

## ORIGINAL ARTICLE

# Phylogeographic evidence for two mesic refugia in a biodiversity hotspot

H Nistelberger<sup>1</sup>, N Gibson, B Macdonald, S-L Tapper and M Byrne

Phylogeographic studies of flora in species-rich south-western Australia point to complex evolutionary histories, reflecting patterns of persistence and resilience to climatic changes during the Pleistocene. We asked whether coastal areas of the mid-west and south, as well as granite outcrops and inland ranges, have acted as major refugia within this region during Pleistocene climatic fluctuations by analysing phylogeographic patterns in the shrub *Calothamnus quadrifidus* R.Br. (Myrtaceae). We determined variation in chloroplast DNA data for 41 populations across the geographic range. Relationships and major clades were resolved using parsimony and Bayesian analyses. We tested for demographic and spatial expansion of the major clades and estimated clade divergence dates using an uncorrelated, lognormal relaxed clock based on two conservative chloroplast mutation rates. Two distinct phylogeographic clades were identified showing divergence during the Pleistocene, consistent with other phylogeographic studies of south-west Australian flora, emphasising the impact of climatic oscillations in driving divergence in this landscape. The southern clade was more diverse, having higher haplotype diversity and greater genetic structure, while the northern clade showed evidence of fluctuation in population size. Regions of high haplotype diversity with adjacent areas of low diversity observed in each clade indicated the locations of two coastal refugia: one on the south coast and another along the mid-west coast. This is the first evidence for major Pleistocene refugia using chloroplast genetic data in a common, widespread species from this region.

*Heredity* (2014) **113**, 454–463; doi:10.1038/hdy.2014.46; published online 2 July 2014

## INTRODUCTION

Phylogeographic studies at the intraspecific level can reveal important aspects of the evolutionary history of species, including areas of refugia during times of climatic stress and areas of expansion that may occur under favourable conditions. Refugia are habitats that species contract to when environmental conditions are unfavourable elsewhere in their range and potentially expand from when those conditions improve (Keppel *et al.*, 2012). Refugia have intrinsic conservation value, because they enable the persistence of species during periods of unfavorable environmental conditions, promote genetic diversification and often contain rare taxa or genotypes (Ashcroft *et al.*, 2012). Hence, the location of refugia and their role in facilitating the diversification and persistence of species throughout Pleistocene climatic oscillations have been broadly studied (see Keppel *et al.*, 2012). Many of these studies have focused on glaciated landscapes of the Northern Hemisphere where broad geographic regions have acted as major refugia for a range of plants and animals during the glacial cycles of the Pleistocene, such as in the temperate regions of Europe and North America (see Hewitt, 2000, 2004). However, an increasing number of studies are investigating how species respond to climatic shifts in the Southern Hemisphere (see Beheregaray, 2008). Studies of species found in non-glaciated landscapes with subdued topography and geological stability, such as southwestern Australia, have also revealed strong phylogeographic signals but have identified long-term persistence in localised refugia as

the major factor in facilitating persistence of species through climatic change (for example, Byrne, 2007; Byrne *et al.*, 2008, 2011). This pattern of localised refugia is quite different to the contraction and migration from major refugia that is a common pattern in Northern Hemisphere glaciated landscapes.

The south-west of Western Australia (herein the Southwest Australian Floristic Region (SWAFR) *sensu* Hopper and Gioia, 2004) is a global biodiversity hotspot well known for the high diversity and endemism of the flora, a result of long evolutionary history in a highly stable landscape that has not been glaciated since the Permian (Hopper, 1979; Myers *et al.*, 2000). The flora also shows high levels of genetic diversity and differentiation among populations in both restricted and widespread species. This genetic structuring has been attributed to complex edaphic patterns that leads to a mosaic of vegetation types and can drive localised selection, along with the influence of Pleistocene climatic fluctuations associated with Milankovitch cycles. Intraspecific phylogeography has highlighted the complex evolutionary history of species in the SWAFR, including divergence of northern and southern clades that is hypothesised to have been driven by major contraction during climatic changes associated with the onset of aridification and formation of desert environments in the mid-Pleistocene (Byrne, 2008; Fujioka *et al.*, 2009). Patterns within the intraspecific phylogeographic clades indicate different responses in the late Pleistocene, with evidence for multiple localised 'micro' refugia dispersed throughout species

Science and Conservation Division, Department of Parks and Wildlife, Bentley, Western Australia, Australia

<sup>1</sup>Current address: School of Animal Biology, The University of Western Australia, M092, 35 Stirling Highway, Crawley, Western Australia 6009, Australia.

Correspondence: H Nistelberger, School of Animal Biology, The University of Western Australia, M092, 35 Stirling Highway, Crawley, Perth, Western Australia 6009, Australia.  
E-mail: heidi.nistelberger@grs.uwa.edu.au

Received 18 February 2014; accepted 3 April 2014; published online 2 July 2014

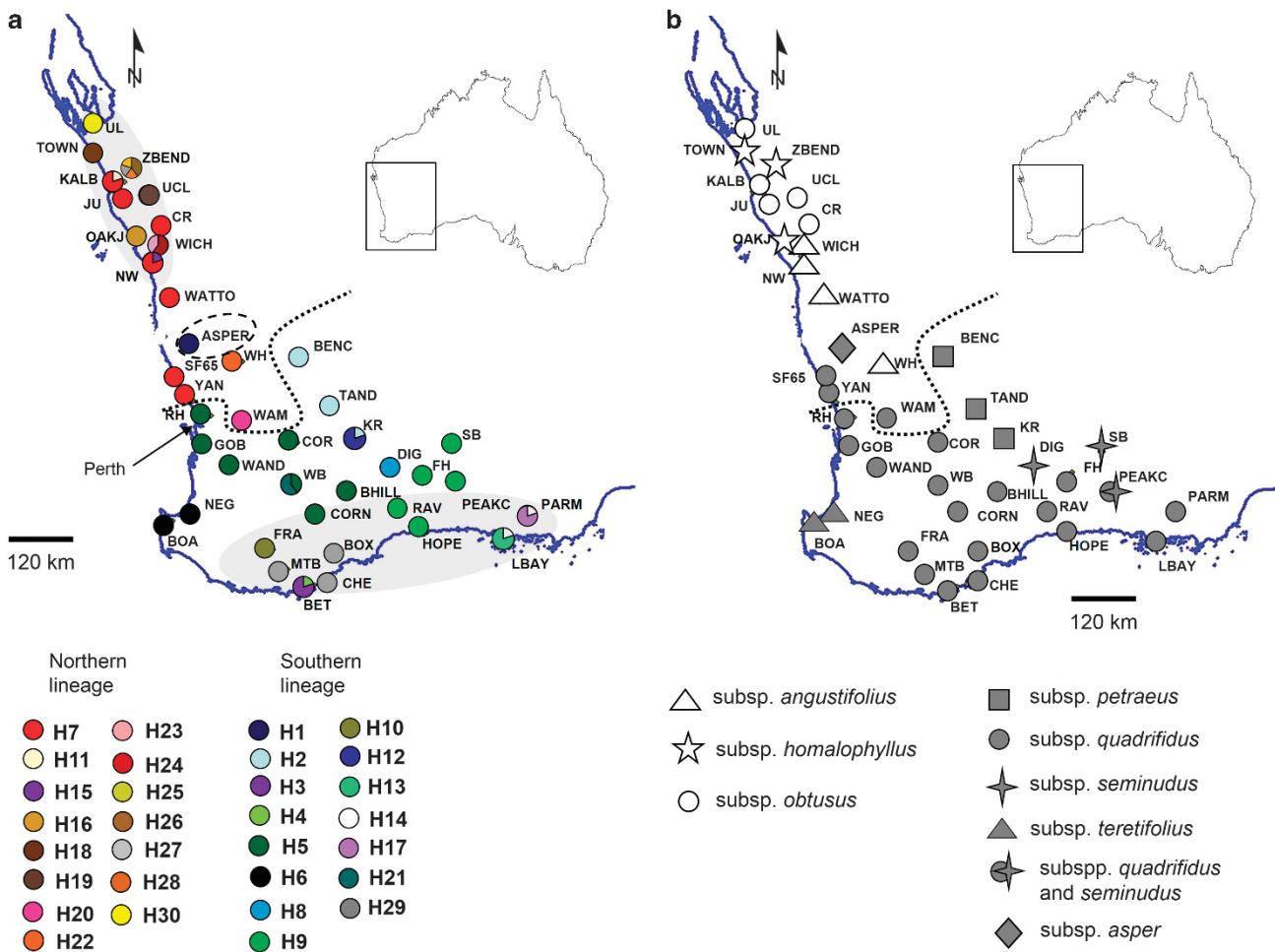
distributions with haplotypes specific to populations and no signal of expansion in most species (Byrne *et al.*, 2008, 2011; Byrne, 2008). Interestingly, these intraspecific phylogeographic studies have not documented regions of major refugia where species have contracted to during arid conditions and expanded from during times of more favourable climate, as would have been expected if large-scale contraction to major refugia was a primary response to climate change (Keppel *et al.*, 2012).

