Research Article

Population genetic analysis reveals barriers and corridors for gene flow within and among riparian populations of a rare plant

Tanya H. Hevroy1,4*, Michael L. Moody2 and Siegfried L. Krauss1,3

1School of Plant Biology, Faculty of Natural and Agricultural Sciences, The University of Western Australia, Crawley, Western Australia 6009, Australia
2Department of Biological Sciences, University of Texas at El Paso, 500 W. University Ave., El Paso, TX 79968, USA
3Kings Park and Botanic Garden, Botanic Gardens and Parks Authority, Fraser Ave., West Perth, Western Australia 6005, Australia
4Present address: Norwegian Radiation Protection Authority, Grinnaerings park 13, 1361 Oslo, Norway

Received: 6 April 2017    Editorial decision: 3 November 2017    Accepted: 19 November 2017    Published: 21 November 2017

Associate Editor: F. Xavier Picó

Citation: Hevroy TH, Moody ML, Krauss SL. 2018. Population genetic analysis reveals barriers and corridors for gene flow within and among riparian populations of a rare plant. AoB PLANTS 10: plx065; doi: 10.1093/aobpla/plx065

Abstract. Landscape features and life-history traits affect gene flow, migration and drift to impact on spatial genetic structure of species. Understanding this is important for managing genetic diversity of threatened species. This study assessed the spatial genetic structure of the rare riparian Grevillea sp. Cooljarloo (Proteaceae), which is restricted to a 20 km² region impacted by mining in the northern sandplains of the Southwest Australian Floristic Region, an international biodiversity hotspot. Within creek lines and floodplains, the distribution is largely continuous. Models of dispersal within riparian systems were assessed by spatial genetic analyses including population level partitioning of genetic variation and individual Bayesian clustering. High levels of genetic variation and weak isolation by distance within creek line and floodplain populations suggest large effective population sizes and strong connectivity, with little evidence for unidirectional gene flow as might be expected from hydrochory. Regional clustering of creek line populations and strong divergence among creek line populations suggest substantially lower levels of gene flow among creek lines than within creek lines. There was however a surprising amount of genetic admixture in floodplain populations, which could be explained by irregular flooding and/or movements by highly mobile nectar-feeding bird pollinators. Our results highlight that for conservation of rare riparian species, avoiding an impact to hydrodynamic processes, such as water tables and flooding dynamics, may be just as critical as avoiding direct impacts on the number of plants.

Keywords: Dispersal; landscape genetics; metapopulation; riparian systems; Southwest Australian Floristic Region; spatial genetic structure.

Introduction

Gene flow is a key evolutionary force affecting spatial genetic structure within species (Ellstrand 2014). Gene flow in plants occurs through diploid embryos in seeds and haploid gametes in pollen (Holderegger et al. 2010). These processes are influenced by the composition of...
landscapes they inhabit, as well as life-history traits such as pollination, mating systems and vectors for dispersal (Storfer et al. 2007; Holderegger et al. 2010).

Landscape features impacting gene flow may be striking, for example mountain ranges (Hu et al. 2008; Trénel et al. 2008), intervening habitats between riparian environments (Kitamoto et al. 2005; Mitsu et al. 2010) or anthropogenic induced fragmentation (Cegelski et al. 2003; Epps et al. 2007). However, in many landscapes, barriers to gene flow may be more cryptic, for instance historical vegetation changes (Jacquemyn et al. 2006), plant physiology (Nistelberger et al. 2015), and/or a combination of both biotic (e.g. pollinator response to habitat features) and abiotic (e.g. wind as a dispersal vector) factors (Holderegger et al. 2010). For riparian habitats, landscape features such as rivers, creek lines and floodplains can provide critical corridors that can both facilitate and limit dispersal (Tero et al. 2003; Storfer et al. 2007; Hu et al. 2008). Establishing an understanding of the landscape features that influence spatial genetic structure and gene flow helps to predict and manage potential impacts on rare species from landscape disturbances.

Through spatial analyses of individual genotypes and genetic clusters, landscape features affecting dispersal can be better understood. Clusters can be viewed as genetically divergent groups of individuals where gene flow is inferred to be impeded (Guillot et al. 2005; Chen et al. 2007; Corander et al. 2008). Clustering models applying Bayesian assignment methods have seen a shift from population-based to individual-based analyses (Storfer et al. 2010). These methods are especially useful for rare species that are localized with individuals not obviously clustered in discrete populations but rather arranged more continuously, such as in many riparian habitats, and when barriers to gene flow may be more cryptic (Manel et al. 2003).

Riparian distributions are typically characterized by largely linear patterns in which gene flow is likely to occur more readily within, rather than among, habitats (Wei et al. 2013). For example, seeds dispersed by water (hydrochory) will promote gene flow within but not among riparian corridors. Hydrochory as a force of dispersal can be responsible for influencing spatial genetic structuring of populations along riparian environments at multiple scales (Kudoh et al. 2006; Liu et al. 2006; Nilsson et al. 2010). However, bidirectional pollinator movement, zoochory or wind dispersal can counteract unidirectional downstream dispersal by hydrochory (Tero et al. 2005; Markwith and Scanlan 2007). While flooding of riparian environments can facilitate seed dispersal (Kitamoto et al. 2005; Nilsson et al. 2010), other studies indicate that the impact and consequences for spatial genetic structure are not necessarily predictable (Prentis et al. 2004).

For rare taxa with a narrow endemic range any new limitation to gene flow could be detrimental to the genetic stability and diversity maintained by the species. Therefore, it is important to understand the genetic structure of such species at a landscape scale, where natural features can provide barriers to gene flow even at a small spatial scale. Here, we implement a genetic approach to describe the spatial structure of genetic variation within and among creek line and floodplain populations of a rare plant species, Grevillea sp. Cooljarloo (https://florabase.dpaw.wa.gov.au/search/current/37180). This taxon (subsequently referred to as G. ‘Cooljarloo’) is listed as Priority 1 (rare and threatened) by the Western Australian Department of Parks and Wildlife (Smith 2013). Although locally abundant, the distribution of G. ‘Cooljarloo’ is restricted to creek lines and floodplains (Fig. 1), in a 20 km² area south of Cervantes, in the Southwest Australian Floristic Region (SWAFR), Western Australia (Hopper and Gioia 2004). In particular, we address the influence of landscape features on the genetic patterns found, and how the expansion of a mineral sand-mining operation may impact on conservation plans as a result of further habitat clearing and impact on drainage lines.

