Chapter 5

Temporal and spatial genetic patterns in the African Wild Silk Moth (*Gonometa postica*) and implications for cyclical population dynamics

"So, if we can observe many genetic variations all of which can be assumed to be the result of the same forces, then the distribution of those variations can be used to estimate those forces."

Richard C. Lewontin (2002)

Temporal and spatial genetic patterns in the African Wild Silk Moth (*Gonometa postica*) and implications for cyclical population dynamics

Wayne Delport, J. Willem H. Ferguson & Paulette Bloomer

[1]Molecular Ecology and Evolution Programme, Department of Genetics, University of Pretoria, Pretoria, 0002, South Africa
[2]Centre for Environmental Studies, Department of Zoology and Entomology, University of Pretoria, Pretoria, 0002, South Africa

Abstract

The African Wild Silk Moth (*Gonometa postica*) exhibits large inter-annual population size changes. Previous studies indicated that the inference of dispersal ability in this species was complicated by the effect of population size fluctuations. In this manuscript we genotype samples collected in three successive years for six microsatellite loci. These data are used to estimate changes in effective population size and migration. Theoretically, migration could be tracked by changes in frequencies and occurrences of rare alleles when temporal samples are analysed. However, the results presented here show little changes in the spatial and temporal distribution and frequency of alleles. We suggest these results are due to the decrease in eruptive populations observed in the species over the temporal period sampled. Further temporal sampling, that includes population eruptions, would allow the inference of dispersal in this species.

Keywords: temporal samples, population cycles, effective population size, continuously distributed

Introduction

The African Wild Silk Moth, *Gonometa postica*, is a species than undergoes large inter-annual population size fluctuations (Veldtman 2004, Delport 2005 Chapter 4). Using simulations, Delport (2005 Chapter 4) showed that population size fluctuations in a continuously distributed species have the potential to generate increased population genetic structure, where demes experience local genetic drift and thus become fixed for alternate alleles (Wright 1940). This potential for population size fluctuations to generate population structure is only evident, however, when dispersal is low. In fact, a dispersal distance of more than 1% of the distribution of a species results in no significant effect of population size fluctuations on spatial genetic pattern versus that of a population with constant size (Delport 2005 Chapter 4). This result is contradictory to the high levels of gene flow inferred from several species with cyclical dynamics (Ehrich *et al.* 2001, Burton *et al.* 2002, Delport 2005 Chapter 4). Burton *et al.* (2002) attribute inferred high levels of gene flow in snowshoe hare (*Lepus americanus*) to a stepping-stone model of gene flow influenced by density cycles, where local bottleneck populations expand to previously unsuitable habitat and thus homogenize genetic diversity across the distribution. The high level of gene flow in collared lemming is potentially explained by long-distance dispersal events (Ehrich *et al.* 2001). However, Delport (2005 Chapter 4) found little indirect evidence in support of long-distance dispersal in a species of moth that survives only a few days, has low flying efficiency and lacks feeding mouthparts. Since population structure may be homogenised by population size fluctuations we consider an alternate method, using temporal genetic sampling, to identify the level of dispersal in this moth species.

Temporal genetic sampling has been shown to be useful in the estimation of effective population size (Krimbas & Tsakas 1971, Nei & Tajima 1981, Pollak 1983, Waples 1989), a measure of particular interest to conservation biologists. In a Wright-Fisher model of population size N, the theoretical variance in allele frequencies can be calculated given a starting allele frequency, and thus subsequently compared to the observed variance in allele frequencies, to estimate the effective population size N_e (Ewens 1979). These moment-based estimators, however, lack precision and typically

