Cooper RD, Russell TL, Burkot TR, Lobo NF, Collins FH, Hii J, Beebe NW. 2014. Microsatellite and mitochondrial markers reveal strong gene flow barriers for *Anopheles farauti* in the Solomon Archipelago: implications for malaria vector control. *International Journal for Parasitology* **44:** 225–233.

Arca, M, Mougel F, Guillemaud T, Dupas S, Rome Q, Perrard A, Muller F, Fossoud A, Capdevielle-Dulac C, Torres-Leguizamon M, Chen XX, Tan JL, Jung C, Villemant C, Arnold G., Silvain JF. 2015. Reconstructing the invasion and the demographic history of the yellow-legged hornet, *Vespa velutina*, in Europe. *Biological Invasions* **17:** 2357–2371.

Asplen MK, Whitfield JB, De Boer JG, Heimpel GE. 2009. Ancestral state reconstruction analysis of hymenopteran sex determination mechanisms. *Journal of Evolutionary Biology* **22:** 1762–1769.

Baines CB, Ferzoco IM, McCauley SJ. 2017. Sex-biased dispersal is independent of sex ratio in a semiaquatic insect. *Behavioral Ecology and Sociobiology* **71:** 119.

Boomsma JJ, Baer B, Heinze J. 2005. The evolution of male traits in social insects. *Annual Review of Entomology* **50:** 395–420.

Carlsson J, McDowell JR, Díaz-Jaimes P, Carlsson JE, Boles SB, Gold JR, Graves JE. 2004. Microsatellite and mitochondrial DNA analyses of Atlantic bluefin tuna (*Thunnus thynnus thynnus*) population structure in the Mediterranean Sea. *Molecular Ecology* **13:** 3345–3356.

Cody ML, Cody TWD. 2004. Morphology and spatial distribution of alien sea-rockets (*Cakile* spp.) on South Australian and Western Canadian beaches. *Australian Journal of Botany* **52:** 175–183.

Conover WJ, Iman RL. 1979. On multiple-comparisons procedures. *Los Alamos Sci. Lab. Tech. Rep. LA-7677-MS*. 1–14.

Cornuet JM, Luikart G. 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144:** 2001–2014.

Cortázar-Chinarro M, Lattenkamp EZ, Meyer-Lucht Y, Luquet E, Laurila A, Höglund J. 2017. Drift, selection, or migration? Processes affecting genetic differentiation and variation along a latitudinal gradient in an amphibian. *BMC Evolutionary Biology* **17:** 189.

Dew RM, Rehan SM, Schwarz MP. 2016. Biogeography and demography of an Australian native bee *Ceratina australensis* (Hymenoptera, Apidae) since the last glacial maximum. *Journal of Hymenoptera Research* **49:** 25–41.

Dew RM, Shell WA, Rehan SM. 2018. Changes in maternal investment with climate moderate social behaviour in a facultatively social bee. *Behavioral Ecology and Sociobiology* **72:** 69.

Eckert CG, Samis KE, Lougheed SC. 2008. Genetic variation across species' geographical ranges: the central–marginal hypothesis and beyond. *Molecular Ecology* **17:** 1170–1188.

Excoffier L, Laval G, Schneider S. 2005. Arlequin (version 3. 0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1:** 47–50.

Friedel A, Paxton RJ, Soro A. 2017. Spatial patterns of relatedness within nesting aggregations of the primitively eusocial sweat bee *Lasioglossum malachurum*. *Insectes Sociaux* **64:** 465–475.

Gerlach G, Jueterbock A, Kraemer P, Deppermann J, Harmand P. 2010. Calculations of population differentiation based on G_{ST} and D: forget G_{ST} but not all of statistics! *Molecular Ecology* **19:** 3845–3852.

Goropashnaya AV, Fedorov VB, Seifert B, Pamilo P. 2007. Phylogeography and population structure in the ant *Formica exsecta* (Hymenoptera, Formicidae) across Eurasia as reflected by mitochondrial DNA variation and microsatellites. *Annales Zoologici Fennici* **44:** 462–474.

Goudet J, Perrin N, Waser P. 2002. Tests for sex-biased dispersal using bi-parentally inherited genetic markers. *Molecular Ecology* **11:** 1103–1114.

Hoffman AA, Blows MW. 1994. Species borders: ecologial and evolutionary perspectives. *Trends in Ecology & Evolution* **9:** 223–227.

