Structural Equation Modelling Analysis of Evolutionary and Ecological Patterns in Australian *Banksia*

Tianhua He

Department of Environment and Agriculture, Curtin University, PO Box U1987, Perth, WA 6845, Australia

Correspondence: Tianhua He, Department of Environment and Agriculture, Curtin University, PO Box U1987, Perth, WA 6845, Australia.

Phone +61 8 9266 7005. Fax +61 8 9266 2495.

Email: Tianhua.He@curtin.edu.au

Number of text pages: 18

Number of figures: 1

Number of tables: 1

Number of electronic supplementary tables: 2

Abstract

Evolutionary history of species, their geographic ranges, ecological ranges, genetic diversity, and resistance to pathogen infection, have been viewed as being mutually linked through a complex network of interactions. Previous studies have described simple correlations between pairs of these factors, while rarely separated the direct effects among multiple interacting factors. This study was to separate the effect of multiple interacting factors, to reveal the strength of the interactions among these factors, and to explore the mechanisms underlying the ecological and evolutionary processes shaping the geographic range, genetic diversity and fitness of species. I assembled comparative data on evolutionary history, geographic range, ecological range, genetic diversity, and resistance to pathogen infection for thirteen *Banksia* species from Australia. I used Structural Equation Modelling (SEM) to test multivariate hypotheses involving evolutionary history, geographic range, genetic diversity and fitness. Key results are: 1) Species with longer evolutionary times tend to occupy larger geographic ranges; 2) higher genetic diversity is directly associated with longer flowering duration in *Banksia*; and 3) species with higher genotypic diversity have higher level of resistance to infection caused by the pathogen *Phytophthora cinnamomi*, whereas heterozygosity has the opposite relationship with capacity of resistance to the infections caused by this pathogen. These results revealed a mutually linked and complex network of interactions among gene, species, environment and pathogen in evolutionary and ecological scales. These findings also have great practical significance and help to provide preemptive management options in pathogen control.

Keywords: Evolutionary history • Flowering duration • Genetic diversity • Geographic ranges • Path Analysis • Pathogen resistance

Introduction

Understanding the factors that best explain variation in geographic ranges among species is one of the central goals at the interface of ecology and evolution (Pigot et al. 2012). Research to date has emphasised the theory of temporal dispersal limitation, which predicts a positive relationship between species age and range size. The key prediction is the age-area hypothesis proposed by Willis (1922): if species initially possess small population sizes and restricted geographic ranges, then many species with restricted geographic ranges could simply be young species. A similar prediction, the dispersal-assembly theory, is based on neutral theory in community ecology (Hubbell 2001). A number of recent studies have shown that the relative geographic range of a species does appear to vary predictably with the evolutionary age of the species (e.g., Liow and Stenseth 2007). Although such "age-area" correlations have been tested in a wide variety of groups, no single model of range evolution appears to apply across groups. Most likely, age-area correlations are clade specific and may depend on dispersal ability of a group (Webb and Gaston 2000). The geographic range of a species is also assumed to be linked to the species' genetic diversity (Sexton et al. 2009). Species with narrow distributions tend to have a lower level of genetic diversity than their widespread congeners (e.g., He et al. 2000; Souza and Lovato 2010). Adding to the complexity of these interactions, the levels of molecular genetic diversity reflect the evolutionary histories of populations and species. A high level of genotypic diversity for nuclear genes may reflect a long evolutionary history. For example, human populations in Africa typically have higher levels of genetic diversity (Tishkoff and Verrelli 2003).

Genetic variation and fitness are linked (Szulkin et al. 2010). Genetic heterozygosity is often assumed to co-vary positively with fitness, generating positive heterozygosity-fitness correlations

(Forestmeier et al. 2012). Previous studies of the relationship between genetic polymorphism and the response to exposure to infectious disease have also predicted a negative correlation between the host's genetic diversity and the consequences of exposure to a pathogen (Radwan et al. 2010; Townsend et al. 2010). A similar assumption made in conservation genetics is that there is a causal relationship between genetic variability and the evolutionary adaptability of a species (Allendorf et al. 1997). However, the connections between neutral genetic diversity, fitness, and adaptability are not always straightforward, and such causal relationships between genetic variability and the evolutionary adaptability of a species have not always been consistent with empirical observations (e.g., He and Lamont 2010).

As indicated above, the evolutionary history of species, geographic ranges, ecological adaptability, genetic diversity, and resistance to pathogen infection have all been proposed and observed to be mutually linked by a complex network of interactions. Conventional bivariate correlations between any two of these factors have been reported in the literature, and strong correlations have been consistently reported. However, few attempts have been made to separate the direct effects occurring among multiple interacting factors, except for those in which one factor was manipulated (e.g., Crawford and Whitney 2010). In this study, I used Structural Equation Modelling (SEM) to test multivariate hypotheses involving evolutionary history, geographic ranges, ecological ranges, genetic diversity and fitness in *Banksia*, an iconic genus in Australian ecosystems. I aimed to separate the effects of multiple interacting factors, to reveal the strength of the interactions among these factors, and ultimately to explore the mechanisms shaping the ecological and evolutionary processes in this genus.