Major areas of species richness and endemism occur in edaphically complex heathlands where the transitional rainfall zone intersects the west and south coastal areas (Hopper, 1979; Hopper and Gioia, 2004). These areas would experience moderating influences of maritime conditions and have been postulated to have provided sufficiently heterogeneous yet broadly climatically stable environments to foster persistence and diversification of genera and species through time (Hopper, 1979). These areas of species richness and endemism have been hypothesised as areas that have remained habitable during Miocene and Pliocene aridification, and would also be predicted to represent major refugia during Pleistocene climatic fluctuations, where populations of species could contract to during arid conditions and expand from in more suitable mesic conditions (Keppel *et al.*, 2012). Topographical features, such as ranges and granite outcrops,

may also provide sufficient environmental complexity for populations of species to persist through Pleistocene climatic fluctuations and expand from during later suitable climatic conditions (Gardner, 1944; Marchant, 1973; Hopper, 2000).

Phylogeographic and geomorphological data would thus suggest areas with features consistent with a role as major refugia during Pleistocene climatic fluctuations, yet localised refugia is the common pattern in many species of southern Australia, including the Southwest Australian Floristic Region (Byrne, 2007; Byrne *et al.*, 2008, 2011). The lack of identification of areas of major refugia may be because most of the studies undertaken to date have not investigated species with widespread distributions that cover the western and southern coastal regions and inland areas, including ranges and granite outcrops. By selecting a species with a widespread distribution that encompasses these areas, we have an opportunity to test hypotheses that coastal areas, inland ranges and granite outcrops have had a role as major refugia in this biodiversity hotspot.

*Calothamnus quadrifidus* R.Br. (Myrtaceae) provides the opportunity to investigate phylogeographic signatures of refugia as it is widely distributed from Shark Bay in the north to Esperance in the south-east and inland to granite outcrops and ranges in the semi-arid region. *C. quadrifidus* forms a dominant component of many



**Figure 1** (a) Map of the 41 sampled populations of *Calothamnus quadrifidus*. Population abbreviations as in Table 1. Colours represent unique haplotypes as identified by cpDNA sequencing, split circles indicate multiple haplotypes per site. The dashed line denotes the border between the identified northern and southern phylogeographic clades. Shading indicates the broad refugial areas. Note that subspecies *C. asper* is placed in the southern clade. (b) Map of *C. quadrifidus* populations as in panel (a) showing the distribution of the eight subspecies identified based on morphological assessment. \*Two subspecies occur at Peak Charles.

vegetation communities in the SWAFR, being only absent from the wetter areas of the south-western forest. This distribution covers northern and southern coastal regions, as well as inland granite outcrops and ranges that show high levels of species richness and endemism (Hopper and Gioia, 2004).

We investigated phylogeographic patterns in *C. quadrifidus* to test the hypothesis that coastal areas of the mid-west and south, and topographic features such as granite outcrops and inland ranges, have acted as major refugia where species have contracted to, and expanded from, during Pleistocene climatic fluctuations. Major refugia would be expected to harbour high diversity of haplotypes as a result of persistence of populations over time in conjunction with low haplotype diversity in surrounding areas where expansion has occurred. Alternatively, signatures of haplotypes specific to populations with no signal of expansion would indicate a pattern of persistence in localised 'micro' refugia without major contraction/expansion dynamics. We also determined climatic variables associated with the current distribution and modelled the species distribution at the Last Glacial Maximum (LGM) to determine whether coastal areas would be predicted to have maintained suitable environments for persistence of the species.

## MATERIALS AND METHODS

### Study species

*C. quadrifidus* is a widespread diploid woody shrub, 2–3 m in height with a generation time of around 50–70 years. It is mass-flowering with multiple inflorescences comprised of 10–30 hermaphrodite red flowers that are protandrous and pollinated by birds, honey possums or honeybees and geitonogamous pollination is possible. The species is serotinous, retaining fruit on the plant for up to 4 years, and the seed is very small and gravity dispersed. *C. quadrifidus* grows in low open heath and shrubland vegetation that is patchily distributed, forming complex landscape mosaics with woodlands and mallee vegetation. The species distribution is centred in the mesic SWAFR, but it can tolerate some degree of water stress as it extends to the sandplains and gorges of Kalbarri and Shark Bay in the north while on its inland boundary it is largely restricted to water-gaining sites around granites and greenstone ranges.

The species has been recently revised and is now considered to consist of eight closely related subspecies that have overlapping geographic ranges (George and Gibson, 2010) (Figure 1b). Samples from all subspecies of *C. quadrifidus* were obtained across the geographic range of the species from 41 different populations (Table 1, Figure 1). The sampling regime did not sample from regions of overlap between subspecies except for the Peak Charles population where two subspecies co-occur. Samples of the closely related species *C. sanguineus* Labill and *C. rupestris* Shauer were collected for outgroups.

### DNA isolation and identification of cpDNA variation

DNA was extracted from fresh leaf material following the methods outlined in Byrne and Moran (1994). The level of variation for 13 non-coding regions of the chloroplast genome has been tested in six plant species from south-west Western Australia, including *C. quadrifidus* (Byrne and Hankinson, 2012). Three of the most polymorphic regions (D4 loop introns *rpl16* and *petB*; intergenic spacer *trnQ-rps16*) were selected for amplification using the PCR on DNA from five plants from each population of *C. quadrifidus*. PCR reactions were carried out in individual, non-multiplexed 50- $\mu$ l volumes containing 40 ng of template DNA, 0.08  $\mu$ M of *rpl16* and *PetB* primers (F71, R1516; sak23F, sak24R, respectively) and 0.1  $\mu$ M of *trnQ-rps16* primers (*trnQ*, *rps16*), 1.5 mM MgCl<sub>2</sub> for *rpl16* and *PetB* and 3 mM for *trnQ-rps16*, 0.3 U of Taq polymerase (Invitrogen, Perth, WA, Australia) for *rpl16* and *PetB* and 0.5 U for *trnQ-rps16*. The optimum amplification protocol involved an initial hot start of 5 min at 80 °C followed by 30 cycles of 95 °C for 1 min, 50 °C for 1 min and 65 °C for 4 min, with a final extension step of 65 °C for 5 min. PCR products were purified using a polyethylene glycol precipitation method (Travis Glenn, University of South Carolina, Columbia, SC, USA). Purified PCR products

(2  $\mu$ l) were then quantified on a 2% agarose gel with 2  $\mu$ l of low mass molecular marker (Invitrogen), using the Genetools software (Syngene, Cambridge, UK). Sequencing reactions (1/4) were carried out in 10- $\mu$ l volumes containing 1  $\mu$ l of 5x Big Dye Terminator buffer, 2  $\mu$ l of Big Dye Terminator, 1  $\mu$ l of 2  $\mu$ M primer for *rpl16* (sak16F) and *petB* (sak23F) and 1.6  $\mu$ l of 2  $\mu$ M *trnQ* primer and 40 ng of polyethylene glycol purified, template DNA. Cycling conditions involved an initial hot start of 96 °C for 2 min followed by 35 cycles of 96 °C for 10 s, 50 °C for 5 s and 60 °C for 4 min. Sequenced products were cleaned before electrophoresis using a standard ethanol precipitation method (Applied Biosystems, Mulgrave, VIC, Australia) and sequenced on an ABI 96-well capillary sequencer (Murdoch University, Perth, WA, Australia).