include confidence intervals that include infinity (Luikart *et al.* 1999, Berthier *et al.* 2002). Recently moment-based estimators have been replaced by maximum likelihood estimators (Williamson & Slatkin 1999, Anderson *et al.* 2000, Wang 2001, Berthier *et al.* 2002, Beaumont 2003, Wang & Whitlock 2003) that have the advantage of utilizing the full distribution of allele frequencies, and not simply summary statistics as in previous methods. In addition, the likelihood-based methods have allowed for changes to the underlying statistical model, such that joint estimation of N_e and migration in a stable population model (Wang & Whitlock 2003), or of N_e and population size changes (Beaumont 2003) can be achieved. Although temporal analysis of genetic data has previously been limited to rapidly mutating viruses (Drummond *et al.* 2002, Fu 2001) and bacteria (Falush *et al.* 2001), or to samples spaced over relatively long periods (Williamson & Slatkin 1999, Wang & Whitlock 2003, Jehle *et al.* 2001, Beaumont 2003), some studies have analysed temporal samples collected in successive years (Begon *et al.* 1980, Wang 2001, Beaumont 2003). These short-term studies, however, are restricted to insects with short generation times, such that 1-2 generations have elapsed between temporal samples. Insects, with short generation times, are particularly suitable for temporal genetic sampling since multiple generations can pass within the duration of a typical study. In addition, several insect species are characterized by large population size fluctuations (see Vandewoestijne *et al.* 1999, Ibrahim 2001, Bjornstad *et al.* 2002, Cooper *et al.* 2003, Turchin 2003, Delport 2005 Chapter 4 for examples), which have the potential to cause substantial changes in allele frequencies over short temporal periods.

Temporal allele frequency changes can provide information on migration patterns if a deme containing a rare allele experiences an eruption, and the rare allele is found in an alternate deme in the subsequent generation. Estimating the migration rate between these temporally spaced demes may allow the inference of dispersal. We use temporal genetic sampling combined with likelihood methods to estimate migration both within and between sampling periods, and changes in effective population size, N_e, over the sampling period. The principal aim of this manuscript is to infer dispersal patterns of the African Wild Silk Moth using temporally spaced genetic samples. Our results indicate little change in the spatial distribution of allele frequencies over the time period considered. We believe these results are due to the observed decrease in *G*.

postica eruptions (Delport 2005 Chapter 2), and we further suggest that long-term genetic monitoring of these populations will assist in identifying levels of dispersal in this species.

Materials and methods

Genetic sampling and laboratory methods

We collected pupae (cocoons) from *Acacia erioloba* and *Acacia mellifera* trees during the species' diapause in June 2002, July 2003 and July 2004 across the core outbreak region for the species (Veldtman *et al.* 2002, Delport 2005 Chapter 2). Adult moths were allowed to emerge in the laboratory and were frozen at -20°C for subsequent DNA extraction. We sampled a total of 206, 53 and 52 individuals in 2002, 2003 and 2004 respectively. Since apparent extinction of several local populations occurred in 2003 and 2004 (Delport 2005 Chapter 2) it was not possible to achieve the same sample sizes, and number of sampling localities (Figure 1) as in the first year. Total genomic DNA was extracted using a Qiagen Dneasy extraction kit and all individuals were genotyped for six polymorphic microsatellite loci (Gon6, Gon60, Gon65, Gon55, Gon107, Gon120). Details of protocols for microsatellite development, amplification conditions and summary statistics for loci are presented in Delport (2005 Chapter 3).

Statistical analysis methods

We conducted preliminary analysis of temporal allele frequencies by plotting distributions of allele frequencies derived from the three years of sampling. These distribution plots allowed for a visual inspection of changes in allele frequencies over the sampling period. Samples were grouped into localities based on geographic proximity (Figure 1 inset) and on the observed persistence of eruptions over the three-year sampling period (Delport 2005 Chapter 2). The practice of grouping sampling sites into demes may be problematic since the underlying distribution of the species appears continuous with spatial variation in local densities (pers obs). Grouping of localities can generate spurious results, such as the Wahlund effect (Wahlund 1928), that results in a deficiency of observed heterozygosity within populations. This effect is, however, only applicable when populations which have an underlying population