Material and methods

Study taxa

Banksia is a genus of approximately 170 species in the plant family Proteaceae (Mast and Thiele 2007), and ranges from prostrate woody shrubs to trees up to 10 m tall. These plants are generally found in a wide variety of landscapes, including sclerophyll forest, (occasionally) rainforest, shrubland, and several additional arid landscape types (Lamont et al. 2007). *Banksia* species are present throughout regions that supply suitable rainfall (annual rainfall > 200 mm) and usually represent the most prominent taxonomic group in the extensive species-rich sandplain flora of southwestern Australia (Lamont et al. 2007). Much of the current knowledge in ecology of *Banksia* concerns patterns of diversity (Cowling and Lamont 1998; Lamont et al. 2007; Merwin et al. 2012) rather than underlying ecological and evolutionary processes (He et al. 2011).

Banksia's proteoid root, which helps it to survive in low-nutrient soils, makes it highly susceptible to infection by *Phytophthora cinnamomi*. The disease caused by the plant pathogen *P. cinnamomi* has been identified as a key threat to biodiversity in the Australian environment (Environment Australia 2001). Almost all (approximately 96%) of the species of Proteaceae rated as priority taxa in Western Australia are susceptible to *P. cinnamomi* (Wills and Keighery 1994). Of greater interest is the extent to which the capacity of particular species to resist this pathogen can be predicted.

Compilation of genetic data

Microsatellite genetic diversity (genotypic richness, N_a, and expected heterozygosity, H_E) was determined for 13 *Banksia* species, species list was given in Table S1 in Electronic Supplementary

Material (ESM). Genetic diversity data for *B. attenuata*, *B. candolleana*, *B. ilicifolia and B. menziesii* were generated for this study. Source data for genetic diversity in nine other species are given in Table S1 in ESM. Genetic data of widespread species were derived from multiple populations collected across its distribution range, and represented the overall genetic diversity of the species.

For *B. attenuata, B. candolleana, B. ilicifolia* and *B. menziesii,* at least 30 samples from each of ten populations were genotyped. The samples were collected from coastal sandplain habitats in Western Australia. Ten to eleven microsatellite primers were evaluated for each species. These microsatellite primers were specifically developed for the species. Primer information and the genotyping protocol for *B. attenuata* were as described by He et al. (2007), and the corresponding information for *B. candolleana* is given in Merwin et al. (2010). Microsatellite primers for *B. ilicifolia* and *B. menziesii* were developed following He et al. (2007) and Merwin et al. (2010). Primer details and genotyping protocols are presented in Table S2 in ESM. Analyses of genetic diversity were implemented in GenAlEx (Peakall and Smouse 2006) for 30 samples per population. For the remaining nine species, the genotypic diversity (N_a) and heterozygosity (H_E) values were taken from the literature or previous study (Table S1 in ESM) but were re-calculated for the sample size of 30 per population using rarefaction analysis (Petit et al. 1998) to eliminate the effect of unequal sample sizes in the cross-species comparisons.

Compilation of evolutionary, ecological and geographic data

The evolutionary times of *Banksia* species were represented by the species ages reported in He et al. (2011). The geographic range of a species was calculated as the area of its geographic distribution by drawing a box around the range, with the longest dimension parallel to the line

between the two most distant points of the range (the distributions of all species were surveyed between 1984 and 1986, Taylor and Hopper 1991).

The parameter termed "ecological range" includes the following four indices: 1) the soil type diversity, measured as the number of soil types occurring in a species' geographical range; 2) the vegetation type diversity, measured as the number of vegetation types in which a species occurs; 3) the landform diversity, measured as the number of major landforms in which a species occurs; and 4) the rainfall range, measured as the range of annual rainfall across the species' geographic range.

The life history and phenological parameter includes the following five indices: 1) post-fire regeneration (resprouting or killed by fire); 2) the growth form (tree, shrub, or both); 3) the pollinator diversity, measured as the number of major pollinator categories (birds, mammals, bees, other insects); 4) the flowering duration, measured as the number of months per year during which a species has open flowers; and 5) the growth duration, measured as the number of months during which a species grows new foliage. Data on the indices for the ecological range, life history and phenology represented variation across the species' range, and were extracted from the comprehensive *Banksia Atlas* (Taylor and Hopper 1991) and from *The Banksia* book (Collins et al. 2008), *The Dryandra* book (Cavanagh and Pieroni 2006) and FloraBase (www.florabase.dec.wa.gov.au).

Banksia species show a wide spectrum of susceptibility to the fungal pathogen *P. cinnamomi*, ranging from full resistance to approximately 100% death after inoculation with this pathogen (McCredie et al. 1985). Data on resistance to *P. cinnamomi* infection were obtained from McCredie et al. (1985) and were calculated as "1 – mortality in 96 days after inoculation" (the highest death rate occurs 35-40 days after inoculation, McCredie et al. 1985). *Banksia* species from eastern Australia are generally resistant to *P. cinnamomi* under Western Australian field

163 conditions, perhaps because certain *Banksia* species from eastern Australia are associated with
164 different pathogen species *P. citricola* (Tynan et al. 1998). For this reason, only resistance data for
165 Western Australian *Banksia* (nine species) were used for the analysis.
166
167 Data analysis
168
169 Conventional bivariate correlation analyses (defined by correlations between pairs of variables)
170 were first implemented between all pairwise parameters/indices. The SEM was then used to detect
171 direct interaction between variables and to test the specific hypotheses proposed. The SEM aims
172 to reveal the possible ecological and evolutionary processes governing the inter-correlated
173 network. The advantage of SEM is allowing the researchers to specify the pathway in the model
174 that represents a working hypothesis thought to reflect the essential casual mechanisms (Mitchell
175 1992; Grace et al. 2010). The modelling process in SEM analysis is guided by a priori and
176 theoretical knowledge and begins with a consideration of expected relationships based on the
177 mechanisms thought to operate in the system. The conceptual SEM models of the expected
178 multivariate relationship were built based on theoretically developed hypotheses.