Restriction fragment length polymorphism (RFLP) analysis was also undertaken in order to enable comparison of this study with previous studies in this landscape that have utilised cpRFLP methodology (see Supplementary Information).

### Data analysis

Sequence data were aligned using Clustal W in BioEdit (Hall, 1999) and finalised by eye. The three chloroplast regions were combined to form a total aligned sequence length of 1948 bp. We estimated evolutionary relationships and time to the most recent common ancestor of all sequence haplotypes using a Bayesian approach implemented in BEAST ver. 1.6.2 (Drummond and Rambaut, 2007). Data were partitioned into outgroups (*Calothamnus rupestris*, CRUP, and *Calothamnus sanguineus*, CSAN) and the ingroup. Initially, the substitution model was set to GTR +  $\Gamma$  as inferred from jModelTest (Posada, 2008) with the coalescent tree prior set to Constant Size; however, this model was found to be too parameter-rich for the data set and prevented stationarity from being reached despite extensive chain lengths. Consequently, we used the lower parameter HKY +  $\Gamma$  model that resulted in faster convergence of runs. As no fossil evidence or geological events were available to calibrate a molecular clock, we used two mutation rates commonly reported for chloroplast plant genes ( $1.2 \times 10^{-9}$  (Graur and Li, 2000) and  $2.0 \times 10^{-9}$  substitutions per site per year (Wolfe *et al.*, 1987)) and assumed an uncorrelated lognormal relaxed clock to estimate divergence times based on our combined, non-coding cpDNA sequence data. Two independent runs of 10 million generations were carried out, sampling every 1000 generations. The log files were analysed in TRACER ver. 1.5.0 (Drummond and Rambaut, 2007) to ensure stationarity had been reached. This was achieved by both visual inspection of the traces and by ensuring the effective sample sizes for all parameters exceeded 200. Tree files were combined in LogCombiner ver. 1.6.2 (BEAST package) and the combined tree file annotated using TreeAnnotator ver. 1.6.2 (BEAST package) and visualised in Figtree ver. 1.3.1.

A maximum parsimony median-joining haplotype network of sequence data was also generated in Network ver. 4.5.1.6 (Fluxus Technology, Suffolk, UK), because networks often provide a more useful method of visualising intraspecific data with low levels of divergence (Templeton *et al.*, 1992). Indels of varying length were treated as one character of equal weight to all other substitutions, epsilon was set to 0. To confirm topology, we also calculated a haplotype network using a statistical parsimony method implemented in TCS ver. 1.21 (Clement *et al.*, 2000).

Descriptive statistics, including nucleotide diversity and haplotype diversity, were estimated using DnaSP ver. 5.10 (Librado and Rozas, 2009). Estimates of  $G_{ST}$  and  $N_{ST}$  were calculated in PERMUT ver. 1.0 (Pons and Petit, 1996) to provide a measure of phylogeographic structure. Tests of neutrality, which also provide useful measures of population growth, decline or stability, Tajima's  $D$  (Tajima, 1989) and  $R_2$  (Ramos-Onsins and Rozas, 2002) were calculated across all *C. quadrifidus* samples and within major clades to examine whether sequence data fit the assumption of neutral evolution (Tajima, 1989) and to determine whether any patterns of haplotype expansion or contraction were evident. Patterns of population growth or stability under a neutral evolution model were tested using coalescent simulations permuted 10 000 times in DnaSP. To further test clade expansion, both spatial and demographic, we conducted mismatch distributions that compare the observed and expected pairwise sequence differences in ARLEQUIN ver. 3.5 (Excoffier and Lischer, 2010). Goodness-of-fit tests using Harpending's raggedness index (HRag) and the sum of squared differences (SSD) were used to test each model.

**Table 1** Details of subspecies and populations sampled within the *Calothamnus quadrifidus* complex and the haplotypes identified through sequence analysis