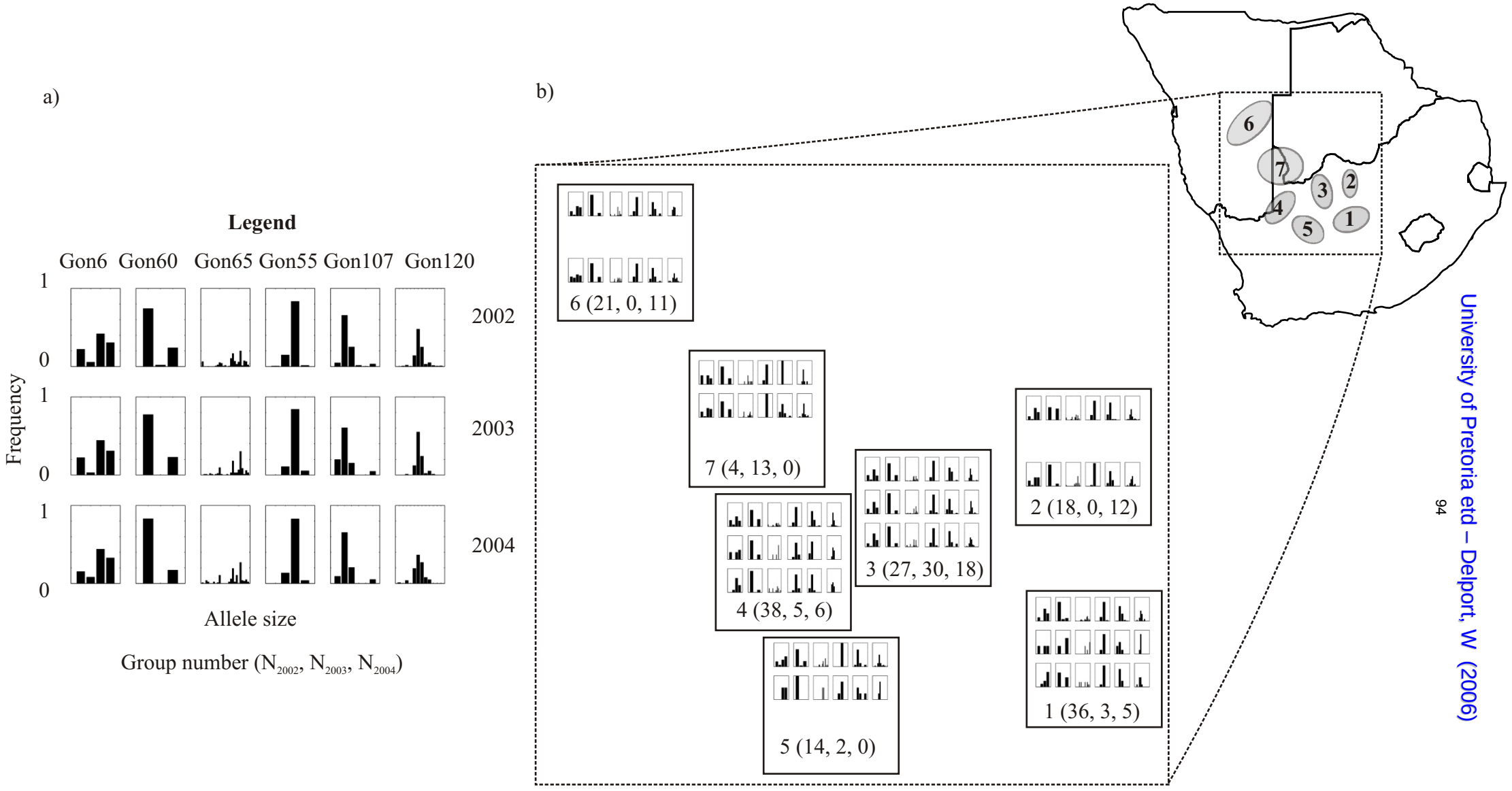


Figure 1: Temporal and spatial distribution of microsatellite allele frequencies. The allele frequency distribution a) for all localities, and b) per locality is represented. Figure 1a also serves as a legend for allele frequency distributions in 1b. Alleles are ordered in size from smallest to largest on the x-axis, with frequency on the y-axis. Locality groupings in subsequent analysis are indicated as gray polygons. Note locality groups 4 and 7 were combined for all subsequent analysis. N_x = sample size in year x. Sample sizes represented here are not consistent with those reported in the methods since several isolated individual samples could not be combined into the locality groups shown in this figure.

genetic structure are combined. Since there is little population genetic structure within *G. postica* (Delport 2005 Chapter 4) we consider the lumping of sampling sites into demes to be acceptable. Clearly, the analysis of continuously distributed data under an island model is not ideal, yet given that there are no alternative genealogical methods for the estimates of migration, we followed this approach. Some support for using an island model to infer gene flow in an apparently continuously distributed species is provided by temporal sampling. *G. postica* does not have overlapping generations over years, and thus temporal sampling of populations is analogous to an island model, where demes are temporally structured. The subsequent analysis comprised three approaches.

