ASSESSMENT OF POPULATION GENETIC VARIATION AND STRUCTURE OF *ACACIA WOODMANIORUM*, AND ITS PHYLOGENETIC RELATIONSHIP TO OTHER *ACACIA* SPECIES

Eighteen month report to Karara Mining Ltd by the Department of Environment and Conservation.

April 2010

Melissa A. Millar and David J. Coates

EXECUTIVE SUMMARY:
18 MONTH REPORT

AN ASSESSMENT OF POPULATION GENETIC VARIATION AND STRUCTURE OF *ACACIA WOODMANIORUM*, AND ITS PHYLOGENETICC RELATIONSHIP TO OTHER ACACIA SPECIES.

Phase 1. Assessing levels and partitioning of population genetic variation within *Acacia woodmaniorum* and the evolutionary relationships and distinctiveness of this species to its closest relatives.

1. *The extent and partitioning of population genetic variation within Acacia woodmaniorum*

This section aims to quantify the potential impact of mining scenarios on genetic variation within *Acacia woodmaniorum* and to establish a baseline for future management of genetic variation. Comprehensive germplasm collections have been conducted across the range of *A. woodmaniorum*, DNA extraction protocols refined and species specific genetic markers (microsatellites) optimised for use in *A. woodmaniorum*. 573 individuals from 33 populations across the species range have been genetically characterised with 15 microsatellite loci and analysis of genetic diversity, population genetic variation, spatial genetic structure and an initial assessment of the potential impact of specific mining scenarios on the species conducted.

A comprehensive report on this component of the project is included below. The main findings reveal,

- Overall genetic diversity within *A woodmaniorum* is relatively high and comparable to other Acacia species (including common species) investigated to date.
- Genetic diversity is concentrated in the main population across Mungada/Windaning ridge. The next greatest level of diversity is found in the Jasper Hills region, followed by the Blue Hills and Terapod populations. Interestingly the Blue Hills population contains a significant number of genetic variants not found anywhere else and more unique variants than do populations in Jasper Hills.
- A pattern of increasing genetic difference with increasing geographic separation is found among populations over the species range suggesting that the populations have been stable within this region for a long period of time.
- Isolation by distance is found within the Jasper Hills region, likely due to a level of restricted gene flow among populations associated with the discrete nature of the habitat i.e. the smaller disjunct hematite ranges providing a non continuous habitat.
- There is no isolation by distance among the Mungada/Windaning ridge populations indicating significant pollen movement and gene flow between individuals across this region.

- Genetic differences between populations and regions are low and overall estimates of gene flow via pollen dispersal are relatively high. This result is unexpected and raises the issue of the pollinators and their apparent capacity to move over significant distances.
- Mating within populations appears to be random with little evidence of inbreeding.
- An initial assessment of habitat removal on genetic diversity indicates that the removal of plants at Terapod, Blue Hills and the most westerly populations across the main range of Mungada/Windaning ridge would result in:
    1. A relatively small drop (13%) in total species genetic diversity.
    2. A relatively large drop (45%) in overall unique genetic variation that is genetic variation unique to a single location.
- The loss of unique genetic variation is of particular concern for the Blue Hills population.
- Populations that may be impacted upon by mining activities, particularly those on Mungada ridge, may play an important role in the maintenance of gene flow and genetic continuity among populations and regions given their central location within the species range.

2. Evolutionary relationships and distinctiveness of Acacia woodmaniorum to its closest relatives in the Acacia alata complex

This element of the project focuses on the putative close relatives in the *A. alata* complex. The closest members of this complex are found some 100km to the west of Karara at the Yandanooka Hills and then occur to the north west in the Morseby Range and to the southwest through the jarrah forest to Albany. The relationship between *A. woodmaniorum* and members of this complex are of particular interest as they may help not only understand the origins of *A. woodmaniorum* but also other plant species that are endemic to the mid west banded ironstone ranges.

Germplasm collections and DNA extraction for 20 populations covering the geographic range of the *Acacia alata* species complex have been completed. Initial DNA sequence studies have been carried out by colleagues on *A. woodmaniorum* and *A. alata* at CSIRO Canberra and the University of Melbourne and we will commence a detailed analysis of DNA variation in the 20 populations within the next week.

Phase 2 Determining how key population genetic processes such as mating systems, gene flow and pollination may influence future levels and patterns of genetic variation in *A. woodmaniorum*.

Although Phase 2 is only now due to formally commence, seed pods from the 2009 flowering season were collected in November that year. 721 seed were germinated for studies of paternity (pollen source) within pods and other aspects of the mating system, including determining the level of outcrossing, as well as for paternity analysis and assessment of patterns of pollen dispersal within and among populations. More specifically the molecular DNA markers employed in this study will allow us track the movement of pollen between plants within populations and perhaps more importantly determine how much pollen is coming into the population

from outside and in some cases where it might be coming from. Under different mining scenarios some populations may become more isolated and it would be possible to assess whether gene flow might need to be artificially enhanced between populations to ensure their persistence.

Genetic characterization and paternity analysis of the seedlings will commence within the next 2 months when these plants are sufficiently large enough for the non-destructive harvest of phyllode material.

In the next flowering season we will also be carrying out pollinator observation studies to determine the importance of individual insect species in the pollination process and to attempt to assess their capacity to disperse pollen over significant geographic distances.

18 MONTH REPORT

AN ASSESSMENT OF POPULATION GENETIC VARIATION AND STRUCTURE OF *ACACIA WOODMANIORUM*, AND ITS PHYLOGENETICC RELATIONSHIP TO OTHER ACACIA SPECIES.

Phase 1. Assessing levels and partitioning of population genetic variation within *Acacia woodmaniorum* and the evolutionary relationships and distinctiveness of this species to its closest relatives.

1. The extent and partitioning of population genetic variation within *Acacia woodmaniorum*

Introduction

Acacia woodmaniorum is a rare and extremely geographically restricted sprawling, prickly shrub first described in 1999 (Maslin, Buscomb, 2007). The species occurs on a series of highly weathered low ranges on the BIFs of the Yilgarn craton. This land system is located on the border of the species-rich Transitional Rainfall Zone and the Arid zone in the Midwest region of Western Australia. Field surveys have shown A. woodmaniorum to be restricted to approximately 29,000 plants with high substrate specificity for the skeletal soils and rock crevices of hematite/magnetite rich outcrops. The geographic distribution of *A. woodmaniorum* covers an area less than 40 km2 and is comprised largely of two discrete banded ironstone ranges, Mungada/Windaning Ridge and Jasper Hills, with two smaller populations at the Blue Hills and Terapod sites (Figure 1). The bulk of plants (less than 25,000) occur across the almost continuous habitat of Mungada/Windaning ridge, with several small (from a single plant to 30 plants) 'outlying' populations occurring on small outcrops of hematite off the main ridge. The habitat of Jasper Hills differs with plants (~3,000 in total) occurring as smaller populations (17 to 1047 plants) on a number of disjunct hematite ranges and smaller outcrops.

The amount of genetic diversity or level of genetic variation within populations of *A. woodmaniorum* and contained within the species overall is a fundamental parameter of the species evolutionary history and conservation biology. High levels of genetic diversity are required in order to maintain the health of individuals and, in the longer term, the ability of populations to respond to forces of selection and a changing environment. In this way, the maintenance of high levels of population genetic diversity acts to maintain the evolutionary potential of species and long term survival.

Loss of population and overall species genetic diversity is predicted with the direct destruction of individuals and of habitat. The degree of loss is related to the severity of habitat destruction and duration of reduced population size. Mining and other activities that include the direct destruction of individuals will result in reduced genetic diversity via the immediate loss of alleles, or allelic richness (the simplest measure of genetic diversity at a locus). The direct destruction of individuals will also result in reduced private allele richness (the number of unique alleles in a

population) which is a measure of genetic distinctiveness of populations. Rare alleles (those that occur at low frequency) will initially be lost from populations, then, more common alleles will be lost.