Subspecies	Population	Population abbreviation	Latitude	Longitude	cpDNA sequence haplotypes	Specimen collection no.	Genbank
<i>angustifolius</i>	North West	NW	-28.8921	115.0584	H7, <b>H15</b>	NG6130;6131	JQ087340;JQ087348;JQ087365 JQ087369
<i>angustifolius</i>	Watto	WATTO	-29.7346	115.41	H7	NG5996;5997	JQ087340;JQ087348;JQ087365
<i>angustifolius</i>	Wicherina	WICH	-28.7202	115.0209	<b>H23, H24</b>	NG5966;5968	JQ087338;JQ087348;JQ087365 KC895505
<i>angustifolius</i>	Wongan Hills	WH	-30.8658	116.7141	<b>H22</b>	NG6040;6041	JQ087340;JQ087356;JQ087373
<i>asper</i>	Asper	ASP	-30.8618	116.6212	<b>H1</b>	NG6010;6011	JQ087335;JQ087343;JQ087361
<i>homalophyllus</i>	Oakajee	OAKJ	-28.5689	114.6522	<b>H16</b>	NG5950;5951	JQ087340;JQ087348;JQ087370
<i>homalophyllus</i>	Town	TOWN	-27.7178	114.1572	<b>H18</b>	NG5905;5906	JQ087340;JQ087348;JQ087372
<i>homalophyllus</i>	Z Bend	ZBEND	-27.5647	114.4373	<b>H25,26,27,28</b>	NG5884;5888	JQ087340;JQ087373;JQ087374 JQ087365;JQ087374; KC895503 KC895504
<i>obtusus</i>	Campbell Rd	CR	-28.4122	115.1916	H7	NG5935;5936	JQ087340;JQ087348;JQ087365
<i>obtusus</i>	Junga	JU	-27.7586	114.3683	H7	NG5890;5891	JQ087340;JQ087348;JQ087365
<i>obtusus</i>	Kalbarri	KALB	-27.6651	114.2927	H7, <b>H11</b>	Ng5860;5861	JQ087340;JQ087348;JQ087351 JQ087365;JQ087366
<i>obtusus</i>	UCL	UCL	-27.8394	114.8825	<b>H19</b>	NG5920;5921	JQ087340;JQ087354;JQ087365
<i>obtusus</i>	Useless Loop	UL	-26.6063	113.9204	<b>H30</b>	NG5845;5846	JQ087340;JQ087351;JQ087365
<i>petraeus</i>	Bencubbin	BENC	-30.8156	117.8714	<b>H2</b>	NG6025;6027	JQ087337;JQ087344;JQ087362
<i>petraeus</i>	King Rock	KR	-32.3203	119.1478	<b>H2, H12</b>	NG6340;6341	JQ087337;JQ087352;JQ087362
<i>petraeus</i>	Tandagin	TAND	-31.6688	118.5055	<b>H2</b>	NG6055;6056	JQ087337;JQ087344;JQ087362
<i>quad/semi</i>	Peak Charles	PEAKC	-32.8857	121.1687	H9	NG6310;6311	JQ087338;JQ087344;JQ087362
<i>quadrifidus</i>	Betty's Beach	BET	-34.9388	118.2074	<b>H3, H4</b>	NG6241;6243	JQ087338;JQ087345;JQ087363 KC895501
<i>quadrifidus</i>	Boxwood Hill	BOX	-34.3696	118.7498	<b>H29</b>	NG6265;6266	JQ087338;JQ087346;JQ087363
<i>quadrifidus</i>	Breakaway Hill	BHILL	-33.2507	118.9519	<b>H5</b>	NG6355;6356	JQ087338;JQ087346;JQ087362
<i>quadrifidus</i>	Cheyne beach	CHE	-34.8808	118.4119	<b>H29</b>	NG6250;6258	JQ087338;JQ087346;JQ087363 JQ087364
<i>quadrifidus</i>	Corneecup	CORN	-33.6525	118.3049	<b>H5</b>	NG6190;6191	JQ087338;JQ087346;JQ087362
<i>quadrifidus</i>	Corrigin	COR	-32.3375	117.8408	<b>H5</b>	NG6071;6077	JQ087338;JQ087346;JQ087362
<i>quadrifidus</i>	Frank Hahn Nat. Res.	FH	-32.8068	120.5238	H9	NG6295;6296	JQ087338;JQ087344;JQ087362
<i>quadrifidus</i>	Frankland	FRA	-34.3016	117.3476	<b>H10</b>	NG6205;6206	JQ087338;JQ087350;JQ087362
<i>quadrifidus</i>	Gobby Rd	GOB	-32.4406	116.0179	<b>H5</b>	NG6400;6401	JQ087338;JQ087346;JQ087362
<i>quadrifidus</i>	Hopetoun	HOPE	-33.941	120.4977	H9	NG6830;6831	JQ087338;JQ087344;JQ087362
<i>quadrifidus</i>	Lucky Bay	LBAY	-33.9947	122.2461	<b>H13, H14</b>	NG5800;5806	JQ087338;JQ087344;JQ087353 JQ087367;JQ087368
<i>quadrifidus</i>	Mt Barker	MTB	-34.6556	117.6475	<b>H29</b>	NG6220;6223	JQ087338;JQ087346;JQ087363
<i>quadrifidus</i>	Parmango	PARM	-33.5533	122.7468	<b>H14, H17</b>	NG5816;5818	JQ087338;JQ087344;JQ087367 JQ087371
<i>quadrifidus</i>	Ravensthorpe Range	RAV	-33.5123	120.0324	H9	NG6280;6281	JQ087338;JQ087344;JQ087362
<i>quadrifidus</i>	Red Hill	RH	-31.845	116.0703	<b>H5</b>	NG6145;6146	JQ087338;JQ087346;JQ087362
<i>quadrifidus</i>	State Forest 65	SF 65	-31.1279	115.4594	H7	NG5982;5983	JQ087340;JQ087348;JQ087365
<i>quadrifidus</i>	Wambyn	WAM	-31.8958	116.6354	<b>H20</b>	NG6160;6161	JQ087340;JQ087355;JQ087365
<i>quadrifidus</i>	Wandering	WAND	-32.7078	116.5941	<b>H5</b>	NG6102;6103	JQ087338;JQ087346;JQ087362
<i>quadrifidus</i>	Wishbone	WB	-33.0739	117.8351	<b>H5, H21</b>	NG6175;6176	JQ087338;JQ087346;JQ087362 KC895502
<i>quadrifidus</i>	Yanchep	YAN	-31.5705	115.7181	H7	NG6373;6374	JQ087340;JQ087348;JQ087365
<i>seminudus</i>	Diggers Rock	DIG	-32.7263	119.827	<b>H8</b>	NG6329;6333	JQ087337;JQ087349;JQ087362
<i>seminudus</i>	South Bremer	SB	-32.6468	120.7808	H9	NG6430;6431	JQ087338;JQ087344;JQ087362
<i>teretifolius</i>	Boallia	BOA	-33.751	115.2687	<b>H6</b>	NG6115;6118	JQ087339;JQ087347;JQ087362
<i>teretifolius</i>	Negus Block	NEG	-33.6655	115.5718	<b>H6</b>	NG6085;6086	JQ087339;JQ087347;JQ087362

Abbreviation: RFLP, restriction fragment length polymorphism.

Haplotypes restricted to a single subspecies are in bold. Plant vouchers are lodged in the Western Australian Herbarium.

**Species distribution modelling**

Modelling the current climatic envelope of *C. quadrifidus* was undertaken using 2.5-min grids (~5 km<sup>2</sup>) of 19 bioclimatic variables available from WorldClim based on the period 1950–2000 (Hijmans *et al.* 2005). Bioclimatic

variables at the LGM were taken from WorldClim and were developed from the Paleoclimate Modelling Intercomparison Project Phase II—CCSM LGM model. The species distribution data for *C. quadrifidus* was extracted from the Western Australian Herbarium database. Some 804 records were available, for

which 775 had accurate spatial coordinates. Duplicates, and any records falling within the same grid, were removed leaving a total of 632 records included in the model.

Species distribution modelling was undertaken in Maxent (ver. 3.3.3k) (Phillips *et al.*, 2006) using 20 000 background points selected from the south-west of WA with the inland boundary approximating the 250 mm rainfall isohyet. These background points were generated from the WorldClim grids using the method outlined in Appendix 4 of the supplement of Elith *et al.* (2011) to correct for the decreasing size of geographical grids with increasing latitude. The model was run using the defaults except that only hinge features were used to ensure smooth response curves. In addition, response curve graphs and jack knife estimates of the contribution of different bioclimatic variables were also generated. The model was projected to LGM conditions using the CCSM bioclim grids.

## RESULTS

### Haplotype variation

Chloroplast sequence analysis showed variation in all three intergenic spacers used. Sequencing of 414 bp of the *rpl16* D loop intron identified four transitions, one multi-state substitution and two indels in *C. quadrifidus* samples. The *petB* D loop intron yielded 656 bp of sequenced product, with six transitions, five transversions and six indels detected. The intergenic spacer *trnQ-rps16* showed the most variation with 12 transitions, eight transversions and five indels identified in 851 bp of sequenced product. Data obtained by combining the three sequenced DNA regions (total 1930 bp including gaps) resulted in the identification of 30 haplotypes from a total of 203 samples (Table 2). The most prevalent haplotype, H5, occurred at a frequency of 15.7%. Other haplotypes ranged in frequency from 0.49% to 15.3% (Table 2).

### Phylogenetic analysis of haplotype relationships

Phylogenetic analyses of the combined sequence data resulted in the identification of two major clades corresponding to the north and south of the species distribution. The northern clade extended along the western coastline from Useless Loop in the north to Wambyn in the Perth area and the southern clade encompassed all other populations within the south-west. The southern clade showed a deeper level of genetic structuring than the northern clade, with three main clusters identified, two of them with very high support (Figure 2). The topology of the cpDNA sequence tree was similar to that generated by parsimony analysis of RFLP haplotypes (Supplementary Figure S1).

The median-joining network of cpDNA sequence haplotypes illustrates a star-like phylogeny in the northern clade with the distribution of 14 haplotypes stemming from the widespread haplotype H7. In comparison, the southern clade had greater structure with a number of common haplotypes and derived haplotype groups (Figure 3). A statistical parsimony network showed an identical topology to the median-joining network.

Population differentiation within both of the regions and across the entire distribution was high (Table 3). Values of  $N_{ST}$  were significantly higher than  $G_{ST}$  across the entire region and within the southern clade, indicating phylogeographic structure.