Firstly, we used a chi-square permutation procedure (Roff & Bentzen 1989) to determine whether the spatial distribution of allele frequencies at localities in each of the three years was significantly different from that expected at random. The method is justified for n x n contingency tables with low frequencies in cells, and for which there are no exact solutions available (Roff & Bentzen 1989).

Secondly, we used migrate-n (Beerli & Felsenstein 1999, 2001) to estimate migration rates between grouped sampling localities from 2002. Localities were grouped based on geographical proximity as before, yet localities 4 and 7 were combined since the sample size of locality 7 was insufficient for this analysis. Only localities for which there were > 5 samples in any year (mean = 20, sd = 11.45) were included in the analysis. The purpose of a subsequent temporal analysis was to identify whether single or multiple populations from 2002 could have sourced eruptions in 2003, and 2004 respectively. We used migrate-n (Beerli & Felsenstein 1999, 2001) to estimate migration rates between populations in successive sampling years. Only localities with sufficiently large sample sizes in 2003 and 2004 were used (2003: locality 3; 2004: locality 2 & 3). In the temporal migrate-n analysis we used a restricted custom migration model (Table 1) that only estimated one-way migration between populations in successive years. This migration model reduced the number of parameters to be estimated from the data thus ensuring the data were not over-analyzed. The use of migrate-n to identify temporal migration rates is questionable since the software does not model recruitment, however, this analysis would still

Table 1: Custom migration model enforced in migrate-n such that only migration between populations in successive years is estimated. This custom migration model was used such that the number of parameters estimated from the data was reduced from 64 to 25. Estimates of population size (θ) are shown along the diagonal and migration estimates above and below the diagonal. Migration from column to row is indicated in each cell where; * = no restriction, m = average migration across all demes (as per Beerli 2004), Y = year sampled, d = deme as in Figure 1.

		Migration from:							
	Y		**2002**	**2002**	**2002**	**2002**	**2002**	**2003**	**2004**	**2004**
		d	**1**	**2**	**3**	**4&7**	**5**	**3**	**2**	**3**
Migration to:	**2002**	**1**	*	m	m	m	m	m	m	m
	2002	**2**	m	*	m	m	m	m	m	m
	2002	**3**	m	m	*	m	m	m	m	m
	2002	**4&7**	m	m	m	*	m	m	m	m
	2002	**5**	m	m	m	m	*	m	m	m
	2003	**3**	*	*	*	*	*	*	m	m
	2004	**2**	*	*	*	*	*	*	*	m
	2004	**3**	*	*	*	*	*	*	m	*

provide some information of the degree of patch consistency over time (pers com. Peter Beerli). All migrate-n analyses were run with maximum-likelihood as the search strategy with ten short chains (50000 genealogies sampled, 1000 recorded) and three long chains (500000 genealogies sampled, 10000 recorded), and a burn-in of 10000 genealogies per chain. Furthermore, in order to check for consistency, migrate runs were repeated with starting parameters set at the maximum-likelihood estimates of the previous run.

Finally, since we observed changes in abundance between the three sampling years we used the temporal data to estimate changes in effective population size (Beaumont 2003). We used the software tmvp (Beaumont 2003) to obtain a maximum likelihood estimate of change in effective population size given the temporal microsatellite data. Preliminary runs of tmvp were conducted to determine the appropriate scale for the estimation of effective population size. We furthermore divided the analysis into two separate searches. Firstly, we used all the data from 2002 (206 individuals), 2003 (53 individuals) and 2004 (52 individuals), and estimated the likelihood surface of change in effective population size across six generations, given that *G. postica* cycles through two generations per annum (Hartland-Rowe 1992). Secondly, since the sample size from the first year (2002) was substantially larger than in subsequent years, we randomly sampled 50 individuals from the 2002 data. These data were combined with that from subsequent years to account for the effects of sample size in estimating change in effective population size. For both tmvp analyses the following importance sampling search options were used. Forty-thousand updates were proposed at a parameter step size of 0.15, and a thinning interval of 10; thus producing 4000 calculations of the likelihood given the observed data. Maxit, the size of the importance sample was set to 500. Preliminary analysis indicated that the likelihood surface for the first analysis should be estimated at a scale of 1000 individuals on each axis, whereas for the second, a scale of 5000 individuals per axis was most appropriate.