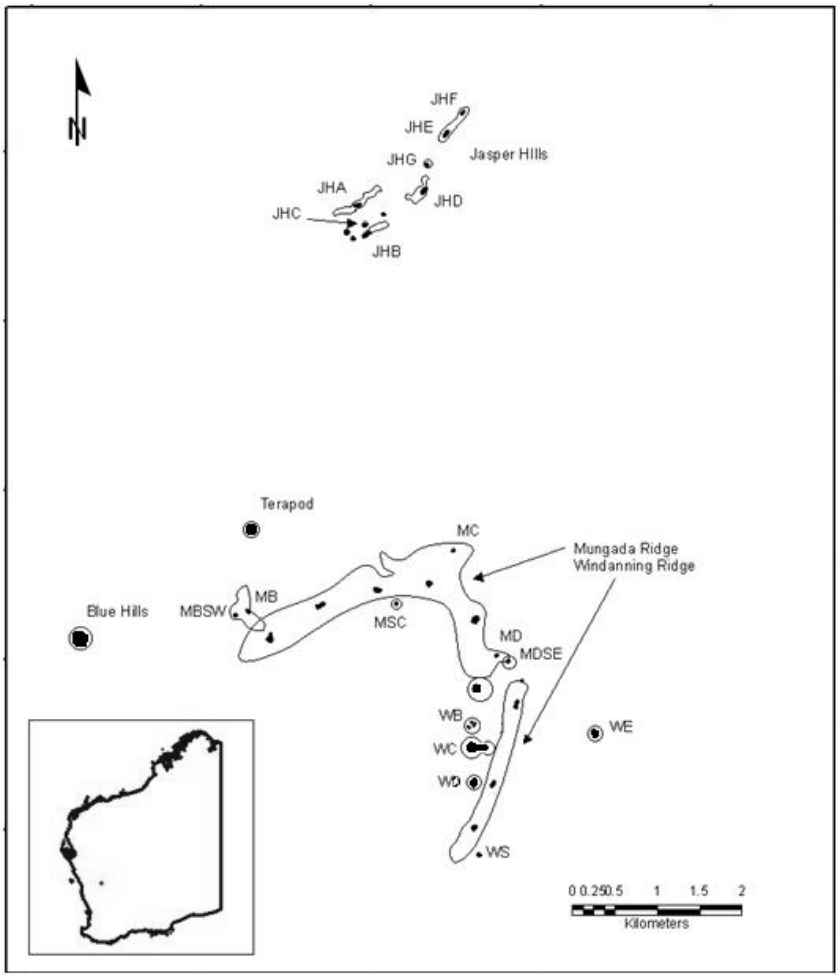


Figure 1 Distribution of *Acacia woodmaniorum* (open shapes) showing regions and locations of sampled populations (dark circles).

Depending on the severity and duration of reduced population size (or genetic bottleneck), genetic diversity may become severely depleted (Ellstrand, Ellam, 1993). Decreased population size is likely to correspond to increased geographic distance and isolation between populations and reduced gene flow between populations. Combined with the effects of random genetic drift, allelic loss and reduced levels of gene flow may result in increased levels of genetic differentiation among disturbed populations (Slatkin, 1987; Young et al., 1996). Small populations may also suffer the negative genetic effects associated with increased levels of inbreeding or mating among related individuals.

Levels of genetic diversity within populations are the result of various processes including forces of selection, genetic mutation, random genetic drift, and the patterns of migration of individuals and of mating. The impacts of various combinations of these forces on *A. woodmaniorum* are unknown and are difficult to determine for a given species. Despite this, broad predictions can be made of the levels and patterns of genetic diversity that may be expected in *A. woodmaniorum*. These

predictions are based largely on the geographical distribution and size of populations and the resulting patterns of migration and mating that may be expected among and within populations. As a rare endemic species with a highly restricted habitat we may expect *Acacia woodmaniorum* to have lower levels of genetic diversity than is found in more widespread and common taxa of the genus. Larger populations are expected to maintain higher levels of genetic diversity and to maintain random patterns of mating. Smaller and more isolated populations are expected to show a degree of genetic impoverishment, lower levels of gene flow or immigration, and a degree of inbreeding due to mating within a restricted number of related individuals
Given the highly specific habitat occupied by *A. woodmaniorum*, and the disjunct nature of the hematite ridges on which it occurs, we expect a degree of genetic differentiation among populations, with increasing levels of genetic differentiation among populations as geographic distance increases.

In this study, we assess the levels of genetic diversity within *Acacia woodmaniorum* and the patterns of genetic differentiation or genetic structure among populations across the species range. We determine a range of parameters describing the patterns of gene flow and genetic connectivity among populations. This information establishes a baseline for future management of *A. woodmaniorum* and gives insight into the processes shaping levels and patterns of genetic diversity in narrow endemic flora of the banded ironstone formations. We also determine the potential impact of several scenarios of population loss on measures of short term diversity loss, including allelic richness and private allele richness, in populations of *A. woodmaniorum*.

Materials and Methods

Sample collection and DNA extraction

Phyllodes were collected from 573 individual plants from across the species range. The sampling of 566 individuals was originally reported but subsequent field trips for seed pod collection revealed 7 additional individuals located in small sites to be used for paternity analysis and gene flow studies. These seven individuals were sampled and the total number of individuals used in the study was 573 (Table 1). Numbers of individuals sampled in each region and population are provided in Table 1. Twenty individuals were sampled at six sites across the large continuous region of Mungada Ridge (MA1-MA6) /Windaning Ridge (WA1-WA3), and from the larger populations on the more disjunct ironstone outcrops of the Jasper Hills region (JHA, JHB, JHD and JHE). Five individuals each were sampled from smaller populations (MB-MD, WC-WD and JHC and JHF-JHG). The smallest 'outlying' populations located off Mungada/Windaning ridge (WD, WE and 'outliers') were exhaustively sampled, as were the Blue Hills and Terapod regions. These sites will be used in future work to delineate patterns of pollen dispersal within the species. The position of all sampled individuals was recorded using a differential Global Positioning Satellite (dGPS) and mapped using Geographical Information Software (GIS) Arcmap™ 9.1 (Figure 1). Fresh phyllode samples were lyophilised and DNA extracted following Millar (2009). All individuals were genotyped using fifteen selected microsatellite primer pairs with PCR conditions previously described for *Acacia woodmaniorum* (Millar, 2009). 16 primer pairs were initially selected for genotyping the main germplasm collection however one primer pair (AwC008) failed

to amplify consistently in previously untested populations and its use was discontinued.

Table 1 Region, population and the number of individuals of *Acacia woodmaniorum* sampled.

Region	Population and number of individuals sampled										Total
Mungada Ridge/ Windaning Ridge	MA1-6	MAB	MAC	MAD	WA1-3	WAB	WAC	WAD	WAE	Outliers	
	120	5	5	5	60	5	5	29	16	30	280
Jasper Hills	JHA	JHB	JHC	JHD	JHE	JHF	JHG			Outliers	
	20	20	5	20	20	5	5			43	138
Blue Hills						145 (all)					
Terapod						10 (all)					
Total											573

Statistical analysis

Utility of loci

Microsatellite genotypes were tested for departures from Hardy-Weinberg Equilibrium (HWE) and linkage disequilibrium (LD) between locus pairs within each region for the 15 loci using exact tests as implemented in the computer program GENEPOP version 4.0 (Raymond, Rousset, 1995). Loci are tested for HWE as microsatellite markers are assumed to be neutral genetic markers not affected by the forces of selection. If a locus is not in HWE selection may be acting on allele and genotype frequencies in that population and population genetic structure may be overestimated. It is also important that microsatellite loci not be in LD. LD means two or more loci are located close together on the same chromosome. In this case alleles will not be independent of each other and will be inherited together during meiosis. Thus the information two linked loci provide becomes redundant for one of the loci. Sequential Bonferroni corrections were applied to alpha values in the determination of significance to correct for multiple comparisons of HWE and LD (Rice, 1989). Loci were assessed for null alleles, which are undetected alleles that fail to amplify, within populations and for the species overall using GENEPOP (Raymond, Rousset, 1995). The presence of null alleles may lead to reduced estimates of allelic richness and overestimation of homozygosity within populations and loci with high frequencies of null alleles are not suitable for population genetic analysis.

Genetic diversity within **A. woodmaniorum**

Genetic diversity parameters including the total number of alleles, number of private alleles, means of the number of alleles per locus (*A*), the number of effective alleles (*Ne*), the proportion of polymorphic loci (*P*), expected heterozygosity (*He*) and observed heterozygosity (*Ho*), were assessed for all populations within regions, and each of the four regions using the program GenAlEx 6.3 (Peakall, Smouse, 2006). To compensate for the bias associated with samples of different size, rarefaction was conducted on measures of allelic richness and number of private alleles (Kalinowski, 2004). This statistical procedure removes bias by standardising for sample size and was implemented using the algorithms in the computer program HP-RARE (Kalinowski, 2005). Sampling was standardised to 20 genes per population with one population per region. Levels of genetic diversity were hierarchically partitioned using the Analysis of Molecular Variance (AMOVA) method implemented

in GenAlEx. Analysis was conducted on F_{ST} values at the individual, population and regional level. Statistical testing was conducted using 999 random permutations.

Genetic differentiation among populations

Genetic differentiation among regions of *A. woodmaniorum* was investigated on the basis of allele identity using pairwise estimates of the intra-class kinship coefficient of Weir and Cockerham (F_{ST}, 1984), and on the basis of differences in allele size assuming stepwise mutation using pairwise R_{ST} values (Slatkin, 1995). Both F_{ST} and R_{ST} are measures of the proportion of total genetic diversity that separates populations. Low values of F_{ST} suggest little genetic differentiation between populations and, if all populations were in HWE with the same allele frequencies, values equal zero. Pairwise values of F_{ST} and R_{ST} were estimated among regions using GENEPOP. The overall F_{ST} among regions, among all populations, and among populations within each multi-population region (Jasper Hills and Mungada/Windaning ridge) was also estimated using GDA (Lewis, Zaykin, 2001) and 95% confidence intervals obtained with 1000 bootstraps.