### Spatial distribution of haplotypes

The distribution of haplotypes was strongly correlated with geography. Within the northern clade, there was a common sequence haplotype (H7, 15.3%) that was widely distributed across seven populations along the coastline (Figure 1). A region of high haplotype diversity among populations was identified in the Kalbarri–Shark Bay region along the mid-west coast.

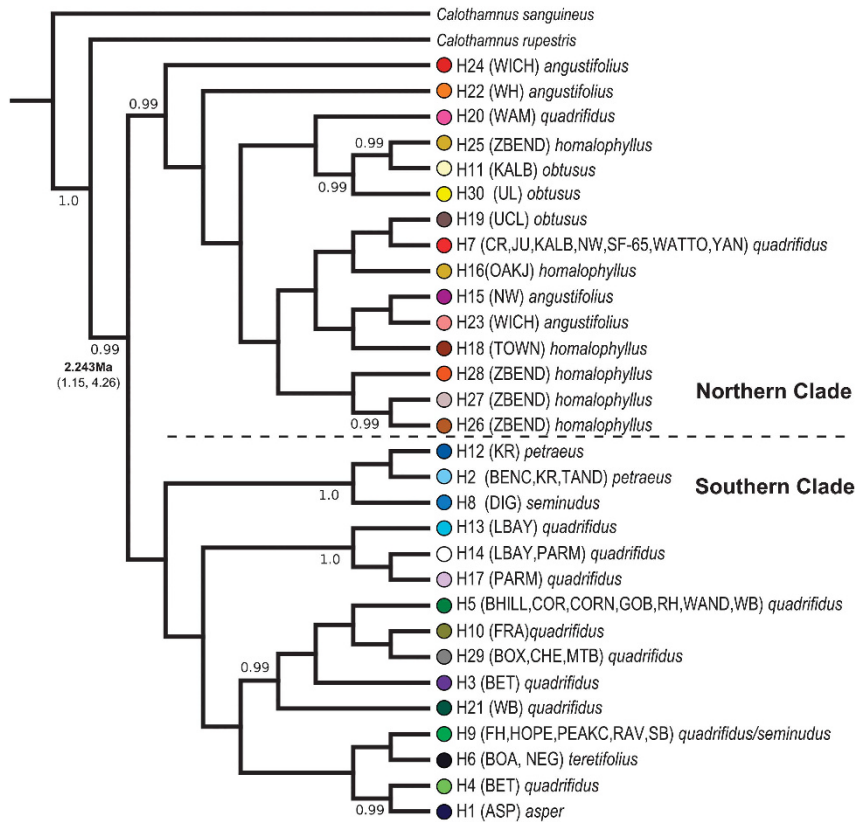
**Table 2** Frequency of cpDNA sequence haplotypes observed in the *Calothamnus quadrifidus* complex

cpDNA sequence haplotype	Haplotype frequency (%)
H1	2.46
H2	5.42
H3	1.97
H4	0.49
H5	15.7
H6	4.93
H7	15.3
H8	2.46
H9	11.8
H10	2.46
H11	0.49
H12	1.97
H13	1.97
H14	0.99
H15	0.99
H16	2.46
H17	1.97
H18	2.46
H19	2.46
H20	2.46
H21	1.48
H22	2.46
H23	0.99
H24	1.48
H25	0.49
H26	0.49
H27	0.99
H28	0.49
H29	7.39
H30	2.46

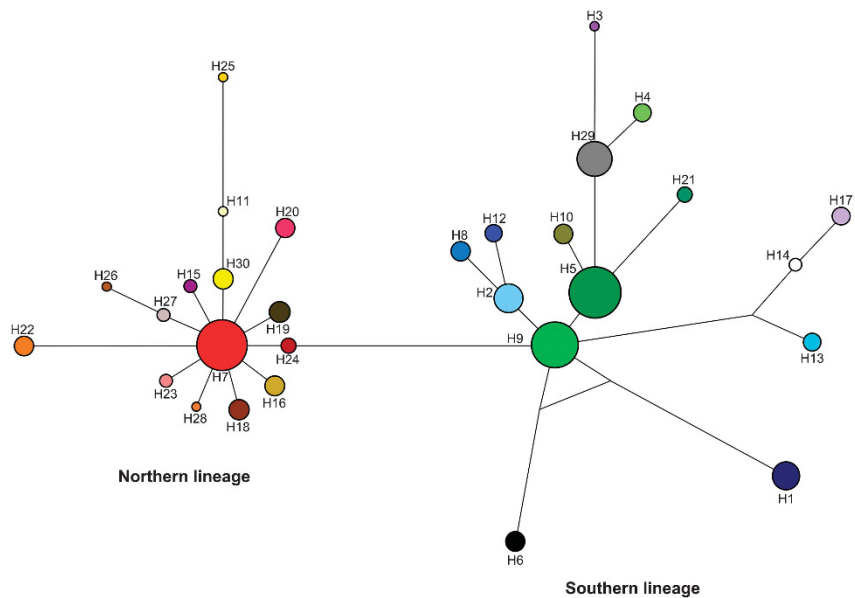
Haplotypes within the southern clade showed greater spatial complexity in distribution. There were several common haplotypes clustered around the central wheatbelt (H2), south-eastern coastal (H9) and central-southern coastal (H5) regions (Figure 1). Within the southern clade, a region of high haplotype diversity was identified along the southern coastline from Betty's Beach in the west to Parmango in the east. There were also haplotypes with more restricted distributions located inland on granite and greenstone outcrops.

### Haplotype and nucleotide diversity

Descriptive statistics averaged across all individuals and for individuals within the two clades using sequence data are listed in Table 3. Haplotype diversity was higher in the southern, more structured clade. Parameters of neutrality and population size fluctuation were both non-significant for populations from the southern clade following coalescent simulations (Table 3). For populations in the northern clade, the neutrality parameter Tajima's  $D$  was significant although Ramos-Onsins and Rozas  $R_2$  was not. The observed mismatch distributions of haplotypes from both clades did not significantly depart from those expected under sudden demographic and spatial expansion models (all  $P_{SSD}$  and  $P_{HRag} > 0.1$ ) (Supplementary Figure S2). Bayesian analysis of time to the most recent common ancestor indicated that haplotypes from the two clades coalesce at 2.24 Ma (95% hpd levels 1.2–4.2 Ma) using the slower mutation rate and 1.38 Ma (95% hpd levels 0.8–2.3 Ma) using the faster rate, placing the phylogeographic split of the northern and southern clade



**Figure 2** Bayesian cladogram of 30 *Calothamnus quadrifidus* haplotypes based on combined cpDNA sequence data, using *C. sanguineus* and *C. rupestris* as outgroups. Numbers on branches indicate posterior probabilities >0.95. The number in bold indicates the divergence estimate of the two clades with confidence intervals in brackets. Haplotypes are coloured according to Figure 1. Subspecies are listed. \*Peak Charles consists of subspecies *quadrifidus* and *seminudus*.



**Figure 3** Median-joining maximum parsimony network for all *Calothamnus quadrifidus* samples. Branch lengths represent the degree of mutational change. Node size is proportional to the frequency of haplotypes. Boxes enscribe clades as identified by phylogenetic analysis of combined cpDNA sequence data. Colours identify sequence haplotypes as shown in Figure 1.

during the Pleistocene. This estimate must be interpreted with caution given that we have only sampled genetic variation at three chloroplast loci. Comparison with the date obtained from RFLP data

(1.76 Ma; Supplementary Information), which sampled across the entire cpDNA genome, provides a valuable cross-check of congruence.