Results

Spatial and temporal distribution of allele frequencies

No clear changes in allele frequency are evident on either a spatial or temporal time scale (Figure 1). In general within a single sampling period, localities have one to two high-frequency alleles that are distributed throughout the range of the species (Figure 1). The observed spatial distributions of alleles for each locus were not significantly different from random as determined with permutations of the chi-square test (Table 2). These results provide support for the lumping of localities within years. This allelic distribution pattern is evident in all three sampling periods, where the high frequency alleles are maintained over time (Figure 1b). However, overall allelic diversity is reduced over time, as is evident in the loss of rare alleles (Figure 1a), particularly in Gon60, Gon65, Gon107 and Gon120. This loss of allelic diversity is attributable to one of two factors; the fewer number of eruptions found subsequent to 2002 and thus indicative of a population crash (Delport 2005 Chapter 2), or directly related to the former; the reduced sample sizes of 2003 and 2004.

Maximum likelihood estimation of demographic parameters

Maximum likelihood estimates (MLE's) of migration indicate the relative contribution of immigration versus mutation in generating new alleles within demes to be generally large (Table 3, Figure 2). Maximum likelihood estimates, however, should not be evaluated without consideration of the confidence limits surrounding such estimates. Confidence limits on the MLE taken at 2 log likelihood units below the maximum likelihood estimate (Edwards 1972) indicated little variation around the MLE and thus greater confidence in this estimate (Table 2). Migration estimates from a second search with the starting parameters set to the MLE of the previous run were in agreement with results from the first search. Very large migration estimates occur between locality 6, and each of localities 4&7 and 3. Lower estimates of migration however occur between adjacent localities 1 & 5, 3 & 5, 4&7 & 3, and 4&7 & 5. These results are generally in contradiction to what one might expect from an isolation by distance model, where an increase in geographic distance is accompanied by an increase in genetic distance (Wright 1943). The estimates of migration provided are given in terms of the contribution that immigration makes to the occurrence of

Table2: The spatial distribution of alleles for each of the six microsatellite loci. The results presented are from 1000 permutations of the chi-square test (Roff & Bentzen 1989) and represent the probability that the observed distribution of alleles among localities is significantly different from random. The Bonferroni corrected rejection level for multiple comparisons is 0.002. χ^2 = observed chi-square, *P* = proportion of random permutations with a χ^2 greater than that observed.

Locus	**Year**	χ^2	***P***
Gon6	2002	6.95	0.980
	2003	3.16	0.990
	2004	8.08	0.819
Gon60	2002	17.09	0.147
	2003	6.16	0.157
	2004	2.48	0.700
Gon65	2002	79.22	0.963
	2003	75.67	0.012
	2004	55.55	0.708
Gon55	2002	10.27	0.915
	2003	7.01	0.521
	2004	6.57	0.622
Gon107	2002	27.72	0.286
	2003	8.03	0.845
	2004	7.77	0.924
Gon120	2002	33.33	0.829
	2003	8.87	0.899
	2004	6.54	0.991

new alleles within a subpopulation/deme. However, it is useful to consider the effective number of migrants, *4Nm*. In this case simply multiplying the estimates of *M* by θ provides an estimate of the effective number of migrants (Beerli 2004) and ranges from less than 1 to approximately 12 (Table 3).

The 2002 migration estimates were inconsistent with the geographical location of samples. In general, higher levels of migration were observed between distant localities than between adjacent localities (Figure 2, Table 3). Therefore, we were unable to combine adjacent localities for the temporal analysis of migration. Rather we identified locality groups in the subsequent years in which there were sufficient sample sizes, and tested for temporal migration between these localities and adjacent localities from previous years. Therefore, the localities 1-5 were included from 2002, locality 3 in 2003 and localities 2 and 3 from 2004. In this way the number of parameters to be estimated was reduced, and localities with small sample sizes removed from the analysis. Smaller sample sizes for localities in successive years are indicative of small local population sizes, and not insufficient sampling, since search effort was consistent across localities and years. The temporal analysis of migration, in general, indicates high levels of gene flow (Figure 3, Table 4). Considering only the samples collected in 2002 and 2003, it appears that two of the five demes (4&7, 1) contributed most to the only deme sampled in 2003. Similarly, considering the results from 2002 and 2004, three of the five demes from 2002 appear to have contributed most to the eruption in 2004. In general, the temporal migration results indicate substantial gene flow across the range of the species. Confidence limits on these estimates of temporal migration were again small (Table 4) and a second run of migrate-n with the MLE's as start parameters resulted in similar estimates.