In order to more clearly visualise the levels of genetic differentiation among populations, the ordination technique of multivariate principal component analysis was also conducted. Covariance genetic distance matrices were constructed by AMOVA on allele frequencies (Excoffier *et al.*, 1992), with 99 randomization steps, using GenAlEx. The matrices were then ordinated in a three dimensional space by principal component analysis (PCA) using a standardised data set and plotted using Statistica software (StatSoft, 2001). Analysis was conducted on both regional and population scales.

Genetic differentiation among populations across the species range was also investigated using the more complex Bayesian method of Pritchard et al. (2000) and Falush et al. (2003), as implemented in the program STRUCTURE version 2.3. This method identifies genetically distinct clusters (*K*) based on allele frequencies. Runs to determine K were conducted with a burn-in period of 10 000, followed by 100 000 iterations. An ancestory model incorporating admixture and correlated allele frequencies was used. This method of analysis allows for the provision of spatial information on populations and we ran the model using prior information on population origin. Ten runs were conducted for each K from K = 1 to K = 6 and the second order rate of change in the estimated Ln probability of the data (ΔK), determined following Evanno et al. (2005), and using the program Structure Harvester v0.56.3 (Earl, 2009). ΔK was then plotted against K to determine the optimal number of clusters. The mean value (Q) of the proportion of membership of individuals (q_i) in each cluster was obtained.

Spatial genetic structure

In order to investigate isolation by distance, Mantel tests were conducted for multi-population regions as well as for all populations across the species range, using GenAlEx 6.3. Linear regression was conducted on matrices of pairwise values of linearised F_{ST} ( ) and of the log geographic distance (Rousset, 1997; Slatkin, 1993). Significance tests for correlation were conducted using 999 random permutations ($p(r_{xy}\text{-random} \geq r_{xy}\text{-data})$(Smouse, Long, 1992; Smouse *et al.*, 1986). In order to investigate the distances over which spatial genetic structure occurs, we conducted linear regression of pairwise population F_{ST}, on geographic distance classes using SPAGeDI 1.3. We conducted significance tests using 10 000

permutations for F_{ST} matrices on 8, 10 or 12 Euclidian distance classes calculated from spatial coordinates. The most appropriate number of distance classes was selected by comparing the proportion of all populations represented in each class and the coefficient of variation of the number of times each population was represented in that distance class. Analysis was conducted for all populations across the species range and for populations within Jasper Hills.

Mating patterns and population sizes

The fixation index (F), a measure of departure from random mating due to inbreeding, was assessed for all populations within regions, and for each of the four regions using GenAlEx. Values of population genetic differentiation (F_{ST}) were converted to the number of migrants following Slatkins island model.

In populations with recently reduced numbers, the level of allelic diversity is reduced faster than the level of expected heterozygosity, and expected levels of heterozygosity will be greater than levels of heterozygosity that may be expected under drift/gene flow equilibrium (Cornuet, Luikart, 1996). To determine whether populations have experienced recent reductions in size, we tested for expected heterozygotic excess at a regional scale as implemented in Bottleneck 1.2.0.2 (Piry *et al.*, 1999). Equilibrium conditions were simulated using 1000 replications assuming a Two-phased model of mutation (TPM) and significance tested using Wilcoxon signed-rank tests.

Impacts of various scenarios

In order to assess the impact of direct destruction of plants of *A. woodmaniorum* on short term levels of genetic diversity, we re-determined measures of allelic richness, private allele richness and heterozygosity for various scenarios of population destruction and compared the results to the baseline data. We assessed five scenarios a) the loss of ten plants at Terapod, b) the loss of the Blue Hills population and c) the loss of the three most westerly populations of the main range (MA1-3) d) the loss of the three most westerly populations of the main range and all plants at MB and MBSW e) the loss of all above mentioned populations. These scenarios were chosen as it seems likely these populations may be at greatest risk of removal by mining activities.

Results

Utility of loci

Fourteen of the fifteen microsatellite loci were polymorphic in all populations of *A. woodmaniourm*. Locus AwD012 was monomorphic in Terapod but polymorphic in other populations. A total of 199 alleles were detected across 15 loci, ranging from a maximum of 23 alleles in AwB109 to a minimum of 5 alleles in AwD010. After adjusting for multiple comparisons (n = 495) significant departures from HWE, in the form of heterozygote excess, were observed for locus AwC001 in most sampling locations across the main range (MA1, 3, 4, 5, 6 and WA1, 2), some larger populations (JHA, JHB, JHD, JHE and Blue Hills), and in two smaller populations (WD and JHBS). Significant departures from HWE in the form of heterozygote deficiency were observed for one or two loci in populations JHB (AwB003), MA1,

MA2, MA4 and WA2 (AwB109), WA3 (AwB107), WD (AwA129 and AwB109), and JHBS (AwB003 and AwB009) and for nine loci in the Blue Hills population (AwA124, AwA129, AwD008, AwB109, AwD010, AwB001, AwB003, AwB107 and AwD012). After correcting for multiple comparisons (n = 3465) significant LD was detected for one locus pair combination (Aw124 and AwD010) in the Blue Hills population. Estimates of Null allele frequencies were low (<0.010) for most loci, reaching a maximum of 0.2441 for D010 in Blue Hills. Given the lack of significant LD and low estimates of the frequency of null alleles we conclude that the loci are suitable for the analysis of population genetic diversity and structure.

Genetic diversity within *A. woodmaniorum*

The total number of alleles observed in regions for *A. woodmaniorum* ranged from a maximum of 171 alleles across Mungada/Windaning ridge to a minimum of 62 alleles for the ten plants at Terapod. The mean number of alleles per locus for populations within regions ranged from a maximum in Blue Hills (8.53) to a minimum in Jasper Hills (3.96, Table 2). At a regional level the mean number of alleles per locus was greatest across Mungada/Windaning ridge (11.40) and lowest for Terapod (4.133). The mean number of effective alleles per locus showed a similar pattern, being highest for Blue Hills (3.102) and lowest for populations in Jasper Hills (2.451), and highest for Mungada/Windaning ridge (3.784) and lowest for Terapod (2.806) at a regional level. The proportion of polymorphic loci was greatest for Blue Hills and lowest for Terapod.

Measures of allelic richness within populations are biased by the number of individuals sampled from each population, as well as the number of populations sampled from each region. Larger populations or populations in which more individuals are sampled are expected to contain more alleles than smaller populations or single individuals. Regions with more populations or sampling locations are also expected to hold more allelic variation than regions with fewer populations. In order to remove this bias rarefaction was conducted on measures of allelic diversity per locus. When the number of alleles per locus within regions were standardised to one population with 20 genes, the number of alleles observed remained greatest for the main population across Mungada/Windaning ridge (5.13) and lowest for Terapod (4.13).

Table 2 Microsatellite diversity in *Acacia woodmaniorum*. Means are shown for the number of individuals sampled (*N*) and genetic diversity parameters including the number of alleles per locus (*A*), the number of effective alleles (*Ne*) the proportion of polymorphic loci (*P*), expected heterozygosity (*He*), observed heterozygosity (*Ho*) and the fixation index (F) for populations within regions and for regions. Values in parenthesis are standard errors.

Region		*N*	*A*	*Ne*	*P*	*He*	*Ho*	F
Jasper Hills	Populations	13.533 (0.599)	3.960 (0.176)	2.451 (0.100)	96.00 (2.67)	0.510 (0.017)	0.536 (0.023)	-0.050 (0.030)
	Region	135.33 (0.637)	7.933 (1.071)	2.984 (0.414)	100.00(0.00)	0.588 (0.044)	0.522 (0.054)	-
Mungada/ Windaning Ridge	Populations	11.575 (0.401)	4.317 (0.130)	2.759 (0.084)	94.60 (1.99)	0.537 (0.012)	0.527 (0.016)	0.004 (0.024)
	Region	243.07 (8.759)	11.40 (1.264)	3.784 (0.660)	100.00 (0.00)	0.627 (0.051)	0.521 (0.049)	-

Blue Hills	Populations	145 (0.00)	8.533 (1.162)	3.102 (0.574)	100.00 (0.00)	0.573 (0.051)	0.509 (0.061)	0.106 (0.080)
	Region	134.13 (3.070)	8.533 (1.162)	3.102 (0.574)	100.00 (0.00)	0.573 (0.051)	0.509 (0.061)	-
Terapod	Populations	10 (0.00)	4.133 (0.496)	2.806 (0.303)	93.33 (0.00)	0.564 (0.057)	0.533 (0.067)	0.043 (0.076)
	Region	9.6 (0.273)	4.133 (0.496)	2.806 (0.303)	93.33 (0.00)	0.564 (0.057)	0.533 (0.067)	-
All populations		15.822 (0.997)	4.331 (0.111)	2.677 (0.064)	95.15 (1.49)	0.531 (0.010)	0.529 (0.013)	-
Species overall		522.133 (11.790)	13.267 (1.631)	3.781 (0.709)	100.00 (0.00)	0.6431 (0.049)	0.516 (0.049)	0.159 (0.071)

Another measure of genetic diversity at microsatellite loci is the expected heterozygosity (H_e) or overall gene diversity. Values of expected heterozygosity were similar for all regions, but again, greatest for the main range across Mungada/Windaning ridge (0.627, Table 2) and lowest at Terapod (0.564).