Methods

Species biology

Grevillea ‘Cooljarloo’ is a spreading lignotuborous shrub to 1.5 m on winter-wet sandy soils. Plants typically produce numerous bright red flowers (Fig. 1), which are self-compatible but, like most Grevillea (England et al. 2003; Roberts et al. 2007), are likely to be predominantly outcrossed. The main pollinators are nectar-feeding birds, specifically honeyeaters (Meliphagidae), but marsupials such as the honey possum (Tarsipes rostratus) and bees (including the introduced honeybee Apis mellifera) are likely to play a role in pollination (Phillips et al. 2010). As for most Grevillea, G. ‘Cooljarloo’ has no obvious adaptations for seed dispersal beyond gravity (Makinson 2000; England et al. 2003).

Sampling strategy and study sites

A 30-m digital elevation model was constructed from 974 spot heights provided by Tronox Inc. (CAD resources) using ArcGIS 10.1 (ESRI, Redlands, CA, USA) and visualized as contour lines to illustrate elevation changes within the study region (Fig. 2). Creek lines were drawn onto the map based on satellite images and previous landscape mapping provided by Tronox Inc. Leaf samples
were collected from three regions within the study area based on GPS points provided by the mine site environmental surveyors, who had previously recorded all populations in the area. Three large creeks occur within the region, here referred to as North (NPC), South (SPC) and West (WPC). Populations within these creeks and three adjoining floodplain populations (CFP) were sampled for this study (Fig. 2).

The three sampled creek lines are found within a Banksia-dominated woodland. The west creek line opens up into a floodplain before continuing west into Nambung National Park (Fig. 2). The north creek line is impacted at both ends by agriculture or mining. The southern creek is impacted by agriculture and mining in the east, and enters Cooljarloo Swamp in the south-west. Two sites where G. ‘Cooljarloo’ individuals have been previously recorded could not be sampled (Fig. 2; indicated by stars). The first site was recently burnt and contained no live plants. The second site in the north was inaccessible due to current expansions of the Tronox Mine.

Within the creek lines, samples were collected based on a constant sampling fraction (Lowe et al. 2009) by defining a 20 m × 20 m perimeter on both sides of the creek every 0.5–1 km. All individuals within the perimeter were sampled, and these were called populations (Table 1). For floodplain populations, every individual from a randomly selected 20 m x 20 m perimeter within each of the three floodplains adjacent to north, south
Genotyping

Genomic DNA was extracted using the modified CTAB extraction method (Doyle 1990). Quantity and quality of DNA were assessed by visualization on a 1.5 % agarose gel and a Nanodrop (Thermo Fisher, Wilmington, DE, USA). Genetic variation was assessed at nine microsatellite loci, which were amplified and scored as reported in Hevroy et al. (2013). Standard genetic diversity parameters, total number of alleles (\(N_a\)), observed and expected heterozygosity (\(H_o\) and \(H_e\)) and inbreeding coefficient (\(F_{IS}\)), were estimated for each population using GenALEX 6.5 (Peakall and Smouse 2012). Allelic richness (\(A_r\)) and private allelic richness (\(p_r\)) were estimated for each population at every locus with HP-Rare (Kalinowski 2005) which uses rarefaction to account for differences in sample size. Significance of differences for genetic diversity parameters between creek line and floodplain populations was assessed by t-tests. Deviations from Hardy–Weinberg equilibrium and linkage disequilibrium were tested using an adjusted sequential Bonferroni correction (Rice 1989) and 10,000 randomizations at the nominal significance level (5 %) in GenALEX 6.5. MicroChecker 2.2.3 (Van Oosterhout et al. 2004) was used to detect presence of null alleles and potential scoring errors.

Population-based structure analyses

Pairwise population genetic differentiation was estimated by \(F_{ST}\) with ARLEQUIN V.3.5 (Excoffier and Lischer 2010), and Jost’s \(D_{EST}\) with SMOGD (Crawford 2010). Genetic distances among populations and spatial collections were also estimated using Nei’s genetic distance (Nei 1972). Ordinations were conducted with all three distance matrices using principal coordinates analysis (PCoA). The total genetic diversity was partitioned into within and among population components, as well as upstream and downstream (based on elevation data), by hierarchical analysis of molecular variance (AMOVA) using ARLEQUIN V.3.5, with 10,000 permutations and significance levels set with nominal values of 0.05 following Bonferroni correction. Isolation by distance among populations and among individuals within all three creek lines was evaluated using Mantel tests.
between log transformed geographic distance and genetic distance [standardized $F_{st}/(1 - F_{st})$] (Rousset 2000) in GenAlEx 6.5. Significance of the correlation was tested using $P$-values from 10 000 permutations.

**Individual-based genetic structure analyses**

Spatial genetic structure was further analysed using STRUCTURE 2.3.3 (Pritchard et al. 2000), which uses a Bayesian algorithm to estimate probability of membership of an individual ($q$) to an assumed genetic cluster, using a Markov Chain Monte Carlo algorithm. Given the restricted range of populations, an admixture ancestry model with correlated allele frequencies was used. The number of potential genetic clusters ($K$) in the data set was tested from 1 to 10, replicated 10 times for 1 million generations and a burn-in of 100 000. Optimum $K$ was identified using the $\Delta K$ function of Evanno et al. (2005), which is based on second rate change of likelihood function with respect to $K$, and was calculated using STRUCTURE HARVESTER (Earl and von Holdt 2012).