179 Seven parameters for each species were considered in SEM models: 1) evolutionary time, 2)
180 geographic range, 3) ecological range, 4) genotypic diversity, 5) genetic heterozygosity, 6) life
181 history and phenotypic diversity, and 7) resistance to pathogen infection. Although the parameter
182 of “ecological range” initially includes four indices (see Method: Compilation of evolutionary,
183 ecological and geographic data), only rainfall range was shown significantly being correlated with
184 other parameter in bivariate correlation analysis. Therefore, rainfall range was introduced into
185 SEM as surrogate for parameter of “ecological range”. Likewise, flowering duration was used in

SEM as surrogate for parameter of "life history and phenotypic diversity". The following hypotheses were tested:

H_1: A longer evolutionary time allows a species to disperse over a larger geographic range.

H_2: A longer evolutionary time allows the accumulation of greater genotypic diversity and increased heterozygosity.

H_3: Life history and phenotypic diversity are positively correlated with genotypic diversity.

H_4: Genetic diversity (both genotypic diversity and heterozygosity) is positively correlated with the ability to resist pathogen infection.

Two SEM models were built to maximize statistic power in a system with small number of observations (Fig. 1). Model A included five parameters (resulting a ratio of observation to parameter of 2.6), and test hypothesis H_1, H_2 and H_3. Model B test H_4 with four parameters (resulting a ratio of observation to parameter of 3.2). Model estimation in SEM was based on Bayesian Markov Chain Monte Carlo (MCMC) procedure with a uniform prior distribution. Bayesian estimation derives a posterior estimation of standard regression weight that summarises the state of knowledge about the parameter, and offer good performance in models with small number of observations. The goodness of fit of the model to the data was evaluated with the model posterior predictive probability. A posterior predictive *P* value near 0.5 indicates a correct model (Gelman et al. 2004). In each Bayesian MCMC estimation, the first 500 generations were discarded, and 100,000 were retained in further analysis.

The conventional bivariates correlation analyses were implemented using SPSS (Statistical Package for the Social Sciences, SPSS Inc. Chicago) software, and $P < 0.05$ was considered to indicate statistical significance. SEM was implemented with path analysis in SPSS AMOS (SPSS Inc. Chicago).

Results

Bivariate analysis

The evolutionary time of *Banksia* species was positively correlated with the geographic range of the species ($P = 0.048$), but not with the genotypic diversity ($P = 0.104$) (Table 1). *Banksia* species with larger geographic ranges tended to have higher levels of genotypic diversity ($P = 0.026$). Species occurring in areas spanning a larger range of annual rainfall were also found to have higher levels of genotypic diversity, though marginally significant ($P = 0.050$) (Table 1). *Banksia* species with longer flowering times also had higher levels of genotypic diversity ($P = 0.012$) (Table 1). A higher level of resistance to *P. cinnamomi* was found in *Banksia* species with higher levels of genetic diversity ($P = 0.008$ for N_a, $P = 0.082$ for H_E). *Banksia* species with longer flowering durations were also more resistant to *P. cinnamomi* infection ($P = 0.034$). Moreover, *Banksia* species occurring in areas spanning a larger range of annual rainfall also tended to be more resistant to this pathogen, though the correlation was not significant ($P = 0.062$).

Structural Equation Modelling results

SEM analysis using Bayesian MCMC procedures revealed direct correlation in multiple-factor situations. Both models shown in Fig. 1 produced a good fit to the data with both posterior predictive *P* values = 0.5. SEM analysis revealed strong direct and positive correlation between

231 evolutionary time and geographic range, confirming that *Banksia* with longer evolutionary times
232 occupied larger geographic areas (Fig. 1). Direct correlation between evolutionary time and
233 genotypic diversity was moderate with standard regression weight of 0.192; while the direct
234 influences of evolutionary time on rainfall range and flowering duration were non-existent.
235 Geographic range and rainfall range had a weak to no direct correlation with genotypic diversity.
236 Species with wider rainfall ranges was direct correlated to a larger geographic distribution with
237 standard regression weight of 0.868 (Fig. 1). Genotypic diversity in *Banksia* was significantly
238 correlated with the flowering duration, and species with longer flowering duration tend to have a
239 higher number of alleles in populations (Fig. 1). Species heterozygosity was significantly
240 correlated with genotypic diversity, while not with any other parameters. SEM analysis revealed
241 that resistance to *P. cinnamomi* infection was directly correlated with genetic diversity. Direct
242 correlation between genotypic diversity and the capacity to resist pathogen infection was
243 significant with standard regression weight of 1.092. Surprisingly, *Banksia* species with more
244 heterozygous genetic variation were more susceptible to *P. cinnamomi* infection, and the direct
245 effect was strong with standard regression weight of - 0.648 (Fig. 1).

246

247 **Discussion**

248

249 A positive and direct correlation between evolutionary age and geographic range was found in
250 *Banksia*, and the hypothesis H_1 (i.e., a longer evolutionary time allows species to disperse over a
251 larger geographic range) is therefore supported. Evidence of a positive relationship between
252 species age and range size has been found in a few taxon groups (e.g., Bohning-Gaese et al. 2006;
253 Paul et al. 2009; Hopkins 2011). A general pattern of an increase in range sizes immediately post-
254 speciation followed by a subsequent decline towards extinction was not supported in *Banksia*.

