

# Biogeography and demography of an Australian native bee *Ceratina australensis* (Hymenoptera, Apidae) since the last glacial maximum

Rebecca M. Dew<sup>1</sup>, Sandra M. Rehan<sup>2</sup>, Michael P. Schwarz<sup>1</sup>

<sup>1</sup> Biological Sciences, Flinders University of South Australia, GPO Box 2100, Adelaide 5001 <sup>2</sup> 46 College Road, Department of Biological Sciences, University of New Hampshire, Durham, NH, USA 03824

Corresponding author: Rebecca M. Dew ([dew0009@flinders.edu.au](mailto:dew0009@flinders.edu.au))

Academic editor: J. Neff | Received 9 February 2016 | Accepted 30 March 2016 | Published 28 April 2016

<http://zoobank.org/C3C89FA2-7DB9-4430-8924-80C4381CD84A>

**Citation:** 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. doi: [10.3897/JHR.49.8066](https://doi.org/10.3897/JHR.49.8066)

## Abstract

The small carpenter bees, genus *Ceratina*, are highly diverse, globally distributed, and comprise the sole genus in the tribe Ceratinini. Despite the diversity of the subgenus *Neoceratina* in the Oriental and Indo-Malayan region, *Ceratina* (*Neoceratina*) *australensis* is the only ceratinine species in Australia. We examine the biogeography and demography of *C. australensis* using haplotype variation at 677 bp of the barcoding region of COI for specimens sampled from four populations within Australia, across Queensland, New South Wales, Victoria and South Australia. There is geographic population structure in haplotypes, suggesting an origin in the northeastern populations, spreading to southern Australia. Bayesian Skyline Plot analyses indicate that population size began to increase approximately 18,000 years ago, roughly corresponding to the end of the last glacial maximum. Population expansion then began to plateau approximately 6,000 years ago, which may correspond to a slowing or plateauing in global temperatures for the current interglacial period. The distribution of *C. australensis* covers a surprisingly wide range of habitats, ranging from wet subtropical forests through semi-arid scrub to southern temperate coastal dunes. The ability of small carpenter bees to occupy diverse habitats in ever changing climates makes them a key species for understanding native bee diversity and response to climate change.

## Keywords

Climate change, dispersal, DNA barcoding, bayesian skyline plot, haplotype network, population structure

## Introduction

Molecular studies have greatly increased our understanding of the antiquity of bees and their historical biogeography, especially with respect to centres of origin and subsequent dispersal routes (e.g. Fuller et al. 2005; Hines 2008; Chenoweth and Schwarz 2011; Rehan and Schwarz 2015). Other studies using museum collection data have implicated very recent climate change as a possible factor underlying changes in bee abundances (e.g. Cameron et al. 2011; Burkle et al. 2013), but there are surprisingly few studies that have attempted to infer changes in bee abundance beyond the last 200 years (but see Wilson et al. 2014). In the face of likely future climate change, it is important to understand how bees have responded to past climates so that we may better predict future trends.

Two recent studies (Groom et al. 2014a; López-Urbe et al. 2014) have used phylogeographic and coalescent Bayesian Skyline Plot analyses to examine changes in bee abundances for tropical halictine (Halictidae) and euglossine (Apidae) bees respectively. Both studies found a strong response to Pleistocene climates, suggesting that these two faunal groups have been impacted by glacial cycles despite their tropical distribution. Small carpenter bees, *Ceratina* (Apidae: Ceratinini), of the subgenus *Zadontomerus* in eastern North America also showed a rapid population expansion approximately 20 kya, linked to post-glacial cycles (Shell and Rehan 2016). However, no studies have examined possible impacts of historical climates on bee species spanning temperate to xeric zones, beyond those using museum records.

The small carpenter bee genus *Ceratina* has a nearly global distribution, occurring on all continents except Antarctica (Rehan and Schwarz 2015). This includes recently colonized remote islands in the Southwestern Pacific and southern Indian Ocean, probably via accidental human agency (Rehan et al. 2010a; Groom et al. 2014b). *Ceratina* originated in Africa in the late Paleocene or early Eocene and showed rapid long distance dispersal events allowing it to eventually colonize the New World by the late Eocene (Rehan and Schwarz 2015). Interestingly, patterns of radiation in this tribe suggest that major long distance dispersal events have been rare and tended to occur more frequently in the early history of this tribe rather than later on, despite there being few geographical barriers to later dispersal events (Rehan and Schwarz 2015). Physical impediments to dispersal, such as water barriers, are actually believed to have decreased in the more recent history of this tribe (Rehan and Schwarz 2015).

*Ceratina* (*Neoceratina*) is the sister clade to all other *Ceratina* subgenera and originated from a dispersal from Africa to the Oriental region in the early Eocene, with some species extending into the Palearctic (Rehan and Schwarz 2015). Surprisingly, there is only a single ceratinine species in Australia, *Ceratina* (*Neoceratina*) *australensis* (Perkins, 1912). This species forms the sister lineage to all other *C. (Neoceratina)* species, and its stem age is dated to the middle Eocene. *Ceratina australensis* has become a model species for understanding simple forms of sociality where both solitary and social forms remain in sympatry (e.g. Rehan et al. 2010b, 2011, 2014). Solitary nests comprise about eighty-five percent of the population and are founded by females that

disperse from their natal nests (Rehan et al. 2014). Dispersal of females could facilitate gene flow between populations across Australia.

Michener (1962) recorded *Ceratina australensis* from subtropical and temperate regions of eastern Queensland and New South Wales but there have been no further studies of its distribution. Here we use haplotype variation at 677 bp of the mitochondrial 'barcoding' region of cytochrome *c* oxidase subunit 1 (CO1) from 102 specimens of *C. australensis*, obtained from southern Queensland, mid-New South Wales, northern Victoria and southern South Australia, to examine historical demography and geographical patterns in population genetic structure. Based on the dispersal of females from natal nests we predicted that there would be gene flow between populations, influencing the genetic structure and historical demography of the species.