Given the geographic clustering of populations, we also employed a Bayesian assignment method that explicitly incorporated geographic prior distributions (Storfer et al. 2007). TESS 2.3 (Chen et al. 2007) implements a Bayesian clustering algorithm that uses a hidden Markov random fields model and allows for a detailed admixture analysis. This method is based on the spatial autocorrelation assumptions that allele frequencies at a geographic site are more likely to be similar to allele frequencies at neighbouring sites than at distant sites (François and Durand 2010). This method provides better estimates of ancestry coefficients when levels of ancestral population divergence are low (Durand 2007), which is a potential issue with our data.

The degree of spatial dependence incorporated in the analyses is determined by a parameter ($\psi$), which represents the strength of spatial autocorrelation, ranging from 0 to 1. In this analysis, the spatial parameter was set $\psi = 0.6$, as recommended by the authors (Chen et al. 2007). Each run was performed under the admixture model using 50 000 cycles with a burn-in of 10 000 for $K = 2–8$ with 100 iterations for each $K$. Based on the 10 highest deviance information criterion (DIC) values for each $K$, the most likely number of clusters ($K$) was determined from where the DIC values plateau (Chen et al. 2007). The proportion of membership of each individual to given clusters at the optimal $K$ ($q$) was obtained to illustrate these values. Similarity among different runs was assessed using the similarity coefficient ($H$) in CLUMPP 1.1.2 (Jakobsson and Rosenberg 2007) for both replicates of STRUCTURE and TESS runs at optimal $K$. 

### Table 1. Sampling information and genetic diversity statistics averaged over nine microsatellite loci for 13 natural populations of Grevillea sp. Cooljarloo, including landscape and position of population along creeks (U = upstream and D = downstream), average elevation (m a.s.l.) and coordinates in latitude/longitude.

| Site          | Collection ID | Landscape | Elevation (m) | N  | $N_a$ | $A_a$ | $pA_a$ | $H_{st}$ | $H_e$ | $F_{IS}$ | Coordinates (UTM) |
|---------------|---------------|-----------|---------------|----|-------|-------|--------|---------|-------|---------|--------------------|
| North         | NPC1 Creek line (U) | 80 | 22 | 5.3 | 5.1 | 0.12 | 0.608 | 0.691* | 0.125 | 30°38'13.2"S 115°17'20.4"E |
|               | NPC2 Creek line (U) | 70 | 17 | 5.2 | 5.0 | 0.27 | 0.606 | 0.684* | 0.114 | 30°38'49.2"S 115°19'30.1"E |
|               | NPC3 Creek line (D) | 60 | 15 | 5.1 | 4.8 | 0.09 | 0.633 | 0.628 | 0.221 | 30°38'19.8"S 115°18'55.0"E |
| South         | SPC1 Creek line (D) | 50 | 23 | 5.7 | 5.0 | 0.14 | 0.629 | 0.677 | 0.116 | 30°40'22.4"S 115°18'28.8"E |
|               | SPC2 Creek line (D) | 55 | 15 | 5.1 | 4.8 | 0.09 | 0.633 | 0.628 | 0.221 | 30°38'19.8"S 115°18'55.0"E |
|               | SPC3 Creek line (D) | 60 | 28 | 5.2 | 3.9 | 0.09 | 0.586 | 0.644 | 0.005 | 30°38'56.1"S 115°18'59.8"E |
|               | SPC4 Creek line (D) | 60 | 15 | 4.0 | 4.3 | 0.09 | 0.633 | 0.628 | 0.079 | 30°38'29.5"S 115°19'10.5"E |
| West          | WPC1 Creek line (D) | 40 | 22 | 6.6 | 6.0 | 0.28 | 0.777 | 0.739 | -0.053 | 30°37'22.8"S 115°15'32.4"E |
|               | WPC2 Creek line (D) | 50 | 26 | 6.8 | 5.9 | 0.09 | 0.566 | 0.723* | 0.221 | 30°37'34.6"S 115°15'0.0"E |
|               | WPC3 Creek line (D) | 60 | 29 | 6.7 | 5.6 | 0.05 | 0.696 | 0.729 | 0.043 | 30°37'9.1"S 115°14'42.7"E |
| Mean          |                | 57.5 | 22 | 5.63 | 5.05 | 0.01 | 0.64 | 0.69 | 0.078 |
|               | CFP(34) Floodplain | 45 | 27 | 6.7 | 5.6 | 0.2 | 0.622 | 0.695* | 0.108 | 30°34'37.9"S 115°16'41.8"E |
|               | CFP(15) Floodplain | 50 | 30 | 6.1 | 5.2 | 0.25 | 0.672 | 0.696 | 0.037 | 30°30'51.3"S 115°18'33.8"E |
|               | CFP(10) Floodplain | 55 | 28 | 6.7 | 5.9 | 0.15 | 0.667 | 0.760 | 0.118 | 30°37'33.4"S 115°15'49.7"E |
| Mean          |                | 50 | 28.33 | 6.50 | 5.57 | 0.20 | 0.65 | 0.72 | 0.090 |
Results

Landscape elevation
Elevation within the study site ranged from 30 to 205 m. Areas where G. ‘Cooljarloo’ populations were mapped and sampled had an average elevation of 51 ± 0.5 m (37–79 m), which was significantly lower than the average elevation of 91 ± 2.3 (52–91 m) (P < 0.001) in habitats where G. ‘Cooljarloo’ was not found. Elevational differences affected the habitat types (creek lines, floodplains, sandplains), which reflect the presence or absence of G. ‘Cooljarloo’.