Most likely, *Banksia* species are relatively young (1 – 18 million *Y*, He et al. 2011) in terms of the average longevity for angiosperms (50 million *Y*, Crisp and Cook 2011). Equally, the sample size (13 species) may be insufficient to reveal the full age-range curve.

Range expansions are driven by dispersal. For this reason, variation in dispersal ability has been hypothesised to explain much of the variation in range sizes among species (Gaston 2003). A general expectation is that species with superior dispersal abilities attain larger range sizes more rapidly (Brown et al. 1996). However, a recent review suggests that dispersal ability may not be particularly important in driving variation in range size in many species (Lester et al. 2007), a principle that appears to hold for *Banksia*. Indeed, the seeds of *B. hookeriana*, a species with a relatively narrow distribution (1.2×10^3 km^2), were found to be as mobile as those of the most widely distributed species, *B. attenuata* (1.4×10^5 km^2). Similar rates and spatial scales of long-distance dispersal of seeds were found for *B. hookeriana* (long distance dispersal rate of 5-6% with spatial scale up to 3 km; He et al. 2004, 2010) and for *B. attenuata* (rate of 6% with spatial scale up to 3 km; He et al. 2009). Hence, the time available for dispersal may imply a central role in explaining range size variation.

The current analysis did not strongly support hypothesis H_2, which asserted that longer evolutionary times cause the accumulation of higher levels of genotypic diversity and increases in heterozygosity in *Banksia*. The evidence clearly supported hypothesis H_3 (Life history and phenotypic diversity, surrogated by flowering duration, are positively correlated with genotypic diversity). Co-occurring *Banksia* species usually have non-overlapping flowering periods (Lamont et al. 2003), and flowering durations in *Banksia* species have been found to show greater variation between species than the duration of shoots extension (Taylor and Hopper 1991). Flowering time is highly genotype-dependent (Pors and Werner 1989). The ability to flower over a longer season is an important aspect of adaptation. Moreover, *Banksia* species are primarily pollinated by birds,

including a wide range of honeyeaters (Meliphagidae) (Collins and Spice 1986), and small mammals (Wooller and Wooller 2001). It is probable that a longer flowering duration will serve to attract sufficient pollinators and facilitate extensive pollen flow, which could, in turn, promote greater genetic diversity by increasing the probability of dispersal and therefore preserve rare alleles.

Genotypically diverse *Banksia* species are more resistant to *P. cinnamomi* infection. The SEM analysis showed that the number of genotypes occurring in a species is closely related to the level of resistance to this pathogen in Western Australian *Banksia* species, supporting suggestions that the processes of selection associated with host and parasite dynamics are strongly dictated by host genetic variability (Teacher et al. 2009). The hypothesis H_4 (genetic diversity is positively correlated with the ability to resist pathogen infection in *Banksia*) is partially supported by the findings. When determining the resistance to *P. cinnamomi*, McCredie et al. (1985) used an isolate of *P. cinnamomi* in the inoculation. A genotypically diverse species may contain fewer individuals susceptible to the particular strain of pathogen used in the experiment. However, further researches are required to determine the actual genetic mechanisms involved. My results for *Banksia* support the assumption in the field of conservation genetics that there is a causal relationship between genetic variability and adaptability to a changing disease (Teacher et al. 2009).

Part of the hypothesis H_4 was not supported because my analysis showed that the ability of *Banksia* to resist *P. cinnamomi* infection was negatively correlated with the species' microsatellite heterozygosity. Although positive heterozygosity-fitness correlations have been reported for several plant and animal taxa (Grueber et al. 2008), our result supports the notion that these positive correlations are not universal (Chapman et al. 2009; Olano-Marin et al. 2011). The negative heterozygosity–fitness correlations could arise from outbreeding depression, either as the result of the breakdown of co-adapted gene complexes through recombination (Lynch and Walsh

1998) or through local adaptation (Szulkin and David 2011). However, outbreeding depression is likely not the cause of the negative correlation between heterozygosity and pathogen resistance found in *Banksia.* Most species in the genus of *Banksia* are almost completely outcrossed with extensive inter-population gene flow (Barrett et al. 2005; Krauss et al. 2009), and recombination may be frequent. As a result, co-adapted gene complexes is likely uncommon. It is probable that the genes for *P. cinnamomi* resistance in *Banksia* are homozygous in resistant plants; that is, only plants homozygous for the resistance genes are able to resist infection. Homozygotes have been reported to be superior to heterozygotes in some studies (Lavie and Nevo 1986; Nevo et al. 1986;). However, the actual biochemical mechanisms responsible for the greater resistance of homozygotes to the pathogen in *Banksia* remain to be determined.

Phytophthora cinnamomi has been of major concern in Australia with more than 1000 native plant taxa are known to be susceptible to infection by *P. cinnamomi* (Shearer et al. 2007). The relationship between the presence of *P. cinnamomi* and the onset of disease is complex due to the considerable variability within and among native plant species in their responses to the presence of *P. cinnamomi* (Shearer et al. 2007). The analysis showed that Western Australian *Banksia* with longer flowering durations tend to be more resistant to *P. cinnamomi* infection, though this effect was mediated by genotypic diversity. This finding has great practical significance and helps to provide preemptive management options.

Acknowledgements

This research was supported by the Australian Research Council (LP100100620 and DP120103389). I acknowledge the support from a Curtin Fellowship during the writing stage of this work. I am indebted to Paul Gioia of the Department of Environment and Conservation (Western Australia) for providing access to the raw data in *Banksia Atlas*, and to JB Wilson and an

anonymous reviewer for constructive comments on earlier versions of the paper.