## Methods

### Collecting localities for genetic samples

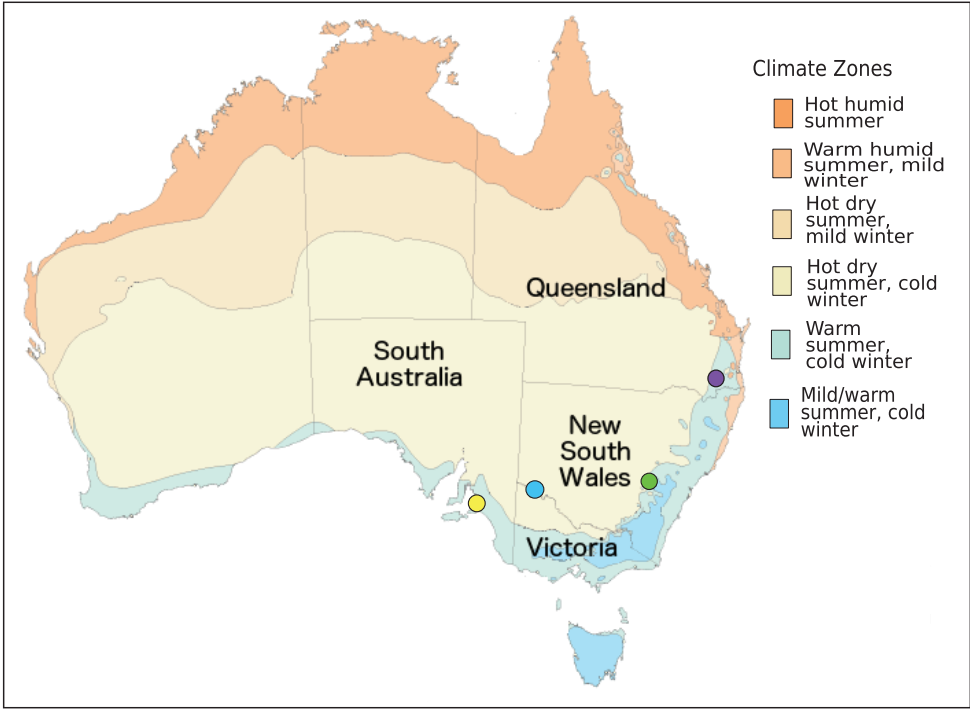
Specimens of *Ceratina australensis* were collected from four populations: (i) Mildura in northwestern Victoria; (ii) West Beach in metropolitan Adelaide, South Australia; (iii) the Cowra region in central New South Wales; and (iv) the Warwick region in southeastern Queensland (Figure 1). Our four study sites represent a substantial proportion of the geographic and climatic conditions for the species, covering warm temperate forest, semi-arid riverine woodland, mediterranean coastal dunes and central tablelands. GPS coordinates, collection dates and the number of barcoded specimens from each population are shown in Table 1. Nests, predominantly found in dead stems of plants from the genera *Cakile* (Brassicaceae), *Senecio* (Asteraceae), *Ferula* (Apiaceae) and the species *Verbena bonariensis* L. (Verbenaceae) were collected during early mornings or late evenings. This ensured that the bees would be present in the nest and not out foraging. Nests were collected, as this species is rarely observed on flowers (Michener 1962). Nests were stored on ice or at 4 °C until processing. One randomly chosen adult from each nest was stored in 100% ethanol for DNA sequencing.

### DNA extraction and sequencing techniques

One leg of each specimen was incubated overnight in arthropod lysis solution with proteinase K. Extractions proceeded using a glass fibre plate and a vacuum manifold to pull the eluates through the membrane, following the procedures detailed in Ivanova et al. (2006). The DNA extract was stored in 50 µl TLE (10mM TRIS, 0.1mM EDTA pH8). Forward and reverse primers M070 (5'-TCC AAT GCA CTA ATC TGC CAT ATT A-3') and M080 (5'-TAC AGT TGG AAT AGA CGT TGA TAC-3') were used for PCR amplification of a 700 bp region of CO1. We used 27.5 µl reactions with 0.1 µl immobilase as the active enzyme, 1 µl of each M070 and M080, 5 µl of MRT Buffer, 15.4 µl water

**Table 1.** Summary of samples collected from each population of *C. australensis*. Includes GPS coordinates, collection dates, total number of specimens sequenced and the number of unique haplotypes recovered.

Population	Latitude (S) / Longitude (E)	Collection Dates	Specimens Barcoded	Haplotypes
Cowra, New South Wales	33°52.78' / 148°45.73'	October 2015	11	8
Mildura, Victoria	34°09.25' / 142°09.58'	June 2013, October 2013, January 2014, April 2014	42	9
Warwick, Queensland	28°12.85' / 152°02.10'	January 2010	30	14
West Beach, South Australia	34°56.28' / 138°29.95'	June 2012, July 2014	19	1



**Figure 1.** Map of Australia with overlaid temperature and humidity climate zones (Bureau of Meteorology 2012) showing sampling locations. New South Wales, green; Queensland purple; Victoria, blue; South Australia, yellow.

and 5  $\mu$ l DNA. The PCR cycle began with 10 min of 94  $^{\circ}$ C. The annealing stage had 5 cycles consisting of 60 s at 94  $^{\circ}$ C, 90 s at 45  $^{\circ}$ C and 90 s at 72  $^{\circ}$ C followed by 35 cycles of 60 s at 94  $^{\circ}$ C, 90 s at 50  $^{\circ}$ C and 60 s at 72  $^{\circ}$ C. Elongation was 10 min at 72  $^{\circ}$ C with a final 2 min at 25  $^{\circ}$ C. The PCR products were cleaned using a vacuum plate with 100  $\mu$ l TLE, with the cleaned products stored in 30  $\mu$ l TLE. Final forward and reverse sequencing of the cleaned PCR products was performed by the Australian Genome Research Facility.