Genetic diversity
The nine polymorphic loci amplified a total of 102 alleles across 355 individuals. Only locus GM13 exhibited potential null alleles. Genetic analyses without this marker did not generate significant differences in results compared to the full data set, so it was included in analyses to increase robustness. Mean estimates (and range) of genetic parameters across nine loci were: number of alleles 5.2 (range = 3.6–6.2), allelic richness 5.1 (3.9–5.9), observed heterozygosity 0.64 (0.58–0.77) and expected heterozygosity 0.69 (0.63–0.76). F<sub>IS</sub> deviated significantly from zero within five populations (Table 1), with fewer heterozygotes than expected. Private alleles were largely absent in the SPC creek line populations, and relatively higher in the floodplain and NPC creek line populations (P < 0.001). Mean private allelic richness was significantly greater (t = 2.3; P < 0.05) for floodplain populations (ρ<sub>p</sub> = 0.2) than creek line populations (ρ<sub>p</sub> = 0.1) (Table 1). All other genetic diversity parameters were not significantly different.

Population-based genetic structure
Mean (and range) pairwise F<sub>ST</sub> across all populations was 0.087 (0.028–0.178), mean pairwise D<sub>EST</sub> was 0.215 (0.072–0.422) (Table 2), and both parameters showed similar patterns of divergence. There was significant genetic differentiation for all pairwise population comparisons except NPC1/NPC2, and SPC1/SPC4 (Table 2). Genetic differentiation was higher for inter-creek population pairs than for intra-creek population pairs (mean F<sub>ST</sub> = 0.112 (inter), 0.064 (intra), P = 0.024; mean D<sub>EST</sub> = 0.273 (inter), 0.137 (intra), P = 0.014). The highest pairwise F<sub>ST</sub> value was 0.178 (NPC2/SPC3) and the lowest was 0.027 (WPC2/WPC3), which reflected geographic distance (12.2 and 0.9 km, respectively). The highest D<sub>EST</sub> value was 0.422 (NPC1/SPC2) and the lowest was 0.072 (NPC1/NPC2) again reflecting geographic distance (11.7 and 0.3 km, respectively). Pairwise F<sub>ST</sub> for intra-creek population pairs for each individual creek was 0.065 (NPC), 0.060 (SPC) and 0.054 (WPC), and D<sub>EST</sub> was 0.145 (NPC), 0.097 (SPC) and 0.170 (WPC). In addition, pairwise F<sub>ST</sub> for inter-creek comparisons when we compare one creek to another was 0.105 (NPC vs. SPC), 0.088 (NPC vs. WPC) and 0.138 (SPC vs. NPC), and D<sub>EST</sub> comparisons are 0.233 (NPC vs. SPC), 0.245 (NPC vs. WPC) and 0.339 (SPC vs. NPC).

Analysis of molecular variance partitioned a small, but significant, amount of the total genetic variation among

| CFP10 | CFP15 | CFP34 | NPC1 | NPC2 | NPC3 | SPC1 | SPC2 | SPC3 | SPC4 | WPC1 | WPC2 | WPC3 |
|-------|-------|-------|------|------|------|------|------|------|------|------|------|------|
| 0.179 | 0.094 | 0.191 | 0.199 | 0.185 | 0.195 | 0.168 | 0.230 | 0.214 | 0.182 | 0.078 | 0.164 |
| 0.054 | 0.152 | 0.292 | 0.325 | 0.235 | 0.097 | 0.235 | 0.175 | 0.211 | 0.156 | 0.229 | 0.201 |
| 0.030 | 0.041 | 0.126 | 0.156 | 0.201 | 0.145 | 0.193 | 0.235 | 0.186 | 0.250 | 0.101 | 0.159 |
| 0.057 | 0.067 | 0.061 | 0.072 | 0.177 | 0.268 | 0.422 | 0.357 | 0.288 | 0.392 | 0.165 | 0.271 |
| 0.073 | 0.119 | 0.080 | 0.032 | 0.186 | 0.270 | 0.384 | 0.395 | 0.366 | 0.359 | 0.150 | 0.218 |
| 0.086 | 0.098 | 0.087 | 0.070 | 0.093 | 0.272 | 0.398 | 0.320 | 0.331 | 0.228 | 0.184 | 0.239 |
| 0.061 | 0.041 | 0.051 | 0.071 | 0.123 | 0.113 | 0.095 | 0.092 | 0.091 | 0.179 | 0.199 | 0.187 |
| 0.057 | 0.110 | 0.084 | 0.146 | 0.166 | 0.175 | 0.046 | 0.074 | 0.144 | 0.197 | 0.200 | 0.179 |
| 0.089 | 0.093 | 0.109 | 0.116 | 0.178 | 0.148 | 0.047 | 0.060 | 0.084 | 0.227 | 0.297 | 0.263 |
| 0.087 | 0.095 | 0.094 | 0.100 | 0.176 | 0.150 | 0.045 | 0.103 | 0.058 | 0.308 | 0.299 | 0.278 |
| 0.049 | 0.063 | 0.065 | 0.090 | 0.112 | 0.091 | 0.068 | 0.073 | 0.096 | 0.120 | 0.227 | 0.204 |
| 0.037 | 0.082 | 0.055 | 0.071 | 0.074 | 0.097 | 0.076 | 0.076 | 0.138 | 0.146 | 0.071 | 0.078 |
| 0.057 | 0.074 | 0.060 | 0.077 | 0.082 | 0.095 | 0.081 | 0.101 | 0.136 | 0.147 | 0.064 | 0.027 |
the three geographic clusters (4.5 %), among populations within geographic clusters (4.6 %) and among individuals within populations (4.1 %), with the remainder (86.8 %) partitioned within individuals (Table 3). When applied to the creek lines alone, 89.4 % of total genetic variation was partitioned within populations, 5.5 % partitioned among the three creek lines and 5.1 % among populations within creek lines (data not shown). When AMOVA was applied to downstream and upstream populations, no significant amount of variation was partitioned among them (2.0 %, P = 0.02; Table 3).

Near identical ordinations were generated from each of the three distance matrices. The first two principal axes in the PCoA of Nei’s D distance (Fig. 3) explained 44.6 and 18.43 % of the total genetic variation, respectively. The PCoA showed the NPC, SPC and WPC populations clustered separately, and the floodplain populations clustered together in between creeks. The PCoAs of \( F_{ST} \) and \( D_{est} \) genetic distance explained 34.8 and 18.8 %, and 34.2 and 19.2 %, respectively. The association of pairwise population differentiation (\( F_{ST} \)) against logarithmic geographic distance was significant when calculated across all sampled populations \( (r^2 = 0.230, P = 0.0001; \text{ Fig. 4}) \), indicating overall isolation by distance. However, the association between genetic distance and geographic distance among individuals within each creek line (NPC, SPC or WPC) was not significant \( (r^2 = 0.008; r^2 = 0.001; r^2 = 0.033, P > 0.05 \) in each case).