References

Allendorf FW, Bayles D, Bottom DL, Currens KP, Frissell CA, Hankin D, Lichatowich JA, Nehlsen W, Trotter PC, Williams TH (1997) Prioritizing Pacific salmon stocks for conservation. Conserv Biol 11:140-152

Barrett LG, He T, Lamont BB, Krauss SL (2005) Genetic variation within the aerial seed bank of the shrub *Banksia hookeriana* (Proteaceae). Mol Ecol 14:4169-4179

Bohning-Gaese K, Caprano T, van Ewijk K, Veith M (2006) Range size: disentangling current traits and phylogenetic and biogeographic factors. Am Nat 167:555-567

Brown JH, Stevens GC, Kaufman DM (1996) The geographic range: size, shape, boundaries, and internal structure. Ann Rev Ecol Syst 27:597-623

Cavanagh T, Pieroni M (2006) *The Dryandras*. Australian Plants Society Inc. Perth, WA.

Chapman JR, Nakagawa S, Coltman DW, Slate J, Sheldon BC (2009) A quantitative review of heterozygosity-fitness correlations in animal populations. Mol Ecol 18:2746-2765

Collins BG, Spice J (1986) Honeyeaters and the pollination biology of *Banksia prionotes* (Proteaceae). Aust J Bot 34: 175-185

Collins K, Collins C, George AS (2008) Banksias. Bloomings Books. Melbourne

Cowling RM, Lamont BB (1998) On the nature of Gondwanan species flocks: diversity of Proteaceae in mediterranean south-west Australia and South Africa. Aust J Bot 46:335-355

Crawford KM, Whitney KD (2010) Population genetic diversity influences colonization success. Mole Ecol 19:1253-1263

352 Crisp MD, Cook LG (2011) Cenozoic extinctions account for the low diversity of extant
353 gymnosperms compared with angiosperms. New Phytol 192:997-1009
354 Environment Australia (2001) Threat abatement plan for dieback caused by the root-rot fungus
355 (*Phytophthora cinnamomi*). Environment Australia, Canberra
356 Forsteimer W, Schielzeth H, Mueller JC, Ellegern H, Kempenaers B (2012) Heterozygosity–
357 fitness correlations in zebra finches: microsatellite markers can be better than their reputation.
358 Mol Ecol 21:3237-3249
359 Gaston KJ (2003) The structure and dynamics of geographic ranges. Oxford University Press,
360 Oxford
361 Gelman A, Carlin JB, Stern HS, Rubin DB (2004) Bayesian data analysis. Chapman and
362 Hall/CRC.
363 Grace JB, Anderson TM, Olff H, Scheiner SM (2010) On the specification of structural equation
364 models for ecological systems. Ecol Monogr 80:67-87
365 Grueber CE, Wallis GP, Jamieson IG (2008) Heterozygosity–fitness correlations and their
366 relevance to studies on inbreeding depression in threatened species. Mol Ecol 17:3978-3984
367 He T, Lamont BB (2010) High microsatellite genetic diversity failed to predict greater
368 population resistance capacity in response to extreme climate change. Conserv Genet
369 11:1445-1451
370 He T, Rao GY, You RL, Ge S, Zhang DM (2000) Genetic diversity of widespread *Ophiopogon*
371 *intermedius* (Liliaceae s.l.): a comparison with endangered *O. xylorrhizus*. Biol Conserv 96:
372 253–257
373 He T, Krauss SL, Lamont BB, Miller BP, Enright NJ (2004) Long distance dispersal in a
374 metapopulation of *Banksia hookeriana* inferred by population allocation from AFLP data.
375 Mol Ecol 13:1099-1109

He T, Krauss SL, Lamont BB (2007) Polymorphic microsatellite DNA marker for *Banksia attenuata* (Proteaceae). Mol Ecol Notes 7:1329-1331

He T, Lamont BB, Krauss SL, Enright NJ, Miller BP (2009) Long-distance seed dispersal between populations of *Banksia attenuata* a fire-surviving species in Southwestern Australia. Ecography 32:571-580

He T, Lamont BB, Krauss SL, Enright NJ (2010) Landscape genetics of *Banksia hookeriana* in a metapopulation system. Ann Bot 106:457-466

He T, Lamont BB, Downs KS (2011) Banksias born to burn. New Phytol 191:184-196

Hopkins MJ (2011) The influence of morphological variation and geographic range size on species longevity in late Cambrian trilobites. Evolution 65:3252-3273

Hubbell SP (2001) The unified neutral theory of biodiversity and biogeography. Princeton University Press, Princeton, NJ

Krauss SL, He T, Barrett LG, Lamont BB, Miller BP, Enright NJ, Henley ME (2009) Contrasting impacts of pollen and seed dispersal on spatial genetic structure in a *Banksia hookeriana* population. Heredity 102:274-285

Lamont BB, He T, Enright NJ, Krauss SL, Miller BP (2003) Anthropogenic disturbance promotes hybridization between *Banksia* species by altering their biology. J Evol Biol 16: 551-557

Lamont BB, Enright NJ, Witkowski ETF, Groeneveld J (2007) Conservation biology of banksias: insights from natural history to simulation modeling. Aust J Bot 55:280-292

Lavie B, Nevo E (1986) Genetic selection of homozygote allozyme genotypes in marine gastropods exposed to cadmium pollution. Sci Total Environ 57:91-98