## Alignment and phylogenetic reconstruction

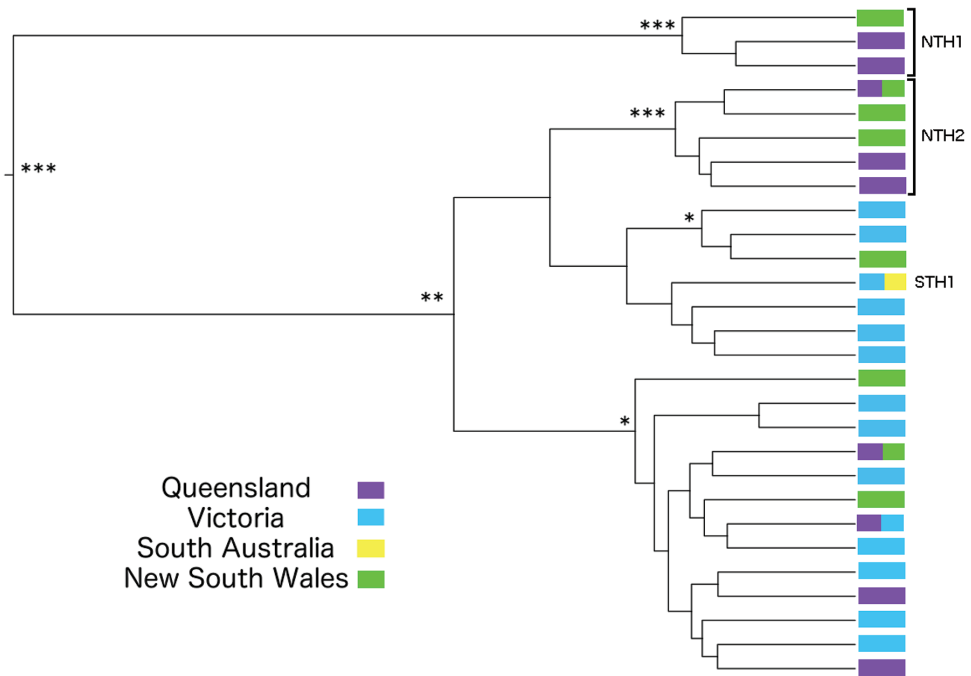
Sequences were imported into GENEIOUS v.6.1.6 (Kearse et al. 2012) for editing and alignment. Reverse and forward sequences were combined into a consensus sequence for each sample. As we were comparing individual base pair changes, no ambiguous or unknown base pairs (including at the end of sequences) could be left in the final alignment. We aimed to maximize both the sequence length and the number of samples. The sequence length was shortened, so that all samples had base pair data covering the same read length without any missing or ambiguous nucleotides, since missing data can lead to spurious results for coalescent analyses (Ho and Shapiro 2011). The final alignment consisted of 102 sequences of 677 bp in length, with 28 unique haplotypes. All edited sequences are submitted to GenBank (accession numbers KR824844-KR824934 and KU664337-KU664347).

An undescribed *Neoceratina* species from the Solomon Islands, *Ceratina* (*Neoceratina*) “Solomons sp.” was included in the alignment as the outgroup to determine the root of the tree. This species has been shown to be phylogenetically distinct to *Ceratina australensis* (Rehan and Schwarz 2015). The sequences available for this species did not cover the full length of the 677 bp alignment, so the *C. australensis* sequences were shortened to 639 bp for this analysis. This sequence attenuation did not remove any of the unique haplotype information present within the *C. australensis* sequences. The phylogenetic tree was reconstructed in BEAST v.1.8.1 (Drummond and Rambaut 2007) with a Yule Process tree prior. The substitution model HKY+I+G model was identified as the most appropriate based on Akaike information criterion in JMODELTEST (Guindon and Gascuel 2003; Darriba et al. 2012). The analysis ran for  $5 \times 10^7$  generations, logged every 1,000 trees, using a fixed mutation rate of 1.0, with all other parameters set to default. The log files were viewed in TRACER v.1.5 to confirm that the posterior had stabilized. A consensus tree was constructed with TREE ANNOTATER v.1.8.1 with a burn-in of 10,000 trees (i.e. 10 million iterations; Suppl. material 1).

The full alignment was then pruned to contain only unique haplotypes (28 sequences). The analysis was run following the conditions described above. TRACER again confirmed the posterior of the analysis had stabilised and a consensus tree with a burn-in of 10,000 trees was generated (Figure 2).

## Haplotype networks

Analysis of Molecular Variance (AMOVA) was implemented in ARLEQUIN 3.11 (Excoffier et al. 2005) to compare genetic variation within and among Victoria, New South Wales, South Australia and Queensland populations. For these analyses we used all sequences and included all codon positions. All four populations were compared in the full model followed by pair-wise comparisons of each possible pairing in subsequent AMOVA analyses. The full alignment was imported into NETWORK (Fluxus