**Table 3** Nucleotide and haplotype diversity, tests for neutrality and demographic and spatial expansion and genetic differentiation estimates in *Calothamnus quadrifidus* and the two major clades therein, based on combined *rpl16*, *petB* and *trnQ-rps16* cpDNA sequences

	n	Haplotype diversity	Nucleotide diversity (%)	Tajima's D	Ramos-Onsins and Rozas $R_2$	$N_{ST}$	$G_{ST}$
All <i>C. quadrifidus</i>	203	0.884 ± 0.01	0.299 ± 0.08	-1.35	0.081	0.952 (0.018)*	0.877 (0.038)
Northern clade	74	0.765 ± 0.05	0.164 ± 0.06	-1.51*	0.102	0.777 (0.124)	0.800 (0.092)
Southern clade	129	0.788 ± 0.03	0.156 ± 0.05	-0.81	0.089	0.946 (0.022)*	0.904 (0.038)

n, number of samples; haplotype diversity based on the number of haplotypes excluding sites with gaps, s.d. shown; nucleotide diversity given as a percentage with s.d. shown,  $N_{ST}$  and  $G_{ST}$  with s.e. in parentheses; asterisk (\*) denotes a significant result  $P \leq 0.05$ .

### Species modelling

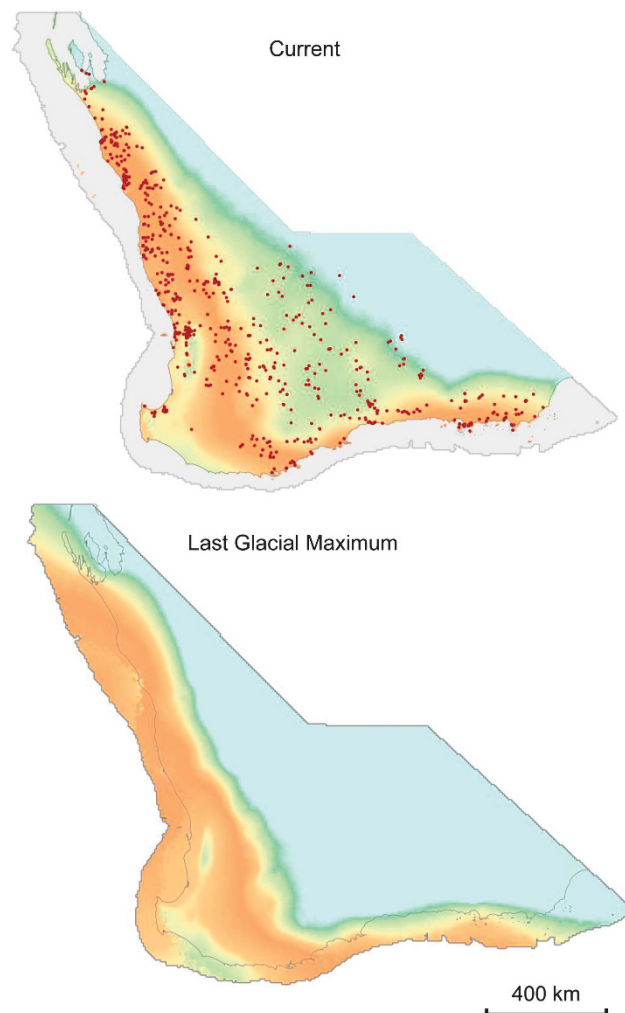
Given the distribution of the species, it was expected that rainfall and perhaps temperature would be important in defining the climate envelope. Preliminary modelling indicated two competing models that achieved area under the curve (AUC) scores of  $>0.76$ , these consisted of one or two variables related to rainfall, and both closely matched current distribution patterns with highest suitability of climate along the west and south coasts and lower suitability inland. The AUC for both models were very similar (the single-term model only resulted in an increase of 0.03) so the single-term model (AUC 0.755) based on Rainfall of coldest quarter was selected to project distribution back to LGM (Figure 4). The projected distribution at the LGM showed some contraction of the areas of greatest climatic suitability towards the west and south and elimination of lower suitability climatic areas inland. At the LGM, the coastline was approximately 30 km further west and south than at present, and projections show these areas to be generally climatically suitable for *C. quadrifidus*.

### DISCUSSION

Phylogeographic analysis of the widespread shrub *C. quadrifidus* has revealed two major clades in the SWAFR, each with areas of high haplotype diversity indicative of the location of major refugia. These refugia occur along the mid-west and southern coastal regions where mesic conditions would have prevailed during arid periods of the Pleistocene and are consistent with areas of climatic suitability derived from species distribution modelling during projected LGM climatic conditions. Other areas of low haplotype diversity suggest expansion of the species from these refugial areas. Although areas of refugial habitat during Miocene and Pliocene aridification have been previously postulated in the SWAFR based on species richness and endemism (Gardner, 1944; Marchant, 1973; Hopper 1979; Hopper *et al.*, 2006), this study provides the first evidence from cpDNA data for the location of major refugia during the Pleistocene for a widespread plant species. It demonstrates a different response to major climatic fluctuations compared with other species with non-coastal distributions, where studies show evidence for multiple localised refugia and little widespread expansion (Byrne, 2008).

### Divergent clades

Despite the topographically subdued nature of the broad landscape of the SWAFR, there is evidence for clade diversification within this species that may have been driven by aridity that has increased since the Pliocene and intensified during the climatic fluctuations of the Pleistocene (Bowler, 1982; Hopper and Gioia, 2004; Martin, 2006). The presence of two major clades in *C. quadrifidus* is consistent with higher level phylogenetic divergence, such as in *Conostylis* spp. (Hopper *et al.*, 2006), and with intraspecific phylogeographic patterns identified in other species in the region (see studies cited in Byrne, 2007, 2008). Like the phylogeographic patterns in the above species,



**Figure 4** Location of herbarium collections of *Calothamnus quadrifidus* and species distribution modelled using Maxent under (a) current climatic conditions and (b) Last Glacial Maximum conditions projected using CCSM estimates. Colours depict probability of presence based on logistic outputs, pale blue low (0–0.2) to orange high (0.6–0.65). Land extent during LGM shown in both maps.

the current distribution of the *C. quadrifidus* clades are not defined by specific topographical features and, although the clade boundaries vary between species, they generally occur along a north-west to south-east axis that broadly correlates with transition zones between mesic and more arid environments. Similar spatial structuring of genetic clades both in the absence of, or in conjunction with, topographical barriers has been noted throughout southern Australia in both flora and fauna (Byrne *et al.*, 2008, 2011), emphasising the

influence of long timescales and climatic variation in driving genetic differentiation even in the absence of topographical barriers to gene flow.

### Northern refugium

The two clades showed evidence of contrasting evolutionary histories with differences in haplotype structuring and genetic diversity. The northern clade, characterised by a star phylogeny, exhibited a high number of localised haplotypes centred about the Kalbarri–Shark Bay region, with lack of phylogeographic signal in this area due to the presence of unrelated haplotypes within populations, and distribution of a common haplotype along the coast to the south. This pattern suggests the presence of a refugium around the Kalbarri–Shark Bay region, and an area of expansion in the south of the distribution. The LGM species distributional modelling showed that climatic suitability for the species was maintained in the Kalbarri region, along the coast and off the current coastline in the mid-west. The Kalbarri area is characterised by an unusual series of gorge systems that provide environments likely to have facilitated the persistence of species throughout periods of climatic stress. In addition to the presence of specialised habitats, the coastal environment of the region would also have provided a buffering effect from the drying climate (Marchant, 1973; Hopper, 1979; Hopper *et al.*, 2006; Byrne *et al.*, 2008). The Shark Bay region is a centre of endemism characterised by high species richness and has been previously hypothesised to have been a refugial area (Keighery *et al.*, 2000). The high haplotype diversity observed in *C. quadrifidus* within this region provides support for this hypothesis. The broader distribution of one haplotype in the south of the current clade distribution implies expansion southward along the coast from this refugium. This spatial signature of expansion was supported with tests of neutrality as well as tests of spatial and demographic expansion using mismatch distributions. The species distribution model at the LGM supported moderate climate suitability along the western coastline. Thus the time of expansion may not be since the LGM but could have occurred earlier during previous interglacial cycles with subsequent persistence of an ancestral haplotype in this area.