Lester SE, Ruttenberg BI, Gaines SD, Kinlan BP (2007) The relationship between dispersal ability and geographic range size. Ecol Lett 10:745-758

Liow LH, Stenseth NC (2007) The rise and fall of species: implications for macroevolutionary and

macroecological studies. Proc R Soc Lond B 274:2745-2752

Lynch M, Walsh B (1998) Genetics and analysis of quantitative traits. Sinauer Associates Inc., Sunderland, Massachusetts

Mast AR, Thiele K (2007) The transfer of *Dryandra* R.Br. to *Banksia* L.f. (Proteaceae). Aust Syst Bot 20:63–71

McCredie TA, Dixon KW, Sivasithamparam K (1985) Variability in the resistance of *Banksia* L.f. species to *Phytophthora cinnamomi* Rands. Aust J Bot 33:629-637

Merwin L, He T, Krauss SL (2010) Isolation and characterization of polymorphic microsatellite DNA markers for *Banksia candolleana* (Proteaceae). Conserv Genet Resources 2:345-347

Merwin L, He T, Lamont BB (2012) Phylogenetic structure in *Banksia* communities of southwestern Australia. J Biogeogr 39:397-407

Mitchell RJ. (1992) Testing evolutionary and ecological hypotheses using path analysis and structural equation modelling. Funct Ecol 6:123-129

Nevo E, Beiles A, Kaplan D, Golenberg EM, Olsvig-Wittaker L, Naveh Z (1986). Natural selection of allozyme polymorphisms: a microsite test revealing ecologic genetic differentiation in wild barely. Evolution 40:13-20

Olano-Marin J, Mueller JC, Kempenaers B (2011) Heterozygosity and survival in blue tits (*Cyanistes caeruleus*): contrasting effects of presumably functional and neutral loci. Mol Ecol 20:4028-4041

Paul JR, Morton C, Taylor CM, Tonsor SJ (2009) Evolutionary time for dispersal limits the extent but not the occupancy of species' potential ranges in the tropical plant genus *Psychotria* (Rubiaceae). Am Nat 173:188-199

Peakall R, Smouse P (2006) GenAlEx V6: Genetic Analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288-295

424 Petit RJ, El Mousadik A, Pons O (1998) Identifying populations for conservation on the basis of
425 genetic markers. Conserv Biol 12:844-855

426 Pigot AL, Owens IPF, Orme CDL (2012) Speciation and extinction drive the appearance of
427 directional range size evolution in phylogenies and the fossil record. PLoS Biol 10:e1001260

428 Pors B, Werner PA (1989) Individual flowering time in a goldenrod (*Solidago canadensis*): field
429 experiment shows genotype more important than environment. Am J Bot 76:1681-1688

430 Radwan JA, Biedrzycka A, Babik W (2010) Does reduced MHC diversity decrease viability of
431 vertebrate populations? Biol Conserv 143:537-544

432 Sexton JP, McIntyre PJ, Angert AL, Rice KJ (2009) Evolution and ecology of species range limits.
433 Ann Rev Ecol Evol Syst 40: 415-36

434 Shearer BL, Crane CE, Barrett S, Cochrane A (2007) *Phytophthora cinnamomi* invasion, a major
435 threatening process to conservation of flora diversity in the South-west botanical province of
436 Western Australia. Aust J Bot 55:225-238

437 Souza HA, Lovato MB (2010) Genetic diversity and structure of the critically endangered tree
438 *Dimorphandra wilsonii* and of the widespread in the Brazilian Cerrado *Dimorphandra mollis*:
439 Implications for conservation. Biochem Syst Ecol 38:49-56

440 Szulkin M, David P (2011) Negative heterozygosity–fitness correlations observed with
441 microsatellites located in functional areas of the genome. Mol Ecol 19:3949-3952

442 Szulkin M, Bierne N, David P (2010) Heterozygosity-fitness correlations: a time for reappraisal.
443 Evolution 64:1202-1217

444 Taylor A, Hopper S (1991) The Banksia Atlas. reprinted with amendments, Australian Flora and
445 Fauna series, No. 8. Australian Government Publishing Service, Canberra

446 Teacher AGF, Garner TWJ, Nichols RA (2009) Evidence for directional selection at a novel major
447 histocompatibility class I marker in wild common frogs (*Rana temporaria*) exposed to a viral

448 pathogen (Ranavirus). PLoS One 4:e4616
449 Tishkoff SA, Verrelli BC (2003) Pattern of human diversity: Implications for human
450 evolutionary history and disease. Ann Rev Genom Human Genet 4:293-340
451 Townsend AK, Clark AB, McGowan KJ, Miller AD, Buckles EL (2010) Condition, innate
452 immunity and disease mortality of inbred crows. Proc R Soc London B 277: 2875-2883
453 Tynan KM, Scott ES, Sedgley M (1998) Evaluation of *Banksia* species for response to
454 Phytophthora infection. Plant Pathol 47:446-455
455 Webb TJ, Gaston KJ (2000) Geographic range size and evolutionary age in birds. Proc R Soc
456 London B 267:1843-1850
457 Willis JC (1922) Age and area. Cambridge University Press, Cambridge
458 Wills RT, Keighery GJ (1994) Ecological impact of plant disease on plant communities. J Roy
459 Soc WA 77:127-132
460 Wooller SJ, Wooller RD (2001) Seed set in two sympatric banksias, *Banksia attenuata* and *B.*
461 *baxteri*. Aust J Bot 49: 597-602
462
463
464
465
466
467
468
469
470
471
472

Figure Caption:

Fig. 1. SEM analysis using Bayesian MCMC on two pre-determined models. The models represent the hypothesised path effects involving the evolutionary history of species, geographic ranges, ecological ranges, genetic diversity and resistance to *Phytophthora cinnamomi* infection. Models A and B were tested separately. In Model B, the parameter that has the most significant regression weight on genotypic diversity (as revealed in Model A) was used. H_1 – H_4: working hypothesis as listed in Methods. Path coefficients are presented as standard direct regression weight. a: Standardised regression weight derived in Model A; b: Standard regression weight derived in Model B.