Comprehensive studies of genetic diversity using microsatellite markers have been conducted for few Acacia taxa, despite their ecological, cultural and economic importance. Levels of microsatellite genetic diversity within *A. woodmaniorum*, including the number of alleles per locus, effective number of alleles per locus, percentage of polymorphic loci, and observed and expected heterozygosities, are however all greater than those found in a comprehensive study of *A. saligna*, a common species with a widespread distribution across south western Australia (M.A. Millar, unpublished results).

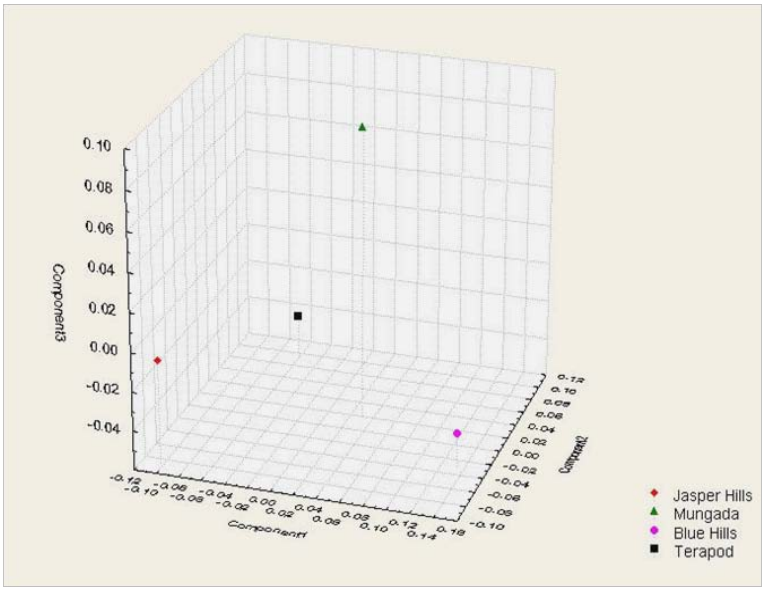
Genetic differentiation among populations

A measure of the distinctiveness of populations is the number of private alleles they contain. Private alleles are alleles present in a given population and not found in any other population. Private alleles were abundant in populations of *A. woodmaniorum*, being detected in 24 of the 33 sampling locations, with the exceptions being JHBW, JHC, JHD, JHF, MA4, MASC, MDSE, WN and WS. All regions contained private alleles; 28 were detected across the main Mungada/Windaning ridge, 12 alleles were unique to Blue Hills, 8 alleles were unique to the Jasper Hills region, and one unique allele was observed at Terapod. The number of private alleles within populations can also be expected to be biased by the number of individuals sampled from each population and the number of populations sampled from each region. However, when rarefaction was conducted on the number of private alleles per region for *A. woodmaniorum*, the number of private alleles remained greatest for Mungada/Windaning ridge (17.4), decreasing through Blue Hills (13.4) and Jasper Hills (11.8) to Terapod (6.4). Values for sampled populations within each region did change (data not presented).

F_{ST} is another measure of the genetic distinctiveness of populations, based on allele identity or allele frequency. A low level of genetic differentiation is evident among populations of *A. woodmaniorum* across the species range (F_{ST} = 0.0977) and genetic differentiation at a regional scale was lower (F_{ST} = 0.0677). The Blue Hills population shows the greatest degree of genetic differentiation from the other regions (Table 3). On a regional scale, genetic differentiation is greatest between Jasper Hills and Blue Hills (F_{ST} = 0.109), low between Jasper Hills and Terapod (F_{ST}

= 0.056) and lowest between Jasper Hills and the main range on Mungada/Windaning ridge (F_{ST} = 0.048). This pattern of genetic differentiation is not surprising, as divergence in allele identity or frequency is a measure of genetic drift among populations and greater levels of genetic drift and genetic differentiation are expected between populations that are more geographically isolated from each other. Populations in the Jasper Hills region are located geographically furtherest from Blue Hills (approximately 5.75km, see Table 4), closer to the main range (4.12 km), and closest to those at Terapod (3.64 km). The population at Blue Hills is most genetically similar to its geographically closest population - the main range of Mungada/Windaning ridge (F_{ST} = 0.061, 1.87 km), and plants at Terapod are most genetically similar to their closest population which is the main range on Mungada/Windaning ridge (F_{ST} = 0.050, 0.66 km).

a



b

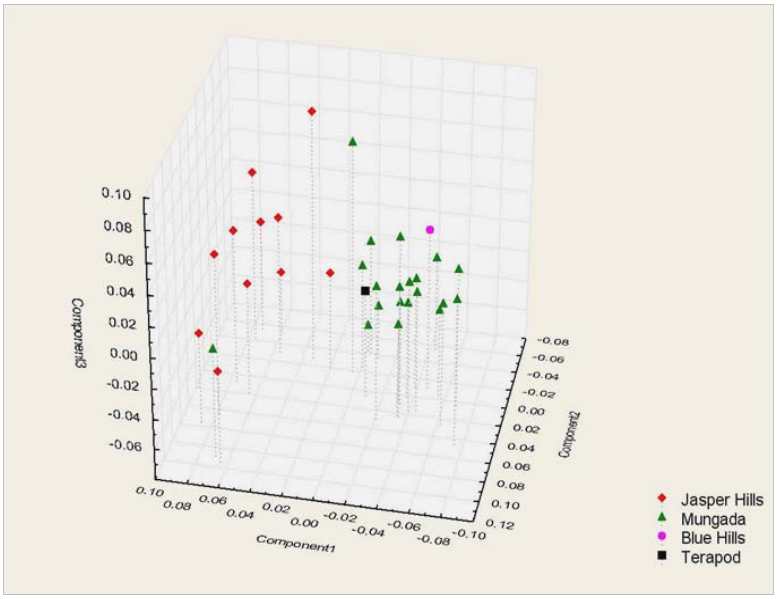


Figure 2. Principle component analysis showing a) the genetic relationships among *Acacia woodmaniorum* a) regions and b) populations.

Table 3. Pairwise F_{ST} values below the diagonal and R_{ST} values above the diagonal for regions of *Acacia woodmaniorum*.

	Jasper Hills	Mungada/Windaning	Blue Hills	Terapod
Jasper Hills	-	0.132	0.186	0.106
Mungada/Windaning	0.048	-	0.046	0.092
Blue Hills	0.109	0.061	-	0.121
Terapod	0.056	0.050	0.097	-

Genetic differentiation among populations is further illustrated in the PCA diagrams (Figure 2). At a regional scale, the first three principal components explained 100% of the variation seen; component 1 53.34%, component 2 23.38% and component 3 18.32% (Figure 2a). By populations the first three principal components explained 71.86% of the variation seen; component 1 37.69%, component 2 21.56% and component 3 12.62% (Figure 2b).

Genetic differentiation among regions was also calculated based on allele size (R_{ST}, Table 3). This measure indicates the relative contribution of stepwise mutation to genetic divergence among populations. Most pairwise population estimates of R_{ST} were greater than estimates of divergence based on allele identity, indicating that stepwise mutation in microsatellite repeats may contribute more to genetic differentiation among regions of *A. woodmaniorum* than random genetic drift does. On the basis of allele size, genetic differentiation is again greatest between Blue Hills and Jasper Hills (R_{ST} = 0.186). Jasper Hills and Mungada/Windaning ridge show greater levels of genetic differentiation due to mutation than due to drift however, and Blue Hills and Mungada/Windaning ridge, as well as Blue Hills and Terapod show lower levels of differentiation due to mutation than due to drift.

Hierarchically, the majority of genetic diversity in *A. woodmaniourm* occurs within individuals. Sixty nine percent of all genetic variation is contained within populations, 24% is contained among populations and 7% is contained among regions. Lower genetic diversity among populations and regions is reflected in the relatively low values of pairwise estimates of differentiation among populations and regions. These results suggest that significant levels of gene flow via pollen dispersal are maintained, and indeed estimates of the number of migrants among populations and regions are high.