Johnstone RA, Cant MA, Field J. 2012. Sex-biased dispersal, haplodiploidy and the evolution of helping in social insects. *Proceedings of the Royal Society B: Biological Sciences* **279:** 787–793.

Jongepier E, Foitzik S. 2016. Ant recognition cue diversity is higher in the presence of slavemaker ants. *Behavioral Ecology* **27:** 304–311.

Jorde LB, Bamshad M, Rogers AR. 1998. Using mitochondrial and nuclear DNA markers to reconstruct human evolution. *Bioessays: News and Reviews in Molecular, Cellular and Developmental Biology* **20:** 126–136.

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

Kirby KS. 1965. Isolation and characterization of ribosomal ribonucleic acid. *The Biochemical Journal* **96:** 266–269.

Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I. 2015. CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across *K*. *Molecular Ecology Resources* **15:** 1179–1191.

Kukuk PF, Decelles PC. 1986. Behavioral evidence for population structure in *Lasioglossum (Dialictus) zephyrum* female dispersion patterns. *Behavioral Ecology and Sociobiology* **19:** 233–239.

Lin N, Michener CD. 1972. Evolution of sociality in insects. *The Quarterly Review of Biology* **47:** 131–159.

Lock CE, Cordingley SP. 2008. *Vegetation Management Plan West Beach Dune Reserve, vol. 129.* SA Urban Forest Biodiversity Program.

López-Uribe MM, Cane JH, Minckley RL, Danforth BN. 2016. Crop domestication facilitated rapid geographical expansion of a specialist pollinator, the squash bee *Peponapis pruinosa*. *Proceedings of the Royal Society B: Biological Sciences* **283:** 20160443.

López-Uribe MM, Morreale SJ, Santiago CK, Danforth BN. 2015. Nest suitability, fine-scale population structure and male-mediated dispersal of a solitary ground nesting bee in an urban landscape. *PLoS One* **10:** e0125719.

López-Uribe MM, Zamudio KR, Cardoso CF, Danforth BN. 2014. Climate, physiological tolerance and sex-biased dispersal shape genetic structure of Neotropical orchid bees. *Molecular Ecology* **23:** 1874–1890.

Luikart G, Allendorf FW, Cornuet JM, Sherwin WB. 1998. Distortion of allele frequency distributions provides a test for recent population bottlenecks. *The Journal of Heredity* **89:** 238–247.

Luikart PS, Cornuet JM. 1999. BOTTLENECK: a computer program for detecting recent reduction in the effective population size using allele frequency data. *Journal of Heredity* **90:** 502–503.

Michener CD. 1962. The genus *Ceratina* in Australia, with notes on its nests (Hymenoptera: Apoidea). *Journal of the Kansas Entomological Society* **35:** 414–421.

Michener CD. 1974. *The social behavior of the bees: a comparative study*. Cambridge, MA: Harvard University Press.

Michener CD. 2007. *The bees of the world*, *2nd edn*. Baltimore, MD: The John Hopkins University Press.

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

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155:** 945–959.

Pritchard JK, Wen X, Falush D. 2010. *Documentation for structure software: version 2.3*. Available at: http://pritch.bsd.uchicago.edu/software/readme_structure2.pdf (accessed 23 May 2018).

Pusey AE. 1987. Sex-biased dispersal and inbreeding avoidance in birds and mammals. *Trends in Ecology & Evolution* **2:** 295–299.

R Core Team. 2017. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at: http//www.R-project.org/ (accessed 13 March 2018).

Rehan SM, Richards MH. 2010. Nesting biology and subsociality in *Ceratina calcarata* (Hymenoptera: Apidae). *Canadian Entomologist* **142:** 65–74.

Rehan SM, Richards MH, Adams M, Schwarz MP. 2014. The costs and benefits of sociality in a facultatively social bee. *Animal Behaviour* **97:** 77–85.

Rehan SM, Richards MH, Schwarz MP. 2010. Social polymorphism in the Australian small carpenter bee, *Ceratina (Neoceratina) australensis*. *Insectes Sociaux* **57:** 403–412.

Rehan S, Schwarz M. 2015. A few steps forward and no steps back: long-distance dispersal patterns in small carpenter bees suggest major barriers to back-dispersal. *Journal of Biogeography* **42:** 485–494.