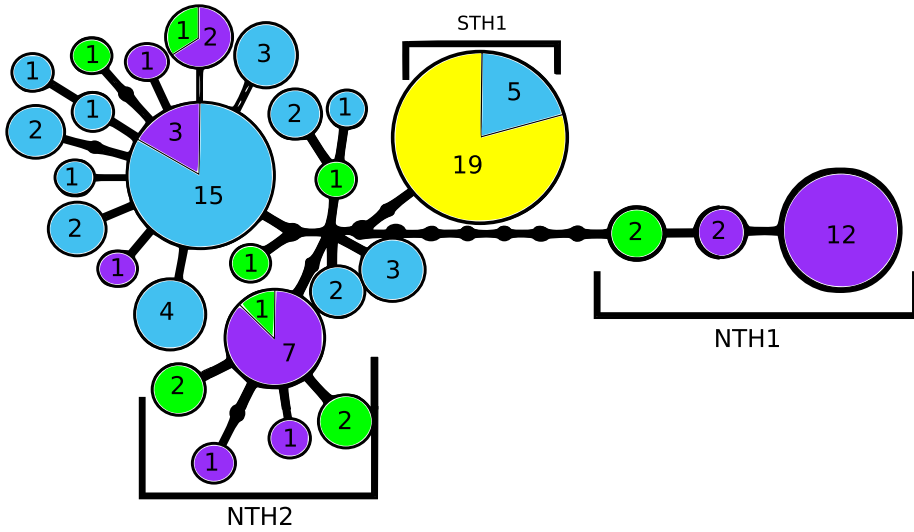


**Figure 2.** Maximum credibility tree of the *C. australensis* haplotypes from Bayesian BEAST analysis. Clades North 1 (NTH1), North 2 (NTH2) and South 1 (STH1) are indicated. Posterior probability values: \*\*\* = 1.0; \*\* ≥ 95; \* ≥ 85.

Engineering 2016) and a haplotype network was constructed using a median-joining analysis with epsilon set as zero (Bandelt et al. 1999; Figure 3). HAPLOVIEWER confirmed the final network and was used to generate a publication quality figure (Salzburger et al. 2011).

### Historical demography

We used Bayesian Skyline Plot (BSP) analyses implemented in BEAST and TRACER to explore historical changes in effective population size of *Ceratina australensis*. For these analyses we included all sequences available including duplicate haplotypes, as analyses of only unique haplotypes can give erroneous results (Grant 2015). BSP analyses assume that genetic markers evolve neutrally (Ho and Shapiro 2011). The very small number of amino acid changes in our alignment suggests that purifying selection may be operating on 1<sup>st</sup> and 2<sup>nd</sup> codon positions, so we restricted BSP analyses to 3<sup>rd</sup> codon positions only. In these analyses we used a GTR model for nucleotide substitutions, but did not include an invariant sites parameter (I) since 3<sup>rd</sup> codon positions should not be constrained by selection. Analyses were run for 100 million iterations, sampling every 10,000<sup>th</sup> iteration to reduce autocorrelation, and were repeated three



**Figure 3.** Haplotype network of *C. australensis* populations. Each circle represents a unique haplotype, with the numerals inside indicating the number of individuals sampled of that haplotype. Each step between haplotypes indicates one base pair substitution.

times to check for convergence. We implemented a strict molecular clock with rate of 1.0, which allows us to readily convert mutations per site per generation into chronological years.

We converted the Bayesian Skyline plot scale to chronological years through dividing it by mutation rate and the number of generations per year. We used the mitochondrial mutation rate observed in *Drosophila melanogaster* Meigen, viz.  $6.2 \times 10^{-8}$  mutations per site per generation as an estimate of the mutation rate for *Ceratina australensis* (Haag-Liautard et al. 2008). This method follows a previous study on demographic history in Fijian halictine bees in the genus *Lasioglossum* (Groom et al. 2014a) and North American *Ceratina* species (Shell and Rehan 2016). We note that the mitochondrial mutation rate for *Caenorhabditis elegans* (Maupas) is estimated as  $9.7 \times 10^{-8}$  mutations per site per generation, close to the rate for *D. melanogaster*, and that they have mitochondrial AT biases of 76% and 82% respectively. Previous studies have reported an AT bias of 74% for the same barcoding region as in our study (Groom et al. 2014a; Shell and Rehan 2016), and our *Ceratina* haplotypes had an AT bias of 78%. The number of generations was determined as two per year based on nest contents data from the Victorian and South Australian sites (Dew and Rehan, unpublished data), which also corresponds to detailed studies on the Queensland population (Rehan et al. 2010b, 2011, 2014).

In order to determine whether inferred changes in historical population size in the BSP analysis were significant we also carried out another coalescent analysis using the same parameter settings as our BSP analysis, but implementing a constant population size model. This was then compared to our BSP analysis using a Bayes Factor test.



## Inferring ancestral distributions

Ancestral distributions were inferred using BEAST ancestral traits reconstruction. The full alignment of 102 sequences was used. Sample location for each sequence was coded as a discrete trait (either New South Wales, Queensland, South Australia or Victoria). The analysis ran for  $2 \times 10^8$  generations, logged every 1,000, with a Yule process tree prior. A HKY+I+G site model with a strict clock of rate 1.0 was employed. All parameters for phylogeny and ancestral trait reconstruction had reached stability, as viewed in TRACER with a burnin of  $1 \times 10^8$  generations. Using this burnin a consensus tree was constructed in TREE ANNOTATER.

## Results

### Haplotype lineages

In total 28 haplotypes of *Ceratina australensis* were found across the four field sites. The haplotype tree along with posterior probability values for node support from our BEAST analysis is given in Figure 2. This analysis indicates the presence of a clade comprising one New South Wales and two Queensland haplotypes, which is highly supported (PP = 1.0) as sister clade to the remaining haplotypes. Our haplotype network analysis (Figure 3) indicates that this clade, which we will refer to as NTH1 (due to their relative northern location), is separated from the common ancestor for the remaining haplotypes by seven nucleotide substitutions, none of which involve amino acid changes.

Both the haplotype tree in Figure 2 and the haplotype network in Figure 3 indicate geographical structuring of haplotypes. Firstly, the haplotypes in NTH1 were not recovered in any of the 99 specimens genotyped from the more southwestern sampling locations of Victoria and South Australia. Secondly, there was another clade comprising five haplotypes that were only recovered from the more northeastern localities of Queensland and New South Wales, and we refer to this clade as NTH2. Lastly, the haplotype, which we refer to as STH1 (due to its southern location), mostly comprised specimens from South Australia, but also some from Victoria, and it was not represented in any of the Queensland or New South Wales specimens. Two haplotypes are shared between Queensland and New South Wales, with one shared between Queensland and Victoria.