Genetic differentiation among populations across the species range was also investigated using the complex Bayesian clustering method. This technique clusters populations on the assumptions that genotypes are at HWE, loci are not in LD and a minimum of between cluster genetic differentiation is present, that is, allele frequencies are cluster specific. Bayesian analysis optimally placed populations into one of two clusters representing Jasper Hills, Mungada/Windaning ridge and Terapod combined (cluster 1), or Blue Hills (cluster 2). Values of mean estimated Ln probability of the data (Ln P(D)) rose steadily with K until K = 3, after which standard errors about the mean increased (Figure 3a). A plot of ΔK against K revealed a peak at K = 2 and this was taken as the optimal number of clusters (Figure 3b). Figure 3c shows a cluster plot of individuals grouped into one of the two clusters with average proportion of membership values (Q) of 0.900 for Jasper Hills, Mungada/Windaning ridge and Terapod, and of 0.874 for Blue Hills. A number of populations across

Mungada/Windaning ridge (MA1-4, MB, MBW, MC, MA6, WA2, WC and WE) showed some genetic affiliation to Blue Hills (maximum Q for Blue Hills was for MABW, Q = 0.329 and the maximum q_i was for an individual from MA2, q_i = 0.625).

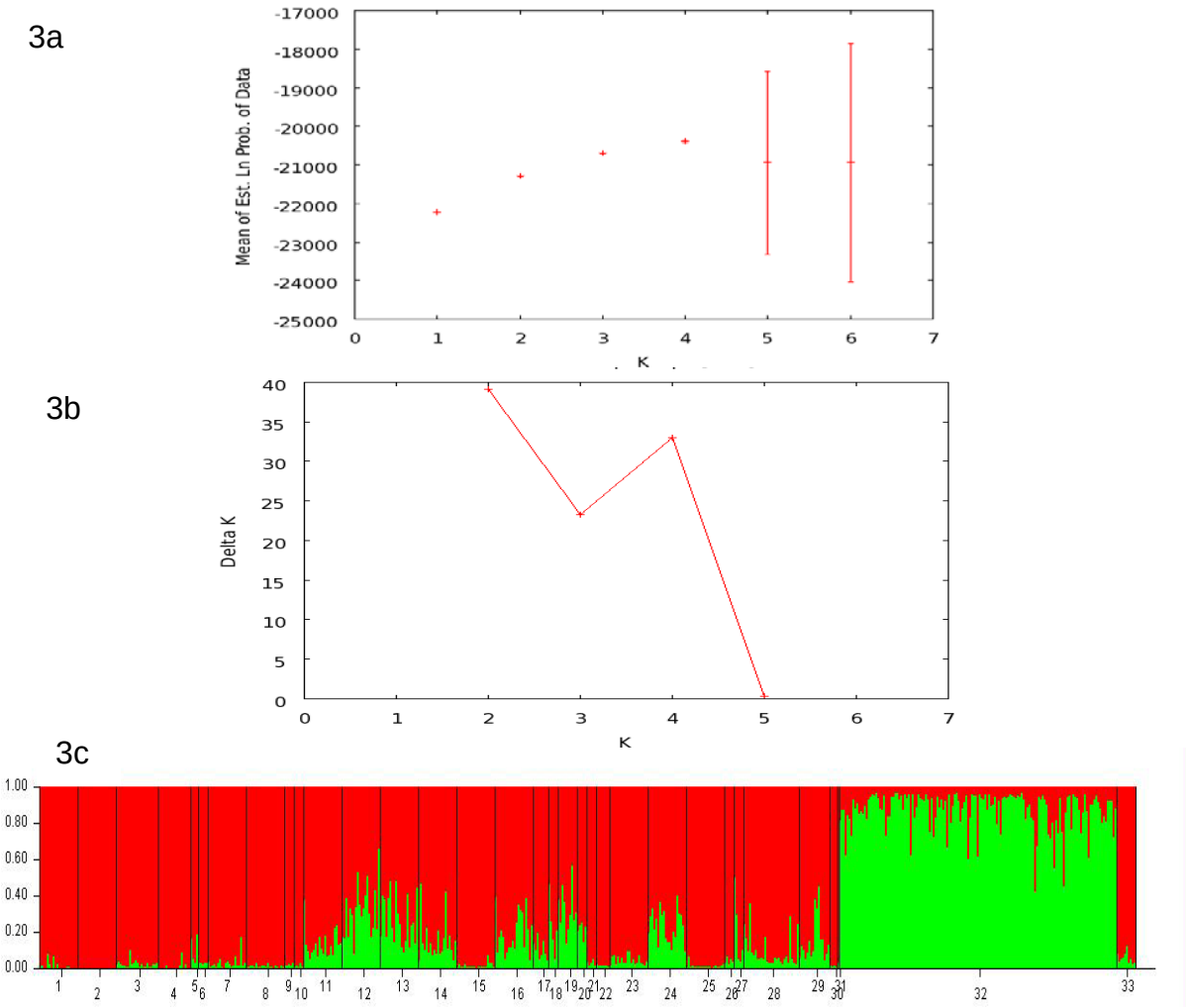


Figure 3 STRUCTURE analyses of sampled individuals of *Acacia woodmaniorum*. a) plot of Ln P(D) against increasing K b) plot of Evannos delta K with increasing K showing a maximum at K = 2 and c) a cluster plot showing the proportion of all sampled individuals in each of two clusters. Populations 1 to 10 are from Jasper Hills, populations 11 to 31 are from Mungada/Windaning ridge, population 32 is Blue Hills and population 33 is Terapod.

Spatial genetic structure

Correlation between genetic and geographic distance was evident for populations across the entire species range (p = 0.001, R^2 = 0.0688, Figure 4a), suggesting an extant populations have persisted long enough to achieve equilibrium between drift and gene flow across the species distribution. Isolation by distance was revealed within the Jasper Hills region (p = 0.003, R^2 = 0.222, Figure 4b) suggesting that in this region the species has had sufficient time to achieve an equilibrium between the effects of genetic drift and gene flow. In Jasper Hills the influence of limited gene flow spreads to all degrees of geographical separation. This result may be attributed to landscape features, namely the disjunct nature of the ironstone ridges in this area that may act to inhibit gene flow via pollen or seed dispersal among populations. At a regional scale, a lack of correlation between genetic and geographic distance was detected across the species main distribution on Mungada/Windaning ridge (p =

0.075, R^2 = 0.0233, Figure 4c), suggesting this large, relatively continuous population is not in drift/gene flow equilibrium. The lack of isolation by distance across the main ridge suggests that that gene flow plays a greater role in shaping genetic structure on the ridge than random genetic drift does and that dispersal via gene flow is not limited in this relatively homogeneous population.

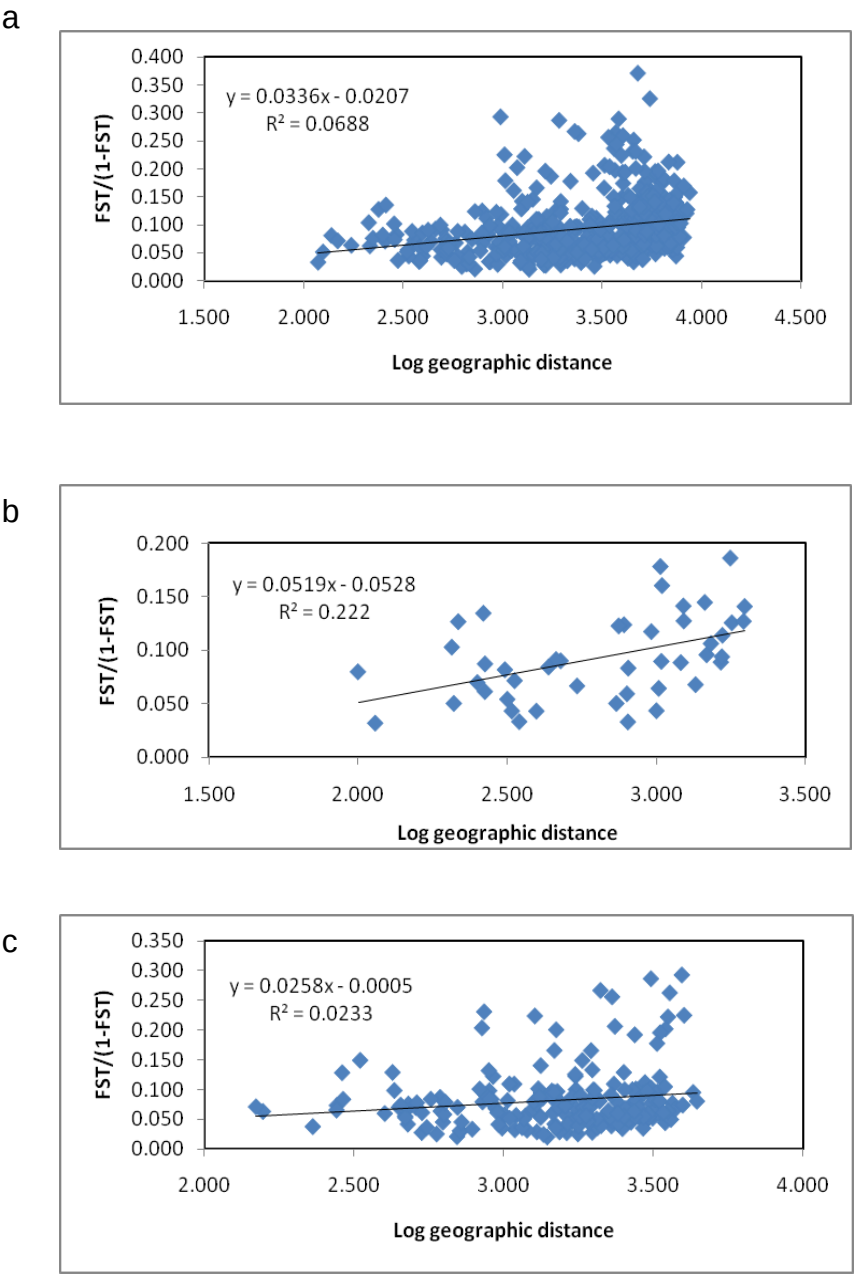
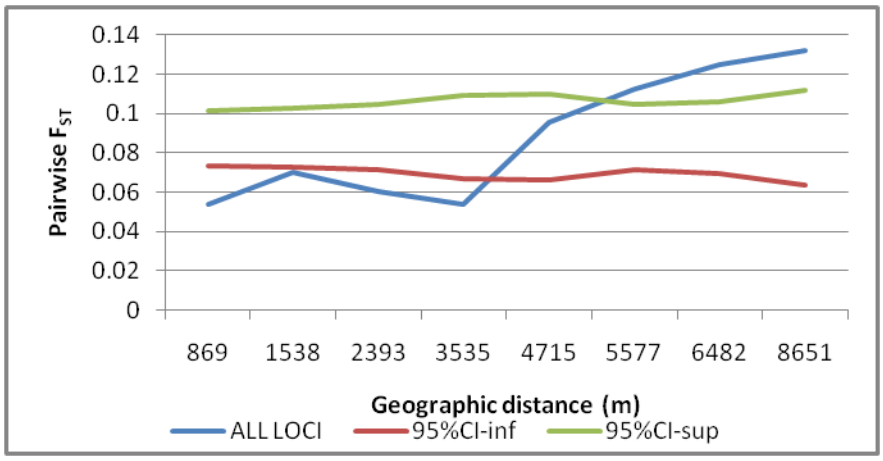