Rehan SM, Schwarz MP, Richards MH. 2011. Fitness consequences of ecological constraints and implications for the evolution of sociality in an incipiently social bee. *Biological Journal of the Linnean Society* **103:** 57–67.

Ross KG, Shoemaker DD. 1997. Nuclear and mitochondrial genetic structure in two social forms of the fire ant *Solenopsis invicta*: insights into transitions to an alternate social organization. *Heredity* **78:** 590–602.

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

Schmid-Hempel P, Schmid-Hempel R, Brunner PC, Seeman OD, Allen GR. 2007. Invasion success of the bumblebee, *Bombus terrestris*, despite a drastic genetic bottleneck. *Heredity* **99:** 414–422.

Schuelke, M. 2000. An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechology* **18:** 233–234.

Shell WA, Rehan SM. 2016. Development of multiple polymorphic microsatellite markers for *Ceratina calcarata* (Hymenoptera: Apidae) using genome-wide analysis. *Journal of Insect Science* **16:** 1–4.

Shell WA, Rehan SM. 2017. Range expansion of the small carpenter bee *Ceratina smaragdula* across the Hawaiian Archipelago with potential ecological implications for native pollinator systems. *Pacific Science* **71:** 1–15.

Sivertsen D, Metcalfe L. 1995. Natural vegetation of the southern wheat-belt (Forbes and Cargelligo 1:250 000 map sheets). *Cunninghamia* **4:** 103–128.

Soro A, Quezada-Euan JJG, Theodorou P, Moritz RFA, Paxton RJ. 2017. The population genetics of two orchid bees suggests high dispersal, low diploid male production and only an effect of island isolation in lowering genetic diversity. *Conservation Genetics* **18:** 607–619.

Tamang R, Chaubey G, Nandan A, Govindaraj P, Singh VK, Rai N, Mallick CB, Sharma V, Sharma VK, Shah AM, Lalremruata A, Reddy AG, Rani DS, Doviah P, Negi N, Hadid Y, Pande V, Vishnupriya S, van Driem G, Behar DM, Sharma T, Singh L, Villems R, Thangaraj K. 2018. Reconstructing the demographic history of the Himalayan and adjoining populations. *Human Genetics* **137:** 129–139.

Tsutsui ND, Suarez AV. 2003. The colony structure and population biology of invasive ants. *Conservation Biology* **17:** 48–58.

Ulrich Y, Perrin N, Chapuisat M. 2009. Flexible social organization and high incidence of drifting in the sweat bee, *Halictus scabiosae*. *Molecular Ecology* **18:** 1791–1800.

Yanega, D. 1990. Philopatry and nest founding in a primitively social bee, *Halictus rubicundus*. *Behavioral Ecology and Sociobiology* **27:** 37–42.

Zayed A, Packer L. 2005. Complementary sex determination substantially increases extinction proneness of haplodiploid populations. *Proceedings of the National Academy of Sciences of the United States of America* **102:** 10742–10746.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Figure S1. Microsatellite summary characteristics averaged across loci for each population. Characteristics were compared using the Conover–Iman test for multiple comparisons. A_e, number of effective alleles considering allele frequency; H_e, expected heterozygosity; H_o, observed heterozygosity; N_a, number of alleles. Asterisks denote comparisons between SA only, as no significance was detected between QLD and VIC. $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

Figure S2. Allele distribution histograms for each population. Bottlenecks are detectable if the frequency of rare alleles (0–0.1) is less than the frequency of alleles occurring at a higher frequency (> 0.1).

Table S1. Primer sequences and locus characteristics of *Ceratina australensis* microsatellite loci. T_a, primer-specific annealing temperature (in degrees Celsius).

Table S2. Summary characteristics of microsatellite markers averaged across loci. A_e, number of effective alleles considering allele frequency; Bonferroni α, level of significance after Bonferroni correction for multiple comparisons, with significant deviations from HWE after correction in bold; H_e, expected heterozygosity; H_o, observed heterozygosity; HWE *P*, probability of deviation from Hardy–Weinberg equilibrium; *N*, number of individuals; N_a, number of alleles. χ^2 test statistics are reported for each characteristic. d.f. = 2 for all comparisons.

Table S3. Results of AMOVA comparing allelic frequencies among 57 non-related females from the three populations of *Ceratina australensis*.