### Population structure

Pair-wise comparisons among all individuals revealed significantly greater sequence divergence between populations than within populations for Victoria to New South Wales and Queensland, and for South Australia to all other populations (Table 2). Queensland



**Table 2.** *Ceratina australensis* regional population structure. Diagonal indicates average pairwise differences within populations, number in parentheses indicates total number of sequences for that region; above diagonal are average pairwise differences between populations; below diagonal are pairwise  $F_{ST}$  values. Significant values ( $p < 0.05$ ) indicated in bold.

Population structure	Queensland	New South Wales	Victoria	South Australia
Queensland	6.88736 (30)	6.5303 (0.49)	<b>6.93968 (&lt;0.0001)</b>	<b>6.5 (&lt;0.0001)</b>
New South Wales	0.0598 (0.19)	5.49091 (11)	<b>5.03247 (&lt;0.0001)</b>	<b>5.09091 (&lt;0.0001)</b>
Victoria	<b>0.35579 (&lt;0.0001)</b>	<b>0.26487 (&lt;0.0001)</b>	2.5331 (42)	<b>3.78571 (&lt;0.0001)</b>
South Australia	<b>0.42823 (&lt;0.0001)</b>	<b>0.55586 (&lt;0.0001)</b>	<b>0.58507 (&lt;0.0001)</b>	0 (19)

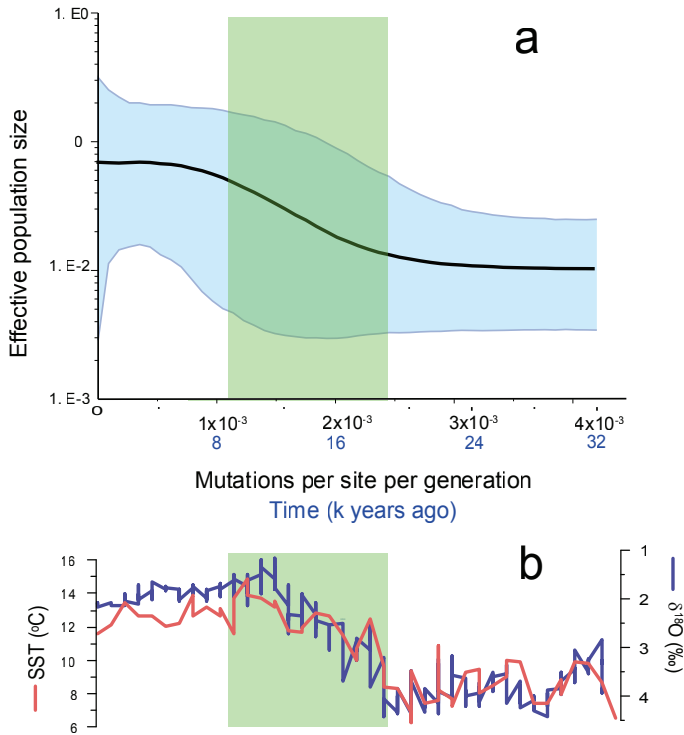
**Table 3.** Tajima's D and Fu's  $F_S$  tests of neutrality within populations. Segregating sites (S), Tajima's D score and significance value (D p-value), and Fu's  $F_S$  value and significance values ( $F_S$  p-value) are presented. Values in bold are statistically significant ( $p < 0.05$ ).

Neutrality tests	Queensland	New South Wales	Victoria	South Australia
S	19	16	16	0
Tajima's D	1.36657	0.02304	-1.01646	0
D p-value	0.937	0.558	0.186	1.000
Fu's $F_S$	<b>-25.15767</b>	<b>-6.57395</b>	<b>-26.633</b>	$3.4 \times 10^{38}$
$F_S$ p-value	<b>0</b>	<b>0.001</b>	<b>0</b>	1

and New South Wales were not significantly different from one another. These results are mirrored by pairwise  $F_{ST}$  calculations (Table 2), suggesting that all populations except New South Wales and Queensland are genetically differentiated. There was only one fixed base pair difference identified between any of the populations. This was at base 486, which was a thymine in all Queensland and New South Wales samples but an adenine in all South Australia samples (Victorian samples varied at this base). Tajima's D value indicated neutral evolution between populations (Table 3).

## Historical demography

Our Bayesian skyline plot (BSP) analysis is summarized in Figure 4 where it is temporally aligned with a graph summarizing two temperature proxies taken from Pahnke et al. (2003). In Figure 4 we have used two x-axis scales, one using mutations per site per generation and the other using years before present, based on a mutation rate of  $6.2 \times 10^{-8}$  mutations per site and two generations per year. Based on the current best estimate for mutation rate these plots suggest an increase in effective population size beginning approximately 20–18 ka, with a peak rate of increase at about 15–8 kya, and a plateauing after about 6 kya. Our BSP plot shows a long period of stasis from approximately 64–20 kya. This is an artifact of the analysis, where signals prior to the last major effective population size change are lost (Grant et al. 2012; Grant 2015). This period of stasis was trimmed in the final figure to show just the plot from 32 kya.