Figure 4. Isolation by distance among (a) all populations of *Acacia woodmaniorum* and (b) among populations within the Jasper Hill and (c) Mungada/Windaning ridge regions. Multilocus estimates of pairwise differentiation are plotted against logarithm of geographic distances.

Spatial autocorrelation analysis of populations across the species range revealed significant correlation of genetic distance and geographic distance at seven of eight optimal geographic distance classes (Figure 5a). Correlation was not significant at distances of 3535 m to 4715 m. Within Jasper Hills, autocorrelation over eight optimal distance classes revealed significant correlation at distances of 307 m to 398 m and at distances of 1540 m or greater ($p < 0.05$).

a



b

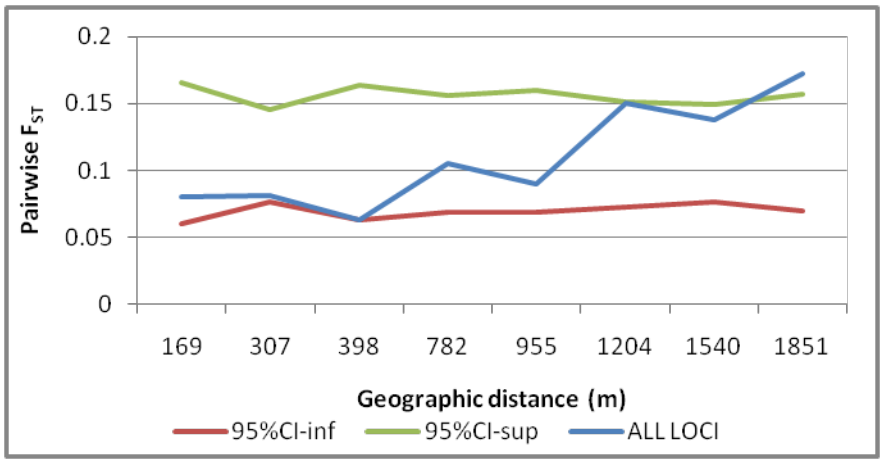


Figure 5. Spatial autocorrelation of a) all sampled populations and b) populations within the Jasper Hills region of *Acacia woodmaniorum*.

Estimates of the fixation index indicate random mating with little inbreeding within populations and regions and this result also suggests that pollen dispersal is not limited in *A. woodmaniorum* (Table 2). Values of the fixation index close to zero indicate random mating with little inbreeding or negative assortative mating within *A. woodmaniourm.* Values of the fixation index close to zero are expected under random mating, with substantial positive values indicating a degree of inbreeding, or breeding among related individuals and negative values indicating a degree of negative assortative mating or heterotic selection.

Random mating with a degree of pollen flow acting to maintain homogeneity among populations is also evident by the high estimates of the number of migrants per generation (N_m). These values are directly related to pairwise estimates of F_{ST} and so to the geographic distance between populations. Regional pairwise values of Nm and of geographic distance between populations are provided in Table 4. Overall 2.31 individuals per generation are estimated to result from migration of either pollen or seed among populations. On a regional scale 3.44 migrants are estimated per generation.

Table 4. Pairwise N_m values below the diagonal for regions of *Acacia woodmaniorum*. Pairwise geographic distances in kilometers above the diagonal.

	Jasper Hills	Mungada/Windaning	Blue Hills	Terapod
Jasper Hills	-	4.12	5.75	3.64
Mungada/Windaning	4.96	-	1.87	0.660
Blue Hills	2.04	3.85	-	2.42
Terapod	4.21	4.75	2.33	-

No evidence of heterozygotic excess was found within regions of *A. woodmaniorum* indicating that population sizes have not been recently reduced and that populations are at mutation/drift equilibrium. A mode shift in allele sizes was detected for plants at Terapod but this is likely to be a result of small population size for which the qualitative method of analysis is unreliable.

Impacts of various scenarios

The distribution of genetic variation measured as allelic richness (number of alleles) and private allele richness (alleles unique to a region) is shown in Fig. 6. The removal of habitat and therefore plants will negatively impact levels of genetic diversity, including levels of allelic richness, private allele richness and heterozygosity within *A. woodmaniorum* (Table 5). Analysis of baseline data after the removal of certain populations reveals that the loss of the ten plants at Terapod will result in a reduction of 0.05% of the species total allelic diversity and one allele private to that population or 2.04% of private alleles. Loss of the Blue Hills population will reduce total allelic diversity by 6.03% and result in the loss of 24.49% of private alleles; the loss of the three most westerly populations of the main range (Mungada – Windaning Ridge, MA1-3) will reduce total allelic diversity by 4.02% and result in the loss of 14.29% of private alleles. If all of the listed populations (Table 5) are removed, 12.56% of the species total allelic diversity will be lost as will 44.90% of private alleles. This represents a significant amount of the species diversity and may impact on its ability for future adaptation and persistence although this is difficult to quantify. Levels of heterozygosity will be reduced slowly with the loss of increasing numbers of individuals due to the high level of polymorphism in the markers. The populations targeted in the above scenarios are located from the western margin to the center of the species distribution. If large numbers of plants are removed from these central locations this may impact not only levels of genetic diversity but also the maintenance by pollen flow among populations as populations become increasingly isolated.

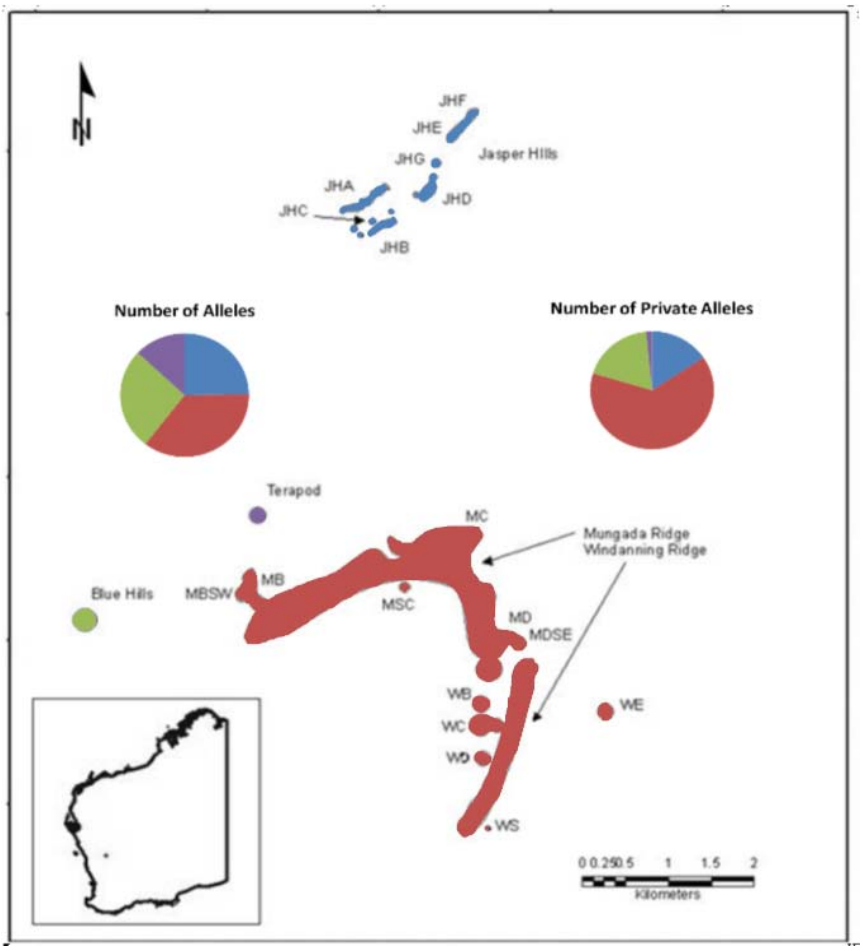


Figure 6. Allelic diversity (number of alleles) and number of private alleles (alleles only found in that region) for the four regions Mungada- Windanning Ridge, Blue Hills, Terapod, and Jasper hills