### Individual-based genetic structure

There was no spatial genetic structure identified within creek lines from the individual-based Bayesian analyses but there was evidence for significant genetic structure between the creek lines. The plot of \( \Delta K \) against \( K \) [see Supporting Information—Fig. S1] values from STRUCTURE indicated two genetic clusters were optimal (Fig. 5A and B), comprising northern and southern populations, respectively.

The distribution of genetic clusters across the creek lines and floodplains as assigned by TESS is illustrated in Fig. 2, and the individual membership proportions are illustrated in Fig. 5C. The DIC values indicated the optimal number of clusters as three [see Supporting Information—Fig. S1]. When using admixture analyses, the DIC values for three clusters dropped considerably (no-admixture lowest run = 26979, admixture lowest run = 20115) suggesting that the admixture model had a better fit (Row et al. 2010). Chen et al. (2007) suggested evaluating posterior probabilities when using spatial components in assignment testing, and to base success on ability of the programs to assign number of individuals to a cluster with membership above 0.8. Under this criterion, TESS assigned 80 % of individuals to a cluster at \( K = 3 \), outperforming STRUCTURE, which assigned 57 % at \( K = 2 \) and 47 % at \( K = 3 \). Individuals from SPC populations had the greatest mean proportion of membership \( (q_i = 0.98) \), followed by NPC \( (q_i = 0.97) \) and WPC \( (q_i = 0.78) \). Two CFP populations, CFP(10) and CFP(34), showed high levels of admixture. CFP (10) showed highest identity with SPC \( (q_i = 0.42) \) and WPC populations \( (q_i = 0.48) \), whereas CFP (34) had highest identity with NPC \( (q_i = 0.61) \) and SPC \( (q_i =0.26) \) populations. CFP (15) was clustered with SCP populations \( (q_i = 0.99) \).

### Discussion

Grevillea ‘Cooljarloo’ is restricted to creek lines and floodplains with an average elevation of 51 m a.s.l., and is absent from intervening habitats that average 91 m in elevation. These landscape features strongly influence the distribution and spatial structure of genetic variation for this species. Spatial genetic analyses supported a conclusion of three genetic clusters that largely corresponded to the disjunct regional distribution of G. ‘Cooljarloo’ in northern, southern and western creek lines and floodplains. A significant isolation by distance relationship was associated with comparisons among

| Source of variation | d.f. | Sum of squares | Percentage of variation | P       |
|---------------------|-----|----------------|------------------------|---------|
| Among clusters      | 2   | 84.8           | 4.1                    | <0.001  |
| Among populations within clusters | 13  | 124.8          | 4.5                    | <0.001  |
| Among individuals within populations | 349 | 1072.4         | 4.6                    | <0.001  |
| Within individuals  | 365 | 1014.5         | 86.8                   | <0.001  |
| Upstream vs. downstream | 1  | 33.4           | 2.0                    | 0.02    |
| Within streams      | 256 | 2072.5         | 98                     | <0.001  |

Grevillea ‘Cooljarloo’ in northern, southern and western creek lines and floodplains. A significant isolation by distance relationship was associated with comparisons among

[Table 3. AMOVA for Grevillea sp. Cooljarloo populations at various hierarchical levels. Clusters correspond to the unique genetic clustering from TESS spatial analyses. Clusters include both creek line and corresponding floodplain populations (NPC + CFP34, WPC + CFP(10) and SPC + CFP(10)). Refer to Table 1 for list of upstream and downstream populations.]

Hevroy et al. – Population genetics of a rare Grevillea

References

Chen et al. (2007) suggested evaluating posterior probabilities when using spatial components in assignment testing, and to base success on ability of the programs to assign number of individuals to a cluster with membership above 0.8. Under this criterion, TESS assigned 80 % of individuals to a cluster at \( K = 3 \), outperforming STRUCTURE, which assigned 57 % at \( K = 2 \) and 47 % at \( K = 3 \). Individuals from SPC populations had the greatest mean proportion of membership \( (q_i = 0.98) \), followed by NPC \( (q_i = 0.97) \) and WPC \( (q_i = 0.78) \). Two CFP populations, CFP(10) and CFP(34), showed high levels of admixture. CFP (10) showed highest identity with SPC \( (q_i = 0.42) \) and WPC populations \( (q_i = 0.48) \), whereas CFP (34) had highest identity with NPC \( (q_i = 0.61) \) and SPC \( (q_i =0.26) \) populations. CFP (15) was clustered with SCP populations \( (q_i = 0.99) \).
creek lines, but not within creek lines. Similar levels of genetic differentiation across similarly small spatial scales have been found in many other riparian plant species (Gornall et al. 1998; Prentis et al. 2004; Jacquemyn et al. 2006; Prentis and Mather 2008) suggesting that extensive dispersal is not generally a feature of these systems (as might be expected from hydrochory).

The relatively even distribution of alleles, allelic richness, moderately high heterozygosity and weak genetic differentiation within creek lines does not support a largely unidirectional, source–sink relationship between upstream (source) and downstream (sink) populations (again, as might be expected from hydrochory). This conclusion is consistent with some previous tests of unidirectional dispersal in other riparian systems, where the absence of downstream accumulation of genetic diversity is proposed to be due to upstream seed dispersal through zoochory and higher seed recruitment opportunities in upstream habitats due to density dependence of recruitment (Honnay et al. 2010; Nilsson et al. 2010).