Table 1 Coefficients of determination and associated probability values derived in conventional bivariate correlation analyses.

	Evolutionary time	Geographic range	Ecological range (rainfall range)	Life history (flowering duration)	Genotypic diversity	Heterozy-gosity
Geographic range (Square rooted)	R^2 = 0.321 P = 0.048					
Ecological range (rainfall range)	R^2 = 0.164 P = 0.184	R^2 = 0.628 P = 0.002				
Life history diversity (flowering duration)	R^2 = 0.025 P = 0.738	R^2 = 0.501 P = 0.006	R^2 = 0.437 P = 0.014			
Genotypic diversity	R^2 = 0.242 P = 0.104	R^2 =0.376 P = 0.026	R^2 = 0.297 P = 0.050	R^2 =0.444 P = 0.012		
Heterozygosity	R^2 = 0.473 P = 0.014	R^2 =0.183 P = 0.144	R^2 = 0.226 P = 0.108	R^2 = 0.193 P = 0.133	R^2 = 0.785 P = 0.001	
Resistance to *P. cinnamomi*	R^2 = 0.032 P = 0.349	R^2 =0.630 P = 0.036	R^2 = 0.406 P = 0.064	R^2 =0.4760 P = 0.038	R^2 = 0.644 P =0.008	R^2=0.309 P =0.082

508

509

510

511

512

513

514

Structural Equation Modelling Analysis of Evolutionary and Ecological Patterns in Australian *Banksia*

Tianhua He

Table S1 List of *Banksia* species, and their genetic diversity, species age, geographic and ecological ranges.

Species	N_a	H_E	Postfire regeneration	Age (Million years)	Resistance to *P. cinnamomi*	Source for genetic data
B. attenuata	9.9	0.76	Resprouter	18.4	28	This study
B. brownii	5.0	0.60	Seeder	5.6	78	McArthur and Coates 2009
B. candolleana	7.6	0.74	Resprouter	13.1	34	This study
B. hookeriana	6.6	0.64	Seeder	1.5	75	He et al. 2010
B. ilicifolia	9.6	0.77	Seeder	0.9	20	This study
B. ionthocarpa	2.6	0.41	Resprouter	N/A	N/A	Millar et al. 2010
B. menziesii	6.1	0.51	Resprouter	1.6	39	This study
B. nivea	7.3	0.57	Seeder	2.3	N/A	Millar and Byrne 2008
B. oblongifolia	4.0	0.43	Resprouter	0.9	0	Urhser et al. 2005
B. paludosa	6.1	0.58	Resprouter	1.2	2	Urhser et al. 2005
B. robur	3.4	0.31	Resprouter	0.9	0	Urhser et al. 2005
B. sphaerocarpa	9.6	0.71	Resprouter	6.0	35	Llorens et al. 2012
B. spinulosa	6.5	0.76	Resprouter	12.8	0	O'Brien et al. 2010

518 （Continued)

519

Species	Soil type diversity	Vegetation type diversity	Landform type diversity	Pollinator diversity	Growing duration (months)	Flowering duration (months)	Rainfall range (mm)	Geographic range ($\times 10^3$ km^2)
B. attenuata	3	3	4	11	12	10	800	140.0
B. brownii	7	4	4	21	3	3	200	5.0
B. candolleana	2	2	2	6	5	4	100	9.0
B. hookeriana	1	2	2	16	8	5	300	1.2
B. ilicifolia	1	4	6	16	12	11	500	24.0
B. ionthocarpa	4	2	3	10	N/A	2	200	0.5
B. menziesii	1	3	2	16	12	6	550	31.0
B. nivea	4	3	3	11	N/A	5	600	90.0
B. oblongifolia	5	3	6	16	12	7	400	30.0
B. paludosa	3	4	6	26	10	3	200	5.5
B. robur	4	3	3	16	12	7	200	10.0
B. sphaerocarpa	7	4	4	16	12	11	500	85.0
B. spinulosa	5	4	5	16	12	8	800	94.0

520

521 **Reference**

522 He T, Lamont BB, Krauss SL, Enright NJ (2010) Landscape genetics of *Banksia hookeriana* in a metapopulation system. Ann Bot 106: 457-466

523 Llorens TM, Byrne M, Yates CJ, Nistelberger HM, Coates DI (2012) Evaluating the influence of different aspects of habitat fragmentation on mating patterns
524 and pollen dispersal in the bird-pollinated *Banksia sphaerocarpa* var. *caesia*. Mol Ecol 21: 314–328

525 McArthur SL, Coates DJ (2009) Characterisation of microsatellite markers for the rare and critically endangered *Banksia brownii* (Proteaceae) Conserv Genet
526 Resources 2: 177-179

527 Millar MA, Byrne M (2008) Characterization of polymorphic microsatellite DNA markers in *Banksia nivea*, formerly *Dryandra nivea.* Mol Ecol Resources 8:
528 1393–1394