The analysis of change in effective population size using all the data collected in 2002 indicated a substantial change as inferred from the MLE of ancestral and recent effective population sizes (Figure 4a). The MLE (L = -323.4971) occurs at ancestral (AN_e) and recent (RN_e) effective population sizes of approximately 863 and 169 individuals, respectively (Figure 4a). The 95% confidence limit for the MLE shown is at –325.5. Evaluation of this confidence limit indicates that there is little confidence in the MLE, with AN_e and RN_e ranging from 300-1000 and 100-1000 respectively. When

Table 3: Maximum likelihood estimates of migration between 2002 demes (Figure 1) and θ ($4N\mu$, N = population size, μ = mutation rate) for each deme, as estimated with migrate-n (Beerli & Felsenstein 1999, 2001). Both M, the relative contribution of immigration versus mutation in generating new alleles ($M = m/\mu$, where m = immigration rate, μ = mutation rate) and $4Nm$, the effective number of migrants (N = population size, m = migration rate), are presented. *MLE*'s (maximum likelihood estimates) of M are shown for each deme/subpopulation pair, with $4Nm$ in parenthesis. MLE's of θ are boxed and shown along the diagonal. Table is read as immigration from column to row. Migration rates shown on Figure 2 are in bold. U = Upper 95% percentile of the estimated likelihood surface, L = lower 95% percentile of the estimated likelihood surface.

Migration from deme:

Migration to deme:		**1**	**2**	**3**	**4&7**	**5**	**6**
1	*MLE*	*0.47*	8.61 (4.1)	5.42 (2.6)	7.15 (3.4)	9.78 (4.7)	**18.72 (8.9)**
	L	*0.44*	7.57	4.61	6.20	8.67	**17.15**
	U	*0.51*	9.75	6.33	8.19	10.99	**20.37**
2	*MLE*	7.03 (3.1)	*0.45*	6.80 (3.0)	9.36 (4.2)	5.39 (2.4)	**14.5 (6.5)**
	L	6.14	*0.40*	5.92	8.33	4.61	**13.22**
	U	8.00	*0.49*	7.76	10.48	6.25	**15.90**
3	*MLE*	6.26 (2.3)	9.41 (3.5)	*0.37*	**19.1 (7.1)**	7.70 (2.9)	**31.9 (11.9)**
	L	5.35	*8.27*	*0.34*	**17.48**	6.67	**29.74**
	U	7.28	10.64	*0.41*	**20.85**	8.82	**34.10**
4&7	*MLE*	9.27 (4.3)	8.77 (4.1)	**18.1 (8.4)**	*0.47*	6.55 (3.1)	**16.49 (7.7)**
	L	8.25	7.77	**16.59**	*0.43*	5.69	**15.10**
	U	10.39	9.85	**19.58**	*0.50*	7.49	**17.96**
5	*MLE*	**11.1 (2.9)**	9.6 (2.5)	**11.0 (2.9)**	**14.0 (3.7)**	*0.26*	**11.0 (2.9)**
	L	**9.74**	8.29	**9.58**	**12.39**	*0.23*	**9.57**
	U	**12.66**	10.99	**12.49**	**15.67**	*0.30*	**12.46**
6	*MLE*	**20.8 (7.5)**	**17.6 (6.3)**	**34.6 (13)**	**25.3 (9.2)**	**12.4 (4.5)**	*0.36*
	L	**18.96**	**15.95**	**32.24**	**23.25**	**11.01**	*0.32*
	U	**22.66**	**19.35**	**37.02**	**27.32**	**13.86**	*0.41*