Within the northern clade, one haplotype (H22) was highly divergent from the others and was restricted to the Wongan Hills area, 200 km north-east of Perth in a population that is morphologically consistent with the more widespread *C. quadrifidus* subsp. *angustifolius* that occurs on the surrounding sandplain. In addition, the one geographical anomaly in the clade distribution was the southern clade haplotype present in the subsp. *asper* population that occurs in the Wongan Hills area. This area is characterised by a complex soil profile and a series of greenstone outcrops that exhibit high floral diversity. Populations in this region may have persisted in isolation on these outcrops throughout fluctuating climate cycles, and such isolation would lead to divergence over time. There was no evidence of subsequent haplotype expansion from these populations.

### Southern refugium

In contrast to the northern clade, the southern clade was more genetically diverse and exhibited greater phylogeographical complexity, seen in the significantly higher  $N_{ST}$  than  $G_{ST}$  and in the high support for intra-clade clusters in the phylogenetic tree. The majority of this diversity was located along the southern coastline with areas of lower diversity found in the inland regions. Several divergent haplotypes were centred on areas at the eastern and western extremes of the clade distribution. The LGM species distributional modelling showed that the areas of climatic suitability for the species contracted

south and highest areas of suitability occurred off the current southern coastline and in the western area. Note that a large part of the area in the west of the southern region is forested and, although it is identified as climatically suitable, it does not support *C. quadrifidus*. The south coast region is characterised by high species richness (Hopper and Gioia, 2004), and some studies have identified high genetic diversity in species that occur in higher elevation areas in the ranges that occur along the coastline, for example, trapdoor spiders (Cooper *et al.*, 2011), assassin spiders (Rix and Harvey, 2012) and the plants, *Acacia verricula* (Byrne *et al.*, 2001) and *Banksia brownii* (D. Coates unpublished data). The high differentiation of haplotypes along the coastal region in *C. quadrifidus* is consistent with persistence in isolated patches along the current coastline. The widespread distribution of Haplotype 5 from the southern coast across to the Perth area occurs in areas that appeared to maintain suitable climate at the LGM. This suggests expansion may have occurred earlier than the LGM or that factors other than climate may have influenced persistence through these time frames.

The low haplotype diversity north of the south coast provides evidence of inland expansion that is also supported by signals of expansion in tests of mismatch distributions. Expansion of two different haplotypes inland suggests that either these re-colonisations occurred at similar times but from different populations or that they occurred from similar areas but across different time frames. However, the species distribution model suggests the inland areas did not maintain suitable climatic conditions at the LGM providing support for the hypothesis of expansion of the species into these areas after the LGM from isolated patches along the coastline.

The divergent haplotypes located in the extremes of the distribution were present in populations that were restricted to the greenstone/granite formations in the east (Bencubbin, Tandagin, King Rock, Diggers Rock; *C. quadrifidus* subsp. *petraeus* and subsp. *seminudus*) and on ironstone formations in the west (Negus, Boallia; *C. quadrifidus* subsp. *teretifolius*). The greenstone/granite and ironstone formations are specific habitats that retain mesic conditions during more arid phases (Gibson *et al.*, 2010). The LGM climate projections suggest the area of the ironstone formations in the west retained suitable climate. In contrast, the climate modelling showed unsuitable LGM climate conditions in the eastern area, suggesting that greenstone/granite formations have facilitated long-term persistence of this species in this area. Opportunities for expansion from these areas of specific habitat may be more limited, and there was no evidence in *C. quadrifidus* that expansion has occurred from these populations.

### Morphological variation

The *C. quadrifidus* complex is morphologically variable but shows consistent geographical patterns. Depending on the historical drivers, it would not be expected that all morphological variation would be reflected in phylogeographic patterns, although some of the major influences might be detected. However, there is broad congruence between the phylogeographic patterns and the morphological variation recognised in the recent revision of this species (George and Gibson, 2010). Haplotypes were specific to subspecies with the exception of one central haplotype that was shared between the widespread subsp. *quadrifidus* and the restricted subsp. *seminudus* at one population. *C. quadrifidus* subsp. *asper*, which was previously recognised as a separate species and sunk as a subspecies in the taxonomic revision, is genetically embedded in the species complex and is not as divergent as the closely related outgroups *C. rupestris* and *C. sanguineus*. With the exception of the widespread



*C. quadrifidus* subsp. *quadrifidus*, morphological subspecies were confined to either the northern or southern genetic clades, suggesting that their differentiation occurred after divergence of the clades. Subspecies *quadrifidus* is morphologically variable, but division within this subspecies was not able to be defined on morphological traits. The identification of two genetic groups within this subspecies provides a basis to revisit this evaluation to see whether subtle morphological characters exist that support further separation. Analysis of nuclear genetic diversity within this subspecies may also assist in resolving the complexity.

The northern clade largely corresponds to *C. quadrifidus* subsp. *obtusus*, subsp. *homalophyllus* and subsp. *angustifolius*, with all three subspecies occurring in the putative refugial area. This current phylogeography at the broad scale does not give any indication of drivers of diversification among these subspecies. The single widespread haplotype (H7) occurs in a number of subspecies, and its central position in the network suggests incomplete lineage sorting of haplotypes from a common ancestor, although admixture may also be a possible explanation. The southern clade includes the widespread subsp. *quadrifidus* (also found in the southern edge of the northern clade), represented by the widespread haplotypes (H5, H9 and H6), as well as the more eastern subsp. *petraeus* and subsp. *seminudus* and the outlying restricted subsp. *teretifolius* and subsp. *asper* that is geographically restricted to a steep slope in the Wongan Hills area. Neither subsp. *petraeus* nor subsp. *teretifolius* share haplotypes with other taxa, consistent with their geographic isolation and habitat specificity. Habitat specificity may have contributed to diversification of subsp. *teretifolius*, *petraeus* and *seminudus*. The general pattern of congruence in the patterns of haplotype diversity and morphological differences provides support for hypotheses of historical processes being a major driver of speciation (Hopper, 1979; Hopper and Gioia, 2004).

#### cpDNA intraspecific phylogeographic patterns

Intraspecific phylogeographic analysis in southern Australia has been very informative in relation to evolutionary patterns and process. Early phylogeographical studies of plant species in this landscape were undertaken using RFLP technology. The general congruence between the RFLP and sequence data sets in resolving clade relationships in this species enables valid comparisons with previous studies using RFLP methodology (see studies cited in Byrne, 2007, 2008).

Although the distinction of two genetic clades in *C. quadrifidus* is similar to that detected in other species (Byrne, 2008), the distribution of haplotype diversity in this species is in contrast to other species where haplotype diversity shows a pattern of localised and highly specific geographic distribution that is indicative of localised, dispersed 'micro' refugia distributed throughout the species range (Byrne, 2008). The reasons for these different patterns is not clear but may be related to species biology or habitat specificity. The identification of major refugia in this study may be, in part, because for the first time we have targeted a species with a wide distribution that covers almost the entire SWAFR, including the coastal regions.

#### CONCLUSIONS

This is the first molecular evidence for major refugial areas in SWAFR, on the mid-west coast and along the southern coast, and contrasts with previous studies that showed evidence for multiple localised refugia throughout species distributions. Phylogeographic analysis of *C. quadrifidus* revealed two genetic clades to the north and south of the species' distribution that are estimated to have diverged in the Pleistocene. The two clades show evidence of contrasting evolutionary histories, with the northern, less structured clade possessing a

geographically restricted core of genetic diversity centred on the mid-west coastal region of Kalbarri–Shark Bay, a suspected refugium for the species. The southern, more structured and diverse clade showed high levels of genetic diversity along the wetter habitats of the south coast, which may indicate a broad refugial area. Identification of response to historical processes has important implications in future conservation efforts as the identification of refugia, and the genetic diversity they harbour, enables these areas to be targeted for protection in climate change adaptation strategies (Keppel *et al.*, 2012).