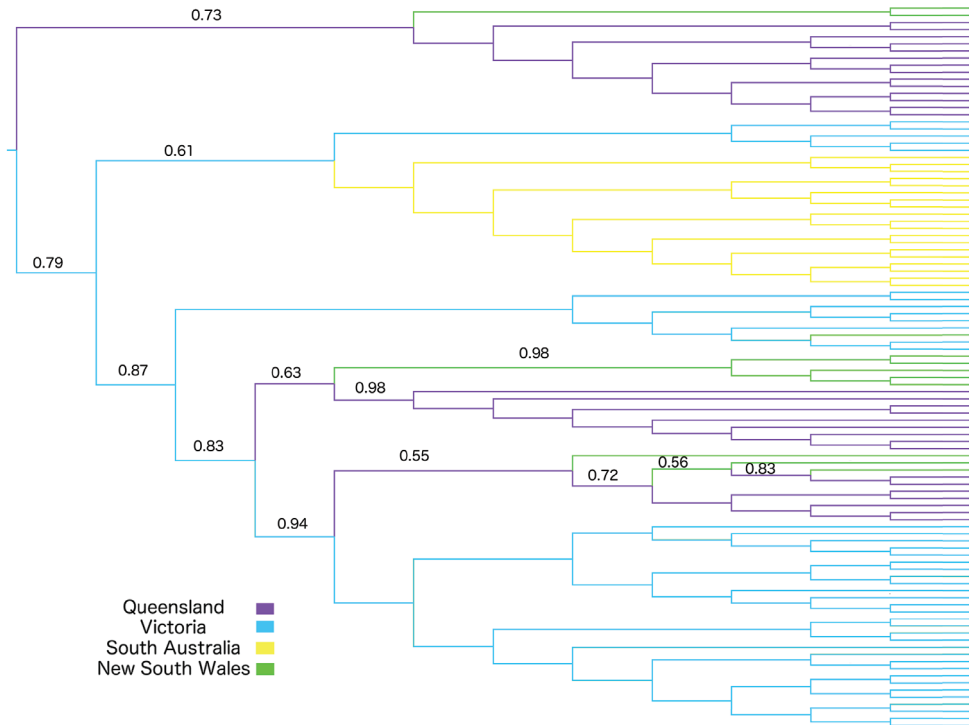


**Figure 4.** Bayesian skyline plot (a), and (b) graphs of two proxies for historical climate in the southern hemisphere (adapted from Pahnke et al. 2003). These proxies are  $\delta^{18}\text{O}$  ‰ and sea surface temperature (SST) based on Mg/Ca ratios. The boxes indicate the approximate period of elevated haplotype accumulation.

The 95% Highest Posterior Densities for the BSP plot in Figure 4 indicate a substantial level of uncertainty in how population size may have changed over time, although a general trend for logistic growth is clear. A Bayes Factor test comparing our BSP model with a constant population size model gave a Bayes factor of 6.164, indicating strong support for increasing population size over time.

### Ancestral distribution

The reconstructed ancestral distribution of haplotypes is shown in Figure 5 with only the probability of the location reconstruction displayed for those branches with a probability below 0.99. The analysis supports a migration from further northeast moving southwest into South Australia. It suggests that there have been multiple dispersals between Victoria and Queensland but one strongly supported dispersal event between the New South Wales and Victorian populations. There is very low support for the ancestral distribution on all branches preceding the Queensland and New South Wales' clades, so we cannot infer directionality of dispersal between these clades and Victoria.



**Figure 5.** Cladogram showing inferred ancestral ranges of haplotypes from a BEAST traits analysis. Posterior probabilities of location reconstruction are shown only on those branches with support less than 0.99.

## Discussion

### Geographic structure

Our haplotype phylogeny, haplotype network and AMOVA analyses suggest geographic structure in haplotypes between the four sample sites. The NTH1 clade consisting of specimens from New South Wales and Queensland forms a sister clade to all other lineages (Figure 2), separated by a minimum of nine base pair mutations (Figure 3). Clade STH1 is restricted to the more southwestern populations of South Australia and Mildura. It is interesting that the South Australian population comprises only a single haplotype. It seems unlikely that a population bottleneck would remove all but one matriline in the population, however without further gene regions we cannot rule out this possibility. Another explanation is that the population has not been in place long enough for new haplotypes to arise and/or that there has been insufficient maternal gene flow from northern populations to promote haplotype diversity above that from a small founder population. The haplotype data are indicative of a southwestwards population expansion.

Interestingly, the second-most common haplotype in our sequences was found in both the Queensland and Victorian samples, and it has given rise to further haplotypes

in both regions and NSW. Our BSP phylogeny (Figure 4) suggests that these haplotypes arose recently, and this might indicate that vagility in *Ceratina australensis* has not been sufficient to completely erode geographical assortment of matriline.

### Historical demography

Our BEAST traits analysis also supports a more northeastern origin with a subsequent introduction into South Australia (Figure 5). The analysis is not able to discern between New South Wales, Queensland and Victoria as the likely origin of *Ceratina australensis* into Australia, but given that the subgenus *C. (Neoceratina)* is a primarily Oriental and Indo-Malayan clade (Rehan and Schwarz 2015), an origin in Queensland seems most likely. Pairwise comparisons indicate that the New South Wales and Queensland populations are not genetically distinct (Table 2). Interestingly the New South Wales population is about equidistant from both the Queensland and Victorian sites, however the Victorian population shows significant genetic distinction from New South Wales. The difference in gene flow between these populations may be due to fragmentation of suitable habitat for *C. australensis* in the semi-arid to arid zone separating the New South Wales and Victorian populations (Figure 1). The traits analysis indicates that multiple dispersals of matriline between the Queensland-New South Wales populations and Victoria have occurred but the direction of movement between populations could not be resolved (Figure 5).

Regardless of when and where *Ceratina australensis* entered Australia, our BSP analyses provide strong support for an increase in effective population size beginning about  $2.5 \times 10^{-3}$  mutations/site ago then plateauing about  $0.8 \times 10^{-3}$  mutations/site ago. Assuming a mutation rate of  $6.2 \times 10^{-8}$  and two generations per year (Shell and Rehan 2016), these values correspond approximately to 20 kya and 6.5 kya respectively.

Our timescale indicates an increase in effective population size approximately 20 kya. This increase could be linked to reduced competition, expansion into a new niche or increased resource availability. To investigate these possibilities we need detailed historical reconstructions, which are not presently available for Australia. Climate reconstructions from 20 kya are available for the southern hemisphere (Pahnke et al. 2003). Climate change may act directly on species, or indirectly, for example by increasing flowering plants and therefore resource availability. In Figure 4 we have contrasted the BSP curve with two climate proxies ( $\delta^{18}\text{O}$  isotopes and estimated sea surface temperatures, SST) for the southern hemisphere reported by Pahnke et al. (2003). These graphs suggest that the major period of increasing  $N_e$  for *Ceratina australensis* coincides with a major period of post-glacial warming, and that the more recent leveling off in  $N_e$  could correspond to a plateauing or slight decline in temperature since 6 kya.