Table 5. Predicted microsatellite diversity in *Acacia woodmaniorum* under different scenarios of population loss. Genetic diversity parameters are provided for the species after removal of the given population. Parameters include the total number of alleles remaining (At) and the percentage of species allelic diversity lost (At%), the number of private alleles remaining (Ap) and the percentage of private alleles lost (Ap%), and the new values of expected heterozygosity (*He*), and observed heterozygosity (*Ho*). Values in parenthesis are standard errors.

	At	*At% lost*	*Ap*	*Ap% lost*	*He*	*Ho*
Extant Species	199		49		0.631 (0.049)	0.516 (0.049)
Population removed						
Terapod	198	0.05	48	2.04	0.630 (0.049)	0.515 (0.049)
Blue Hills	187	6.03	37	24.49	0.627 (0.049)	0.519 (0.048)
MA1-3	191	4.02	42	14.29	0.629 (0.049)	0.514 (0.049)
MA1-3, MB, MBW	189	5.025	40	18.37	0.631 (0.31)	0.516 (0.49)
Terapod, Blue Hills and MA1-3, MB, MBW	174	12.56	27	44.90	0.625 (0.049)	0.519 (0.048)

Conclusion

Despite its restricted distribution and extreme habitat specificity, *A. woodmaniorum* is not genetically depauperate, as may have been expected for a narrow range endemic, and the species shows high levels of polymorphism at microsatellite loci. The majority of genetic variation is contained within populations and is concentrated in the main populations across Mungada/Windaning ridge, followed by the Jasper Hills region, the Blue Hills and the Terapod populations. This follows expectations of larger populations maintaining greater levels of genetic diversity. A pattern of increasing genetic differentiation with increasing geographic separation is found over the species range and among populations within the Jasper Hills region. Despite this, genetic differentiation among populations and regions is generally low and gene flow via pollen dispersal among populations appears to be high. Populations show little evidence of inbreeding or negative assortative mating and there is no genetic evidence of recent population bottlenecks suggesting habitat destruction, disturbance or anthropogenic fragmentation has not occurred. Populations likely to be impacted by mining operations contain a high degree of genetic diversity and private allele richness and may also be important for the maintenance of gene flow and genetic continuity among populations and regions given their geographic location. This should also be considered in assessing the impact of likely mining scenarios and shall be investigated further in specific studies of pollen flow.

2. Evolutionary relationships and distinctiveness of *Acacia woodmaniorum* to its closest relatives in the *Acacia alata* complex

Introduction

The *Acacia alata* species complex is currently thought to be the closest extant relative to *Acacia woodmaniorum*. The complex is comprised of four subspecies or variants, *A. alata* var *alata*, var *biglandulosa*, var *platyptera* and var *tetrantha*. Phylogenetic analysis is currently being conducted using a range of well known sequences in order to confirm the relationship between the two taxa.

If the two species are confirmed sister taxa then the *A. alata* species complex will be useful for a phylogeographic study comparing it with *A. woodmaniorum*. Such a study may provide insight into the lineages and relationships among the two taxa and the evolutionary history of *A. woodmaniorum*.

Materials and Methods

Sample collection and DNA extraction

Germplasm has been collected from five individuals from each of 20 populations of *A. alata* across the species range, and comprising all subspecies, from north of Geraldton south to Albany (Figure 6 and Figure 7). DNA has been extracted from these individuals using the methods described previously by Millar (2009). DNA from two *A. woodmaniorum*, two *A. alata* var *biglandulosa*, one var *platyptera* and one var *tetrantha* accession have been sent to Dr Gillian Brown of the University of Melbourne and the Royal Botanic Gardens Melbourne who is conducting a phylogenetic review of the Pulchelloidea clade of *Acacias* using a range of sequences including matK, rpl32, ITS and ETS. This clade includes members of the section Alatae to which *A. woodmaniourm* and *A. alata* belong. *Acacia woodmaniorum* DNA has also been provided to Dr Paul Neville of the Botanic Gardens and Parks Authority and Joe Miller of the Commonwealth Scientific and Industrial Research Organisation, Canberra who are also working on Acacia systematics and barcoding using regions including trnk and rbcl.

The above mentioned DNA sequences are suitable for the analysis of intersubspecific variation but generally are found to contain insufficient variation for use in intrasubspecific studies. Work is about to begin testing a number of different sequences that show promise of having high levels of intraspecific variation in *A. woodmaniourm* and *A. alata*.

Results

We are currently awaiting results of the phylogenetic analysis of the Pulchelloidaea clade of Acacia being conducted at the University of Melbourne and the Royal Botanic Gardens Melbourne and of testing for sequence variation in barcoding sequences being conducted at CSIRO, Canberra.

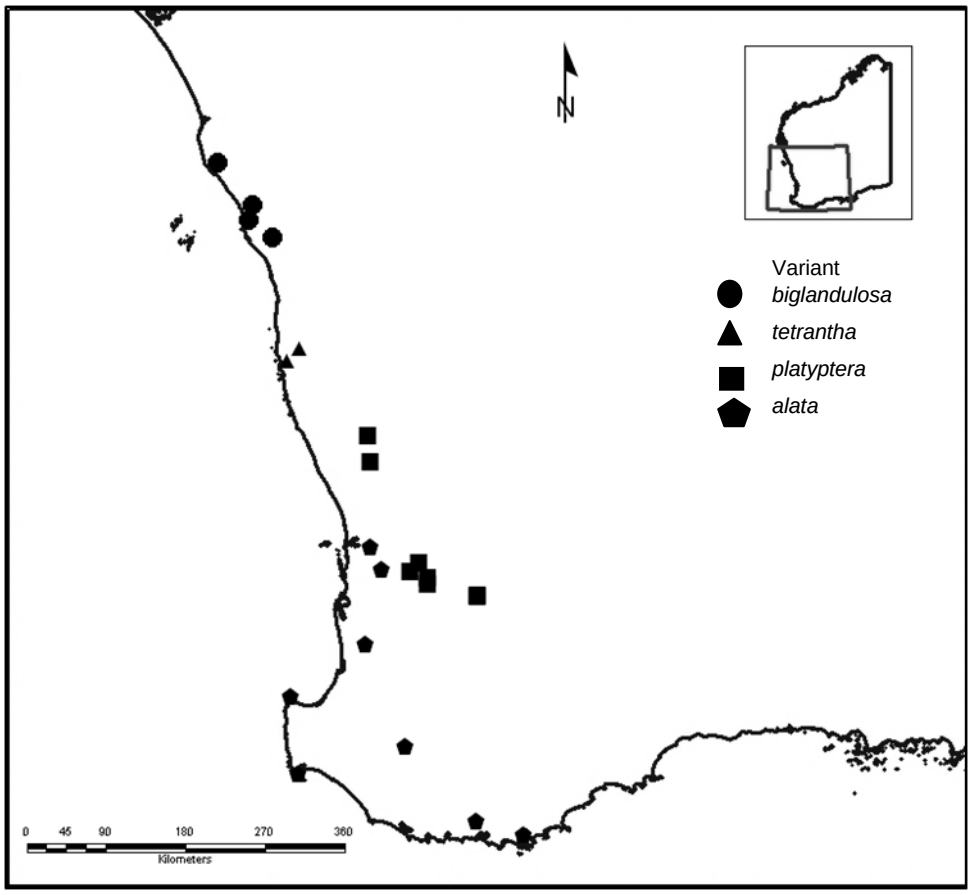


Figure 6 Sampling sites of the *Acacia alata* species complex.

Figure 7. *A. alata* variant *platyptera* growing in *Eucalyptus wandoo* woodland, road verge, Gillingarra.