529 Millar MA, Byrne M, Coates DJ (2010) The maintenance of disparate levels of clonality, genetic diversity and genetic differentiation in disjunct subspecies of
530 the rare *Banksia ionthocarpa*. Mol Ecol 19: 4127–4227

531 O'Brien EK, Aguilar LA, Ayre DJ, Whelan RJ (2010) Genetic tests of the isolation of rare coastal dwarf populations of *Banksia spinulosa*. Aust J Bot 58:
532 637–645

533 Usher AV, Ayre DJ, Whelan RJ (2005) Microsatellites for eastern Australian *Banksia* species. Mol Ecol Notes 5: 821–823

534

535

536

537

538

539

540

541 Table S2 Details of microsatellite primers for *Banksia ilicifolia* and *B. menziesii*. T_m: melting temperature; N_a: genotypic diversity; H_E expected
542 heterozygosity.

543

544 *Banksia ilicifolia*

Locus		Primer sequence	Repeats	T_m °C	N_a	H_E	Allele size (base pairs)
BI – A3	Forward:	5'AGGCCAACAGAGATTATGC'3	$(CA)_{13}$	49	12	0.85	197—219
	Reverse:	5'ATACGAAAGCACGATACATACA'3					
BI – A5	Forward:	5'AGCATTTAGACCCCAAATATG'3	$(AC)_{22}$	52	4	0.59	168-186
	Reverse:	5'CGCCCATACTTTGTAAACTTAG'3					
BI – A110	Forward:	5'ATCCCGATTACTTCAAAAACC'3	$(CA)_{13}$	53	13	0.84	149-185
	Reverse:	5'GTGAGCAGGCTGCCATAT'3					
BI – A111	Forward:	5'TGGATGCTTGATTTATGTCC'3	$(CA)_{43}$	49	19	0.89	263-309
	Reverse:	5'TTCTCCCTGAAACTTGTGAG'3					
BI – B6	Forward:	5'TTTCCTCTTACCCATCAGATG'3	$(TC)_{14}$	52	6	0.67	248-258
	Reverse:	5'GACGGGGAGTAGTAAATAATGC'3					
BI – B104	Forward:	5'CACACTTTCACTGCTCACAC'3	$(AG)_{15}$	51	12	0.83	219-243
	Reverse:	5'CGTAACCCGAAAATGTGTAC'3					
BI – B106	Forward:	5'TAGGGCTTTGGGCATCTTAG'3	$(TC)_{21}$	52	9	0.74	125-129

	Reverse:	5'GGATGGGTGTTGGAAGAAGT'3					
BI – B108	Forward:	5'CTGGTGGGTTTGAGGATCT'3	$(TC)_{36}$	52	12	0.90	186-212
	Reverse:	5'CTAATGGAAACAGCACTGACTG'3					
BI – C103	Forward:	5'CGTTTGTCAAGTCTGGTGATC'3	$(CAA)_9$	51	5	0.77	259-281
	Reverse:	5'TGCTCTTTTGGATCTATGTGG'3					
BI – D3	Forward:	5'TCAGCCTATCACTGCTACATC'3	$(GAT)_{15}$	51	8	0.83	112-155
	Reverse:	5'TTCTGCTCACCACATAAACTC'3					

545

546 *Bankisa menziesii*

Code		Sequence	Repeats	T_m °C	N_a	H_E	Allele size (base pairs)
BM-A1	Forward:	5'GCTGGAGATGGGATTTCTAG'3	$(AC)_{18}$	52	11	0.84	180-194
	Reverse:	5'TGCTGAACCATGTTCCTTAT'3					
BM-B6	Forward:	5'CTCGCCCTTCTATTGGTG'3	$(CT)_{19}$	54	12	0.82	260-284
	Reverse:	5'GGTGGCTGAGTGAGATGG'3					
BM-C2	Forward:	5'AGCGTCGTGTTTTCTTCTG'3	$(CAA)_8$	52	3	0.07	242-248
	Reverse:	5'GAGAGCCGATGAATGTCTATC'3					
BM-D1	Forward:	5'CGGAATCCTGTAATCACCTT'3	$(GAT)_7$	52	3	0.30	157-163
	Reverse:	5'TCCCAGTGGAAAGAACAAC'3					

BM-D4	Forward:	5'TCCTGCTCATCATCACAGTC'3	$(TCG)_6$	54	4	0.27	283-292
	Reverse:	5'CAACCAACCACCAACAGTC'3					
BM-A101	Forward:	5'TCCTGTTCTTACCAAATTCATG'3	$(AC)_{16}$	52	4	0.37	216-228
	Reverse:	5'GATACCATGCTCAAATTCAATC'3					
BM-B102	Forward:	5'CGAACCCTCGTCAATGAAC'3	$(TC)_{25}$	54	11	0.86	168-192
	Reverse:	5'TGAGCAGAACCAGAGCAGA'3					
BM-C104	Forward:	5'AAGACCGTTTCTGTGATTGTG'3	$(TGT)_{11}$	54	4	0.54	195-207
	Reverse:	5'GAGGATTTGGTGGAAGAGTTC'3					
BM-D103	Forward:	5'AACTGCATCAGAACAACTCATG'3	$(CAT)_9$	54	5	0.57	249-267
	Reverse:	5'TGTGGAGGTACTAATGCTGTTG'3					
BM-D105	Forward:	5'TTGTTTACCCTTCCGACTTTA'3	$(ATC)_7$	52	3	0.10	232-243
	Reverse:	5'TGATGGTTTGATTAAGAGGATG'3					

547