Unfortunately, there are very few detailed studies of paleoclimates in Australia beyond a very small number of sites, limiting further analyses. In one of the most thorough studies, Ayliffe et al. (1998) reconstructed climate in the Naracoorte region, in the South East of Australia, over the past 500,000 years. This geographical location,

however, is well removed from our study sites. Reconstructions of Australia-wide paleoclimates are summarized by Byrne et al. (2008), and while this provides evidence for very broad changes in Australian climates, those studies do not permit reconstruction of paleoclimates in a way that permit refugial areas for *Ceratina australensis* during the last glacial maximum (LGM) to be identified with confidence. However, it seems likely that during the LGM, climatic regimes that now occur in southern Queensland would have had a more northerly distribution.

The inferred increase in  $N_e$  for *Ceratina australensis* coincides closely with the timing of dramatic increases in population size for three independent halictine bee clades in Vanuatu, Fiji and Samoa, each of which corresponded to interglacial warming (Groom et al. 2013, 2014a). It is difficult to imagine a factor other than global climate that would be able to influence isolated bee populations in a similar way across the southwestern Pacific (SWP) and Australia, especially when it is considered that *C. australensis* is in a different family to the SWP halictine bees and has very different nesting and floral-adaptation biologies to halictine bees (stem nesting versus ground nesting, and long-tongued versus short-tongued, respectively). On the other hand, we did not find evidence for a dramatic decline in  $N_e$  of *C. australensis* at the LGM, and this contrasts strongly with studies on SWP halictine bees (Groom et al. 2014a). It is possible that this contrast is due to *C. australensis* occurring on a continent, where it may have been able to persist in a wide range of refugial habitats, which would have not been as abundant for bee species restricted to small islands.

The expansion of population sizes in *Ceratina australensis* in the current interglacial is consistent with expectations for a Mediterranean or subtropical adapted species responding to warming climates in the southern hemisphere, where southern latitudes retreated from glacial conditions experienced at the LGM. This is also concordant with our historical biogeography analyses, which suggests a northeastern origin, followed by accumulation of haplotype diversity in the semi-arid population in northern Victoria and a recent dispersal to South Australia, indicated by the lack of haplotype diversity and BEAST ancestral traits reconstruction.

## Conclusion

Our results suggest that *Ceratina australensis* has responded in major ways to climatic changes since the LGM, but there are two important questions that need resolution. Firstly, because bees are pollinators, historical changes in their diversity and abundance are likely to have impacted angiosperm reproduction in the past, and this may help understand current angiosperm communities. Secondly, if past climates have had large impacts on bee populations in the past, it is important to understand these so that we can anticipate the effects of future climate change. We can only interrogate museum records for impacts of climate change to very limited extents: for Australian insects this will be mostly limited to the last 200 years. In contrast, genetic methods can be used to examine changes well before the recent past and for species that were not covered by

early collectors. Our results suggest that genetic approaches to historical demographics of native bees may hold important insights for understanding how climate change has impacted pollinating biota and plant-pollinator relationships.

## Acknowledgements

We thank Elisabetta Menini for help with PCR protocols and Simon Tierney, Sentiko Ibalim, Olivia Davies and Rebecca Kittel for help with field collections. We thank the Holsworth Wildlife Trust and the Mark Mitchell Foundation for funding this research.

## References

- Ayliffe LK, Marianelli PC, Moriarty KC, Wells RT, McCulloch MT, Mortimer GE, Hellstrom JC (1998) 500 ka precipitation record from southeastern Australia: evidence for interglacial relative aridity. *Geology* 26: 147–150. doi: 10.1130/0091-7613(1998)026<0147:KPRFSA>2.3.CO;2
- Bandelt H-J, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* 16: 37–48. doi: 10.1093/oxfordjournals.molbev.a026036
- Bureau of Meteorology (2012) Australian Climate Averages – Climate Classifications. Government of Australia. [http://www.bom.gov.au/jsp/ncc/climate\\_averages/climate-classifications/index.jsp](http://www.bom.gov.au/jsp/ncc/climate_averages/climate-classifications/index.jsp) [viewed 29 January 2016]
- Burkle LA, Marlin JC, Knight TM (2013) Plant-pollinator interactions over 120 years; loss of species, co-occurrence, and function. *Science* 339: 1611–1615. doi: 10.1126/science.1232728
- Byrne M, Yeates DK, Joseph L, Kearney M, Bowler J, Williams AJ, Cooper S, Donnellan SC, Keogh JS, Leys R, Melville J, Murphy DJ, Porch N, Wyrwoll K-H (2008) Birth of a biome: insights into the assembly and maintenance of the Australian arid zone biota. *Molecular Ecology* 17: 4398–4417. doi: 10.1111/j.1365-294X.2008.03899.x
- Cameron SA, Lozier JD, Strange JP, Koch JB, Cordes N, Solter LF, Griswold TM (2011) Patterns of widespread decline in North American bumble bees. *Proceedings of the National Academy of Sciences of the United States of America* 108: 662–667. doi: 10.1073/pnas.1014743108
- Chenoweth LB, Schwarz M (2011) Biogeographical origins and diversification of the exoneurine allodapine bees of Australia (Hymenoptera, Apidae). *Journal of Biogeography* 38: 1471–1483. doi: 10.1111/j.1365-3113.2008.00432.x
- Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* 9: 772. doi: 10.1038/nmeth.2109
- Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* 7: 214. doi: 10.1186/1471-2148-7-214
- Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* 1: 47–50.