Phase 2 Determining how key population genetic processes such as mating systems, gene flow and pollination may influence future levels and patterns of genetic variation in *A. woodmaniorum*.

Introduction

Results from Phase 1 of the project suggest that high levels of gene flow are maintained among populations of *A. woodmaniourm*. Levels and patterns of gene flow via pollen are linked to the mating system where they influence levels of mating among individuals and levels of selfing and outcrossing. Gene flow and mating systems in turn are affected by pollinators and their abundance and behavior which are undescribed for *A. woodmaniorum*. Phase 2 of the project aims to describe the mating system of *A. woodmaniorum* in terms of the level of outcrossing or selfing, levels of seed set and degree of correlated paternity within pods as well as investigating levels and patterns of gene flow for populations across the species range. These processes will be investigated in populations of both small and large size and different degrees of isolation, as these factors are expected to impact on the mating system and patterns of gene flow via pollen. Pollinator studies aim to determine the main pollinators of the species providing further insight on the maintenance of genetic continuity among populations and patterns of pollen dispersal within the species.

Phase 2 of the project was not due to commence until March 2010. This phase of the project is well underway however, due to the nature of the work, e.g. seeds were collected from the 2009 flowering season.

Materials and Methods

Seed pods from the 2009 flowering season were collected in December 2009. Pods were collected from one site in the main range and from nine smaller populations. A total of 721 seed have been processed, germinated in growth cabinets and are currently being grown in shadehouse conditions (Figure 9).

A total of 139 seed representing all seeds within each of two pods from each of ten mother plants at the MA2 sampling site within the main range have been sown for a pilot study assessing correlated paternity within pods. The site represents a subpopulation within the large continuous population occurring across Mungada Ridge where the effects of small population processes are expected to be minimised i.e. pollinator abundance should not be reduced and outcrossing rates are expected to be high. In the Acacia genus pollen occurs as composite units or polyads with several grains, or as pollinia with very large numbers of grains. The pollen grain number usually exceeds the ovule number slightly so that a single polyad pollinating a stigma can fertilise all ovules within a flower (Kenrick, Knox, 1982). In this case, all seed within a pod will be full sibs with the same father and correlated paternity within a pod will be high. It is important to assess the degree of correlated paternity in order to design studies of gene flow via pollen that are not biased by a large number of full sibs.

A further 582 seeds representing progeny arrays of up to 22 seed from mother plants in each of nine small populations (Figure 8) have been sown for studies of pollen flow. As the maternal genotypes are already known, genotyping of progeny allows the paternal (pollen) contribution to the progeny to be determined. In small populations, where all trees have been genotyped, the paternal tree can then often be assigned. If paternal trees are not assigned to progeny from within the population, then the paternal tree is assumed to be located outside the stand, and the progeny is a result of the immigration of pollen into the stand. When combined with spatial geographic data, patterns of gene flow and pollen dispersal can then be assessed. Progeny arrays will also be used in the same manner to determine the level of outcrossing or selfing within populations.

When seedlings are sufficiently large enough for the non-destructive harvest of phyllode material DNA will be extracted and individuals genotyped following the methods of (Millar, 2009)

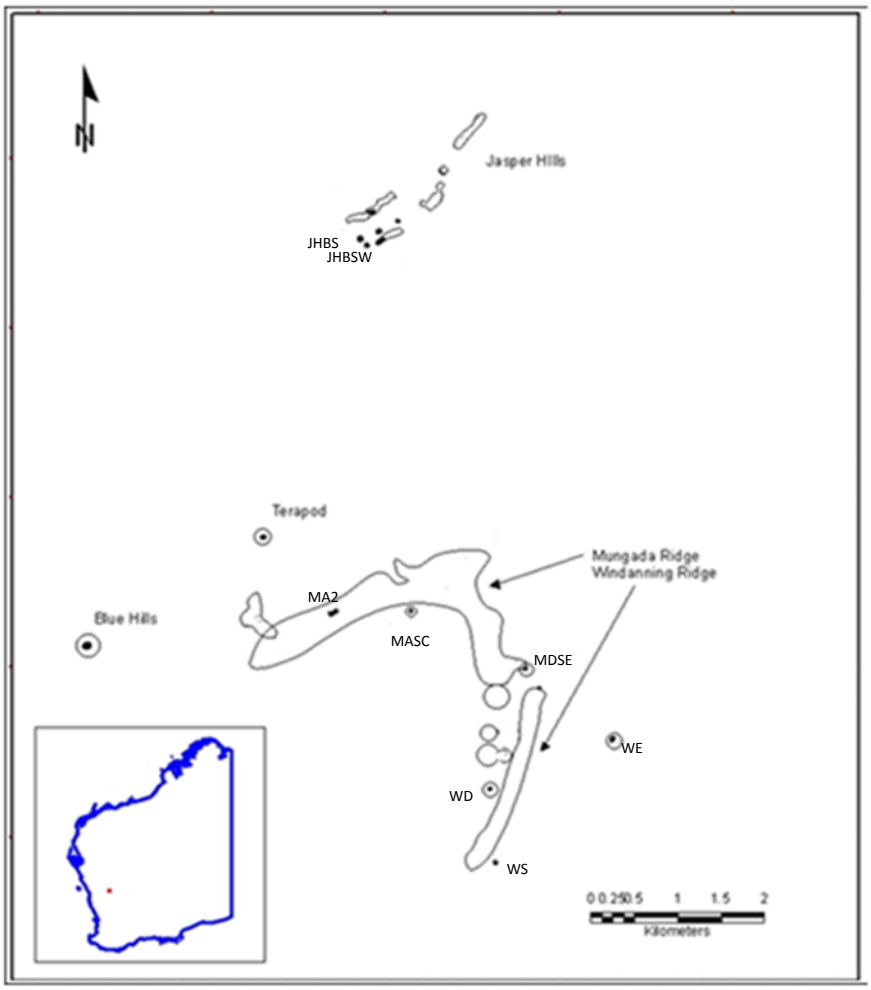


Figure 8. Range of *Acacia woodmaniorum*(open shapes) showing one correlated paternity site within the main range (MA2) and nine small gene flow sites (JHBS, JHBSW, MASC, MDSE, WD, WE, WS, Terapod and Blue Hills) where pods were collected (dark circles). Regions are labeled.

Figure 9. *A. woodmaniorum* progeny being grown for gene flow studies.

References

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frquency data. *Genetics* **144**, 2001-2014.

Earl D (2009) Structure Harvester v0.56.3.

Ellstrand NC, Ellam, D (1993) Population genetic consequences of small population size: implications for plant conservation. *Annual Review of Ecology and Systematics* **24**, 217-242.

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, Smouse P, Quattro J (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**, 479-491.

Falush D, Stephens M, Pritchard J (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164**, 1567-1587.

Kalinowski S (2004) Counting alleles with rarefaction: Private alleles and hierarchical sampling designs. *Conservation Genetics* **5**, 539-543.

Kalinowski S (2005) HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes* **5**, 187-189.

Kenrick J, Knox B (1982) Function of the polyad in reproduction of *Acacia*. *Annals of Botany (London)* **50**, 721-727.

Lewis PO, Zaykin D (2001) Genetic Data Analysis: Computer program for the analysis of allelic data.

Maslin B, Buscomb C (2007) Two new *Acacia* species (Leguminosae: Mimosoideae) from banded ironstone ranges in the Midwest region. *Nuytsia* **17**, 263-272.

Millar MA (2009) Characterisation of microsatellite DNA markers for the rare *Acacia woodmaniorum* (Leguminosae: Mimosaceae). *Conservation Genetics Resources* **1**, 441-445.

Peakall R, Smouse P (2006) GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6**, 288-295.

Piry S, Luikart G, Cournet J (1999) Bottleneck: a computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity* **90**, 502-503.

Pritchard J, Stephens M, Donnely P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.

Raymond M, Rousset F (1995) GENEPOP (version 3.4): population genetics software for exact test and ecumenicism. *Journal of Heredity* **86**, 248-249.

Rice W (1989) Analysing tables of statistical tests. *Evolution* **43**, 223-225.

Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* **145**, 1219-1228.

Slatkin M (1987) Gene flow and the geographic structure of natural populations. *Science* **236**, 787-792.

Slatkin M (1993) Isolation by distance in equilibrium and non-equilibrium populations. *Evolution* **47**, 264-279.

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**, 457-462.

Smouse PE, Long JC (1992) Matrix correlation analysis in anthropology and genetics. *Yearbook of Physical Anthropology* **35**, 187-213.

Smouse PE, Long JC, Sokal RR (1986) Multiple regression and correlation extensions of the Mantel test of matrix correspondence. *Systematic Zoology* **35**, 627-632.

StatSoft (2001) Statistica (Data Analysis Software System).

Weir B, Cockerham C (1984) Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358-1370.

Young A, Boyle T, Brown A (1996) The population genetic consequences of habitat fragmentation for plants. *Trends in Ecology & Evolution* **11**, 413-418.