Grevillea ‘Cooljarloo’ has no obvious mechanism for seed dispersal by zoochory, and hydrochory appears to play a minor role in dispersal. Rather, we suggest our results reflect limited, predominantly gravity, dispersal of seeds

Figure 3. A principal coordinates (PCoA) plot of the first two components calculated using Nei’s D genetic distance measure between creek line populations and floodplain populations of Grevillea sp. Cooljarloo. Squares and circles refer to creek line populations and floodplain populations, respectively. Colours correspond to TESS-inferred optimal K clusters, and labels correspond to populations in Table 1. Note the positions reflecting genetic similarity of the three floodplain populations (CFP).

Figure 4. Scatter plot of pairwise population genetic differentiation ($F_{ST}/(1 - F_{ST})$) against log geographic distance among populations of Grevillea sp. Cooljarloo. Comparisons between populations within creek lines are illustrated by coloured squares corresponding to TESS-inferred optimal K clusters. Comparisons between populations among creek lines are illustrated by triangles in green (NPC vs. SPC), purple (NPC vs. WPC) and orange (SPC vs. WPC).
and wide bi-directional movement of pollen along creek line populations by highly mobile nectar-feeding bird pollinators (discussed further below).

The mean genetic differentiation among populations within creeks ($F_{ST} = 0.060$) was significantly lower than that for pairs of populations among creeks ($F_{ST} = 0.110$), which indicates that landscape features among creek lines inhibit gene flow relative to within creek lines. However, the very wide range of pairwise population genetic distance values between creek lines ($F_{ST} = 0.032$–$0.178$), despite an overall similarity in geographic distance within creek line populations (Fig. 4), indicates other factors in addition to distance are influencing gene flow here. In particular, floodplain populations showed surprisingly high genetic similarity ($F_{ST} = 0.030$–$0.054$), much higher than those among creek lines (mean $F_{ST} = 0.110$), despite similar geographic distances. These results suggest that floodplains may be a landscape feature facilitating gene flow independently of distance, and that the non-floodplain landscape between creek lines acts as barriers to gene flow. These results suggest that floodplain flooding events, although less frequent than creek line flow events, provide for more effective and extensive dispersal of seeds. It may also suggest that pollinators are more mobile across floodplains than across creek lines (discussed further below).

Unlike the creek line populations, the admixture evident in the west and north (but not the south) floodplain populations suggests a directional pattern of dispersal to these habitats. This pattern suggests a source–sink relationship from multiple creek line populations into these floodplain populations. Given that multiple sources are involved, the classic source–sink model predicting lower genetic diversity in sink populations coming from a single succession of source populations (Tero et al. 2003) is not evident. Instead, these floodplain populations on average possess more private alleles and similar levels of heterozygosity and allelic richness to creek line populations. Gene flow to floodplains is likely induced by occasional flooding of creek lines and maintained by subsequent animal pollination. South-west Australia experiences a Mediterranean type climate, where the bulk of rainfall occurs through winter and spring prior to seed release by G. ‘Cooljarloo’. Summers typically experience very low rainfall, but high intensity and short duration storms, derived from tropical airflows, which can lead to episodic recruitment events as floodwaters disperse seed, provide moist seedbeds and create gaps in the vegetation (Pettit et al. 2001). Flooding has been shown to play an important role in seed dispersal in many riparian studies (Gornall et al. 1998; Kitamoto et al. 2005; Kudoh et al. 2006; Van Looy et al. 2009).

Figure 5. Bar plots representing admixture coefficients for Grevillea sp. Cooljarloo populations from assignment tests performed (A) in STRUCTURE with $K = 2$; and (B) in STRUCTURE $K = 3$; and (C) with spatial data in TESS with $K = 3$. Each vertical bar represents an individual, and bars are divided into proportions based upon the probability of assignment to each population.
Pollinators, and in particular highly mobile nectar-feeding honeyeater birds, are likely to play a critical role in extensive multidirectional gene flow by pollen in G. ‘Cooljarloo’. Bird assemblages in riparian habitats in Australia tend to be significantly greater in richness, diversity and abundance than adjacent non-riparian sites (Palmer and Bennett 2006). This includes honeyeaters, which are major pollinators for many Grevillea species (Ford et al. 1979; Celebrezze and Paton 2004). Bird pollination can result in far more extensive pollen movement than from insect pollination (Krauss et al. 2009; Phillips et al. 2010; Krauss et al. 2017). We suggest that wide pollen dispersal may act to erode spatial genetic structure arising from predominantly gravity dispersal of seed within creek lines, but this requires further testing through, for example, paternity assignment studies (Krauss et al. 2009).

Conclusions

The distribution of G. ‘Cooljarloo’ is associated with lower elevation habitats that are characterized by floodplains and creek lines. These landscape features strongly influence the structure of genetic variation within and among populations within and among creek lines, where distance is a driver of weak genetic structure within creek lines, but landscape features drive genetic differentiation among creek lines. However, our results also demonstrated a complexity that is consistent with a metapopulation model of dispersal in riparian systems (Tero et al. 2003; Prentis and Mather 2008; Honnay et al. 2010), where geographic distance alone is a poor predictor of genetic distance between populations, and where some, but not all, floodplain populations displayed a surprisingly high level of admixture. The floodplain result suggests a critical role of irregular flooding events in the spatial distribution of genetic variation in this riparian species. Connectivity relating to flood events is an important issue with regards to climate change, as the average rainfall in the southwest of Western Australia has been decreasing over the last 50 years (Bates et al. 2008). This suggests the possibility of a contracting distribution for the riparian G. ‘Cooljarloo’, due not only to landscapes altered by a changing climate, but also because of a reduced capacity for seed dispersal. These water relationships, including flooding, may also be a critical determinant involved in overcoming seed germination inhibitors (Mo et al. 2015) and enabling germination and seedling survival, critical issues requiring further research.