- Fluxus Engineering (2016) Free Phylogenetic Network Software. <http://www.fluxus-engineering.com/sharenet.htm> [viewed 29 January 2016]
- Fuller S, Schwarz M, Tierney S (2005) Phylogenetics of the allodapine bee genus *Braunsapis*: historical biogeography and long-range dispersal over water. *Journal of Biogeography* 32: 2135–2144. doi: 10.1111/j.1365-2699.2005.01354.x
- Grant WS (2015) Problems and cautions with sequence mismatch analysis and Bayesian skyline plots to infer historical demography. *Journal of Heredity* 106: 333–346. doi: 10.1093/jhered/esv020
- Grant WS, Liu M, Gao T, Yanagimoto T (2012) Limits of Bayesian skyline plot analysis of mtDNA sequences to infer historical demographics in Pacific herring (and other species). *Molecular Phylogenetics and Evolution* 65: 203–212. doi: 10.1016/j.ympev.2012.06.006
- Groom SVC, Ngo HT, Rehan SM, Skelton P, Stevens MI, Schwarz MP (2014b) Multiple recent introductions of apid bees into pacific archipelagos signify potentially large consequences for both agriculture and indigenous ecosystems. *Biological Invasions* 16: 2293–2302. doi: 10.1007/s10530-014-0664-7
- Groom SVC, Stevens MI, Schwarz MP (2014a) Parallel responses of bees to pleistocene climate change in three isolated archipelagos of the southwestern pacific. *Proceedings of the Royal Society B* 281: 20133293. doi: 10.1098/rspb.2013.3293
- Groom SVC, Stevens MI, Schwarz MP (2013) Diversification of Fijian halictine bees: insights into a recent island radiation. *Molecular Phylogenetics and Evolution* 68: 582–594 doi: 10.1016/j.ympev.2013.04.015
- Guindon S, Gascuel O (2003) A simple, fast and accurate method to estimate large phylogenies by maximum likelihood. *Systematic Biology* 52: 696–704. doi: 10.1080/10635150390235520
- Haag-Liautard C, Coffey N, Houle D, Lynch M, Charlesworth B, Keightley PD (2008) Direct estimation of the mitochondrial DNA mutation rate in *Drosophila melanogaster*. *PLoS Biology* 6: e204. doi: 10.1371/journal.pbio.0060204
- Hines H (2008) Historical biogeography, divergence times, and diversification patterns of bumble bees (Hymenoptera: Apidae: *Bombus*). *Systematic Biology* 57: 58–75. doi: 10.1080/10635150801898912
- Ho SYW, Shapiro B (2011) Skyline-plot methods for estimating demographic history from nucleotide sequences. *Molecular Ecology Resources* 11: 423–434. doi: 10.1111/j.1755-0998.2011.02988.x
- Ivanova NV, Dewaard JR, Hebert PDN (2006) An inexpensive, automation-friendly protocol for recovering high-quality DNA. *Molecular Ecology Notes* 6: 998–1002. doi: 10.1111/j.1471-8286.2006.01428.x
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A (2012) Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28: 1647–1649. doi: 10.1093/bioinformatics/bts199
- López-Urbe 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. doi: 10.1111/mec.12689



- Michener CD (1962) The genus *Ceratina* in Australia with notes on its nests (Hymenoptera: Apoidea). Journal of the Kansas Entomological Society 35: 414–421.
- Pahnke K, Zahn R, Elderfield H, Schulz M (2003) 340,000-year centennial-scale marine record of Southern Hemisphere climatic oscillation. Science 301: 948–952. doi: 10.1126/science.1084451
- Perkins RCL (1912) Notes with descriptions of new species, on Aculeate Hymenoptera of the Australian region. Annals and Magazine of Natural History 9: 96–121. doi: 10.1080/00222931208693112
- Rehan SM, Chapman TW, Craigie AI, Richards MH, Cooper SJB, Schwarz MP (2010a) Molecular phylogeny of the small carpenter bees (Hymenoptera: Apidae: Ceratinini) indicates early and rapid global dispersal. Molecular Phylogenetics and Evolution 55: 1042–1054. doi: 10.1016/j.ympev.2010.01.011
- 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 doi: 10.1016/j.anbehav.2014.08.021
- Rehan SM, Richards MH, Schwarz MP (2010b) Social polymorphism in the Australian small carpenter bee, *Ceratina* (*Neoceratina*) *australensis*. Insectes Sociaux 57: 403–412. doi: 10.1007/s00040-010-0097-y
- Rehan SM, Schwarz MP (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. doi: 10.1111/jbi.12439
- Rehan SM, Schwarz MP, Richards MH (2011) Fitness consequences of ecological constraints for the evolution of sociality in an incipiently social bee. Biological Journal of the Linnean Society 103: 57–67. doi: 10.1111/j.1095-8312.2011.01642.x
- Salzburger W, Ewing GB, von Haesler A (2011) The performance of phylogenetic algorithms in estimating haplotype genealogies with migration. Molecular Ecology 20: 1952–1963. doi: 10.1111/j.1365-294X.2011.05066.x
- Shell WA, Rehan SM (2016) Recent and rapid diversification of the small carpenter bees in eastern North America. Biological Journal of the Linnean Society 117: 633–645. doi: 10.1111/bij.12692
- Wilson JS, Carril OM, Sipes SD (2014) Revisiting the great American biotic interchange through analyses of amphitropical bees. Ecography 37: 791–796. doi: 10.1111/ecog.00663

## Supplementary material I

### **Cladogram including *Ceratina* (Neoceratina) Solomons sp. as the outgroup to root the tree**

Authors: Rebecca M. Dew, Sandra M. Rehan, Michael P. Schwarz

Data type: EPS file

Explanation note: Cladogram including *Ceratina* (Neoceratina) Solomons sp. as the outgroup to root the tree. Posterior probability values: \*\*\* = 1.0; \*\*  $\geq$  95; \*  $\geq$  85.

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.