#### DATA ARCHIVING

Sequence data have been submitted to GenBank; accession numbers are listed in Table 1.

#### CONFLICT OF INTEREST

The authors declare no conflict of interest.

#### ACKNOWLEDGEMENTS

We thank David Coates and Anne-Laure Jacquemart for assistance with field collections, Eleftheria Dalmaris for assistance with the HAPLO program analysis and J Dale Roberts for comments on the manuscript.

- Ashcroft MB, Gollan JR, Warton DI, Ramp D (2012). A novel approach to quantify and locate potential microrefugia using topoclimate, climate stability, and isolation from the matrix. *Glob Change Biol* **18**: 1866–1879.
- Beheregaray LB (2008). Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. *Mol Ecol* **17**: 3754–3774.
- Bowler J (1982). Aridity in the late Tertiary and Quaternary of Australia. In: Barker WR, Greenslade PJM (eds). *Evolution of the Flora and Fauna of Arid Australia*. Peacock Publications: Frewville, SA, Australia, pp 35–45.
- Byrne M (2007). Phylogeography provides an evolutionary context for the conservation of a diverse and ancient flora. *Aust J Bot* **55**: 316–325.
- Byrne M (2008). Evidence for multiple refugia at different time scales during Pleistocene climatic oscillations in southern Australia inferred from phylogeography. *Quat Sci Rev* **27**: 2576–2585.
- Byrne M, Hankinson M (2012). Testing the variability of chloroplast sequences for plant phylogeography. *Aust J Bot* **60**: 569–574.
- Byrne M, Moran GF (1994). Population divergence in the chloroplast genome of *Eucalyptus nitens*. *Heredity* **73**: 18–28.
- Byrne M, Steane D, Joseph L, Yeates DK, Jordan GJ, Crayn DM *et al.* (2011). Decline of a biome: evolution, contraction, fragmentation, extinction and invasion of the Australian mesic zone biota. *J Biogeogr* **38**: 1635–1656.
- Byrne M, Tischler G, Macdonald B, McComb J, Coates D (2001). Phylogenetic relationships between two rare *Acacias* and their common widespread relatives in south-western Australia. *Conserv Genet* **2**: 157–166.
- Byrne M, Yeates DK, Joseph L, Kearney M, Bowler J, Williams MAJ *et al.* (2008). Birth of a biome: insights into the assembly and maintenance of the Australian arid zone biota. *Mol Ecol* **17**: 4398–4417.
- Clement M, Posada D, Crandall K (2000). TCS: a computer program to estimate gene genealogies. *Mol Ecol* **9**: 1657–1660.
- Cooper SJB, Harvey MS, Saint KM, Main BY (2011). Deep phylogeographic structuring of populations of the trapdoor spider *Moggridgea tingle* (Migidae) from southwestern Australia: evidence for long-term refugia within refugia. *Mol Ecol* **20**: 3219–3236.
- Drummond AJ, Rambaut A (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* **7**: 214.
- Elith J, Phillips SJ, Hastie T, Dudik M, Chee YE, Yates CJ (2011). A statistical explanation of MaxEnt for ecologists. *Divers Distrib* **17**: 43–57.
- Excoffier L, Lischer HEL (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Res* **10**: 564–567.
- Fujioka T, Chappell J, Fifield LK, Rhodes EJ (2009). Australian desert dune fields initiated with Pliocene-Pleistocene global climatic shift. *Geology* **37**: 51–54.
- Gardner CA (1944). The vegetation of Western Australia with special reference to the vegetation and soils. *J Roy Soc WA* **28**: 11–87.
- George AS, Gibson N (2010). A revision of *Calothamnus quadrifidus* (Myrtaceae). *Nuytsia* **20**: 57–77.
- Gibson N, Yates CJ, Dillon R (2010). Plant communities of the ironstone ranges of South Western Australia: hotspots for plant diversity and mineral deposits. *Biodiv Conserv* **19**: 3951–3962.
- Graur D, Li WH (2000). *Fundamentals of Molecular Evolution*, 2nd edn. Sinauer Associates Inc.: Sunderland, MA, USA.

- Hall TA (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp* **41**: 95–98.
- Hewitt GM (2000). The genetic legacy of the Quaternary ice ages. *Nature* **405**: 907–913.
- Hewitt GM (2004). Genetic consequences of climatic oscillations in the Quaternary. *Phil Trans R Soc Lond B* **359**: 183–195.
- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005). Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol* **25**: 1965–1978.
- Hopper SD (1979). Biogeographical aspects of speciation in the southwest Australian flora. *Ann Rev Ecol Syst* **10**: 399–422.
- Hopper SD (2000). Floristics of Australian granitoid inselberg vegetation. In: Porembeski S, Barthlott W (eds). *Inselbergs*. Springer-Verlag: Berlin, Germany, pp 391–407.
- Hopper SD, Chase MW, Fay MF (2006). A molecular phylogenetic study of generic and subgeneric relationships in the south-west Australian endemics *Conostylis* and *Blancoa* (Haemodoraceae). *Monocots: Comp Biol Evol Aliso* **22**: 527–538.
- Hopper SD, Gioia P (2004). The southwest Australian floristic region: evolution and conservation of a global hot spot of biodiversity. *Ann Rev Ecol Syst* **35**: 623–650.
- Keighery GJ, Gibson N, Lyons MN, Burbidge AH (2000). Flora and vegetation of the southern Canarvon Basin, Western Australia. *Rec WA Mus Supp* **61**: 77–154.
- Keppel G, Van Niel KP, Wardell-Johnson GW, Yates C, Byrne M, Mucina L *et al.* (2012). Refugia: identifying and understanding safe havens for biodiversity under climate change. *Glob Ecol Biogeogr* **21**: 393–404.
- Librado P, Rozas J (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**: 1451–1452.
- Marchant NG (1973). Species diversity in the south western flora. *J Roy Soc WA* **56**: 23–30.
- Martin HA (2006). Cenozoic climatic change and the development of the arid vegetation in Australia. *J Arid Environ* **66**: 533–563.
- Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J (2000). Biodiversity hotspots for conservation priorities. *Nature* **403**: 803–808.
- Phillips SJ, Anderson RP, Schapire RE (2006). Maximum entropy modeling of species geographic distributions. *Ecol Model* **190**: 231–259.
- Pons O, Petit RJ (1996). Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics* **144**: 1237–1245.
- Posada D (2008). jModelTest: phylogenetic model averaging. *Mol Biol Evol* **25**: 1253–1256.
- Ramos-Onsins SE, Rozas J (2002). Statistical properties of new neutrality tests against population growth. *Mol Biol Evol* **19**: 2092–2100.
- Rix MG, Harvey MS (2012). Phylogeny and historical biogeography of ancient assassin spiders (Araneae: Archaeidae) in the Australian mesic zone: evidence for Miocene speciation within tertiary refugia. *Mol Phyl Evol* **62**: 375–396.
- Swofford DL (2003). *PAUP\* Phylogenetic Analysis Using Parsimony (\*and Other Methods)*. Version 4. Sinauer Associates: Sunderland, MA, USA.
- Tajima F (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585–595.
- Templeton AR, Crandall KA, Sing CF (1992). A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* **132**: 619–633.
- Wolfe KH, Li WH, Sharp PM (1987). Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proc Natl Acad Sci USA* **84**: 9054–9905.

Supplementary Information accompanies this paper on Heredity website (<http://www.nature.com/hdy>)