Genetic diversity underpins evolutionary potential of a species and is a measure of population fitness (Spielman et al. 2004). Given the strength of genetic differentiation among the few existing G. ‘Cooljarloo’ populations across creek lines, the loss of any population will have a significant impact on total genetic diversity. NPC populations in particular are under threat of impact from mining (Fig. 2). These populations have relatively high levels of private alleles and show significant differentiation from other creek line populations, indicating a disproportionally high loss of species genetic diversity with the loss of these populations. While overall plant numbers are plentiful in the region, maintaining the genetic structure and metapopulation processes identified through this research will have an important role in preserving the long-term potential of this taxon. Here, avoiding an impact to hydrodynamic processes, such as water tables and flooding dynamics, is just as critical as avoiding direct impacts on the number of plants.

Supporting Information

The following additional information is available in the online version of this article—

Figure S1. Likely number of genetic clusters in Grevillea telemannaniana ssp. Cooljarloo data set from spatial analyses programs based on (A) delta K in STRUCTURE, $K = 2$ and (B) deviance information criterion (DIC) in TESS, $K = 3$.

Sources of Funding

T.H.H. was supported by an Australian Postgraduate Award (APA) and the University Postgraduate Award (UPA) from the University of Western Australia. This work was supported in part by a contract to M.L.M. from Tronox Limited.

Contributions by the Authors

All authors have made a substantial contribution to the paper.

Conflict of Interest

None declared.

Acknowledgements

We would like to acknowledge B. MacDonald and S. McArthur from the Department of Parks and Wildlife (DPaW) and J. Anthony from the Botanical Gardens and Parks Authority (BGPA) for their laboratory assistance and support, P. Nevill (BGPA), J. Sampson, G. Keighery and T. Llorens (DPaW) for their helpful discussions and advice, and D. Coates (DPaW) for support and advice.
Literature Cited

Bates BC, Hope P, Ryan B, Smith L, Charles S 2008. Key findings from the Indian Ocean Climate Initiative and their impact on policy development in Australia. *Climatic Change* 89:339–354.

Cegelski CC, Waits LP, Anderson NJ. 2003. Assessing population structure and gene flow in Montana wolverines (*Gulo gulo*) using assignment-based approaches. *Molecular Ecology* 12:2907–2918.

Celebreze T, Paton DC. 2004. Do introduced honeybees (*Apis mellifera*, Hymenoptera) provide full pollination service to bird-adapted Australian plants with small flowers? An experimental study of *Brachyloma ericoides* (Epacridaceae). *Australian Ecology* 29:129–136.

Chen C, Durand E, Forbes F, François O. 2007. Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study. *Molecular Ecology Notes* 7:747–756.

Corander J, Sirén J, Arjas E. 2008. Bayesian spatial modeling of genetic population structure. *Computational Statistics* 23:111–129.

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

Doyle JJ. 1990. Isolation of plant DNA from fresh tissue. Focus 12:13–15.

Earl D, von Holdt B. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4:359–361.

Ellstrand NC. 2014. Is gene flow the most important evolutionary force in plants? *American Journal of Botany* 101:737–753.

England PR, Whelan RJ, Ayre DJ. 2003. Effects of seed bank disturbance on the fine-scale genetic structure of populations of the rare shrub *Grevillea macleayana*. *Hereditas* 91:475–480.

Epps CW, Weihausen JD, Bleich VC, Torres SG, Brashares JS. 2007. Optimizing dispersal and corridor models using landscape genetics. *Journal of Applied Ecology* 44:714–724.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14:2611–2620.

Excoffier L, Lischer HE. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* 10:564–567.

Ford HA, Paton DC, Forde N. 1979. Birds as pollinators of Australian plants. *New Zealand Journal of Botany* 17:509–519.

François O, Durand E. 2010. Spatially explicit Bayesian clustering models in population genetics. *Molecular Ecology Resources* 10:773–784.

Gornall RJ, Hollingsworth PM, Preston CD. 1998. Evidence for spatial structure and directional gene flow in a population of an aquatic plant, *Potamogeton coloratus*. *Hereditas* 80:414–421.

Guillot G, Estoup A, Mortier F, Cosson JF. 2005. A spatial statistical model for landscape genetics. *Genetics* 170:1261–1280.

Hevroy TH, Moody ML, Krauss SL, Gardner MG. 2013. Isolation, via 454 sequencing, characterization and transferability of microsatellites for *Grevillea thelemanniana* subsp. *thelemanniana* and cross-species amplification in the *Grevillea thelemanniana* complex (Proteaceae). *Conservation Genetics Resources* 5:887–890.

Holderegger R, Buehler D, Guerler F, Manel S. 2010. Landscape genetics of plants. *Trends in Plant Science* 15:675–683.

Honnay O, Jacquemyn H, Nackaerts K, Breyne P, Van Looy K. 2010. Patterns of population genetic diversity in riparian and aquatic plant species along rivers. *Journal of Biogeography* 37:1730–1739.

Hopper SD, Gioia P. 2004. The Southwest Australian Floristic Region: evolution and conservation of a global hot spot of biodiversity. *Annual Review of Ecology, Evolution and Systematics* 35:623–650.

Hu LJ, Uchiyama K, Shen HL, Saito Y, Tsuda Y, Ide Y. 2008. Nuclear DNA microsatellites reveal genetic variation but a lack of phylogeographical structure in an endangered species, *Fraxinus mandshurica*, across North-east China. *Annals of Botany* 102:195–205.

Jacquemyn H, Honnay O, Van Looy K, Breyne P. 2006. Spatiotemporal structure of genetic variation of a spreading plant metapopulation on dynamic riverbanks along the Meuse River. *Heredity* 96:471–478.

Jakobsson M, Rosenberg NA. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23:1801–1806.

Jost L. 2008. G(ST) and its relatives do not measure differentiation. *Molecular Ecology* 17:4015–4026.

Kalinowski ST. 2005. hp-rare 1.0: a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes* 5:187–189.

Kitamoto N, Honjo M, Ueno S, Takenaka A, Tsumura I, Ohsawa R. 2005. Spatial genetic structure among and within populations of *Primula sieboldii* growing beside separate streams. *Molecular Ecology* 14:149–157.

Krauss SL, He T, Barrett LG, Lamont BB, Enright NJ, Miller BP, Hanley ME. 2009. Contrasting impacts of pollen and seed dispersal on spatial genetic structure in the bird-pollinated *Bankia hookeri ana*. *Heredity* 102:274–285.

Krauss SL, Phillips RD, Karron JD, Johnson SD, Roberts DG, Hopper SD. 2017. Novel consequences of bird pollination for plant mating. *Trends in Plant Science* 22:395–410.

Kudoh H, Shimamura R, Takayama K, Whigham DF. 2006. Consequences of hydrochory in *Hibiscus*. *Plant Species Biology* 21:127–133.

Lowe A, Harris S, Ashton P. 2009. *Ecological genetics: design, analysis, and application*. Oxford, UK: Blackwell.

Liu Y, Wang Y, Huang H. 2006. High interpopulation genetic differentiation and unidirectional linear migration patterns in *Myricaria laxiflora* (Tamaricaceae), an endemic riparian plant in the Three Gorges Valley of the Yangtze River. *American Journal of Botany* 93:206–215.

Ma X, Guo J, Han X, Yan G. 2015. *Grevillea* (Proteaceae) seed coats contain inhibitors for seed germination. *Australian Journal of Botany* 63:566–571.

Makinson R. 2000. *Grevillea*. Flora of Australia 17A. Oxford, UK: Blackwell.

Mitsui Y, Isagi Y, Setoguchi H. 2010. Multiple spatial scale patterns of genetic diversity in riparian populations of *Ainsliaea faurieana* (Asteraceae) on Yakushima Island, Japan. *American Journal of Botany* 97:101–110.
Nei M. 1972. Genetic distance between populations. *The American Naturalist* **106**:283–292.

Nilsson C, Brown RL, Jansson R, Merritt DM. 2010. The role of hydrochorous seed dispersal in structuring riparian and wetland vegetation. *Biological Reviews of the Cambridge Philosophical Society* **85**:837–858.

Nistelberger H, Byrne M, Coates DJ, Roberts JD. 2015. Genetic drift drives evolution in the bird-pollinated, terrestrial island endemic *Grevillea georgeana* (Proteaceae). *Botanical Journal of the Linnean Society* **178**:155–168.

Palmer GC, Bennett AF. 2006. Riparian zones provide for distinct bird assemblages in forest mosaics of south-east Australia. *Biological Conservation* **130**:447–457.

Peakall R, Smouse PE. 2012. GenAlEx 6.5: genetic analysis in Excel. *Population genetic software for teaching and research—an update. Bioinformatics* **28**:2537–2539.

Pettit N, Froend R, Davies P. 2001. Identifying the natural flow regime and the relationship with riparian vegetation for two contrasting western Australian rivers. *Regulated Rivers: Research & Management* **17**:201–215.

Phillips RD, Hopper SD, Dixon KW. 2010. Pollination ecology and the possible impacts of environmental change in the southwest Australian biodiversity hotspot. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **365**:517–528.

Prentis PJ, Mather PB. 2008. Fine-scale patterns of genetic variation indicate non-equilibrium gene frequency divergence in the stream lily, *Helmholtzia glaberrima*. *Freshwater Biology* **53**:973–980.

Prentis PJ, Vesey A, Meyers N, Mather P. 2004. Genetic structuring of the stream lily *Helmholtzia glaberrima* (Philydraceae) within Toolona Creek, south-eastern Queensland. *Australian Journal of Botany* **52**:201–207.

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

Roberts DG, Ayre DJ, Whelan RJ. 2007. Urban plants as genetic reservoirs or threats to the integrity of bushland plant populations. *Conservation Biology* **21**:842–852.

Rice WR. 1989. Analyzing tables of statistical tests. *Evolution* **43**:223–225.

Rousset F. 2000. Genetic differentiation between individuals. *Journal of Evolutionary Biology* **13**:58–62.

Row JR, Blouin-Demers G, Lougheed SC. 2010. Habitat distribution influences dispersal and fine-scale genetic population structure of eastern foxsnakes (*Mintonius gloydi*) across a fragmented landscape. *Molecular Ecology* **19**:5157–5171.

Smith M. 2013. Threatened and priority flora list for Western Australia. Kensington, Western Australia: Department of Parks and Wildlife.

Spielman D, Brook BW, Frankham R. 2004. Most species are not driven to extinction before genetic factors impact them. *Proceedings of the National Academy of Sciences USA* **101**:15261–15264.

Storfer A, Murphy MA, Evans JS, Goldberg CS, Robinson S, Spear SF, Dezzani R, Delmelle E, Vierling L, Waits LP. 2007. Putting the “landscape” in landscape genetics. *Heredity* **98**:128–142.

Tero N, Aspi J, Siikamäki P, Jäkäläniemi A. 2005. Local genetic population structure in an endangered plant species, *Silene tatarica* (Caryophyllaceae). *Heredity* **94**:478–487.

Tero N, Aspi J, Siikamäki P, Jäkäläniemi A, Tuomi J. 2003. Genetic structure and gene flow in a metapopulation of an endangered plant species, *Silene tatarica*. *Molecular Ecology* **12**:2073–2085.

Trénel P, Hansen MM, Normand S, Borchsenius F. 2008. Landscape genetics, historical isolation and cross-Andean gene flow in the wax palm, *Ceroxylon echinulatum* (Arecaceae). *Molecular Ecology* **17**:3528–3540.

Van Looy K, Jacquemyn H, Breyne P, Honnay O. 2009. Effects of flood events on the genetic structure of riparian populations of the grassland plant *Origanum vulgare*. *Biological Conservation* **142**:870–878.

Van Oosterhout C, Hutchison WF, Wills DP, Shipley P. 2004. *MICRO-CHECKER*: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4**:535–538.

Wei X, Meng H, Jiang M. 2013. Landscape genetic structure of a streamside tree species *Euptelea pleiospermum* (Eupteleaceae): contrasting roles of river valley and mountain ridge. *PLoS One* **8**:e66928.