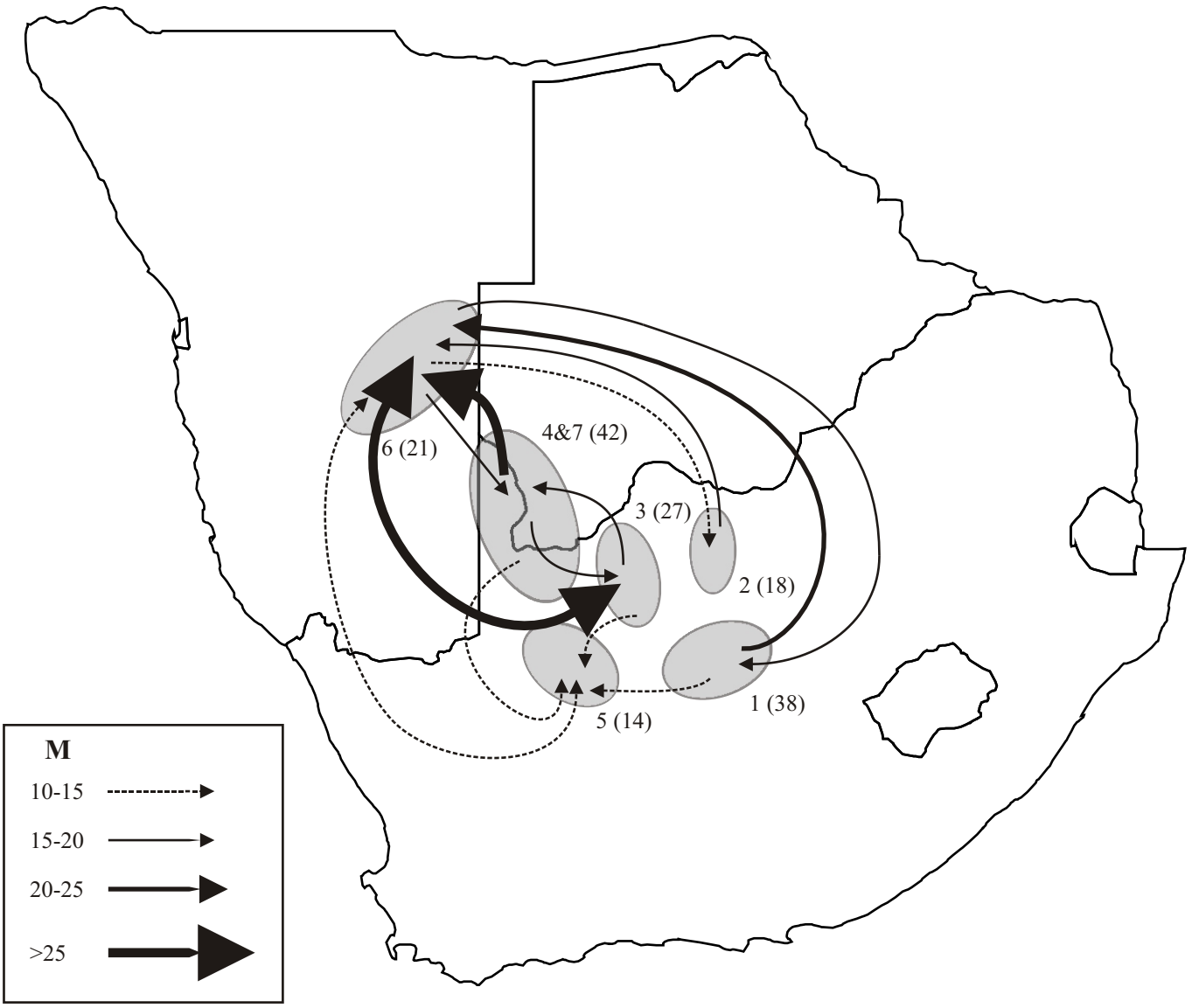


Figure 2: Maximum likelihood estimates of migration (from 2002 samples) using migrate-n (Beerli & Felsenstein 1999, 2001). The relative importance of migration versus mutation in generating new alleles in demes is presented as M, where $M = m/\mu$ (m = immigration rate, μ = mutation rate). Only M values > 10 are presented, whereas remaining immigration levels are shown in Table 2.

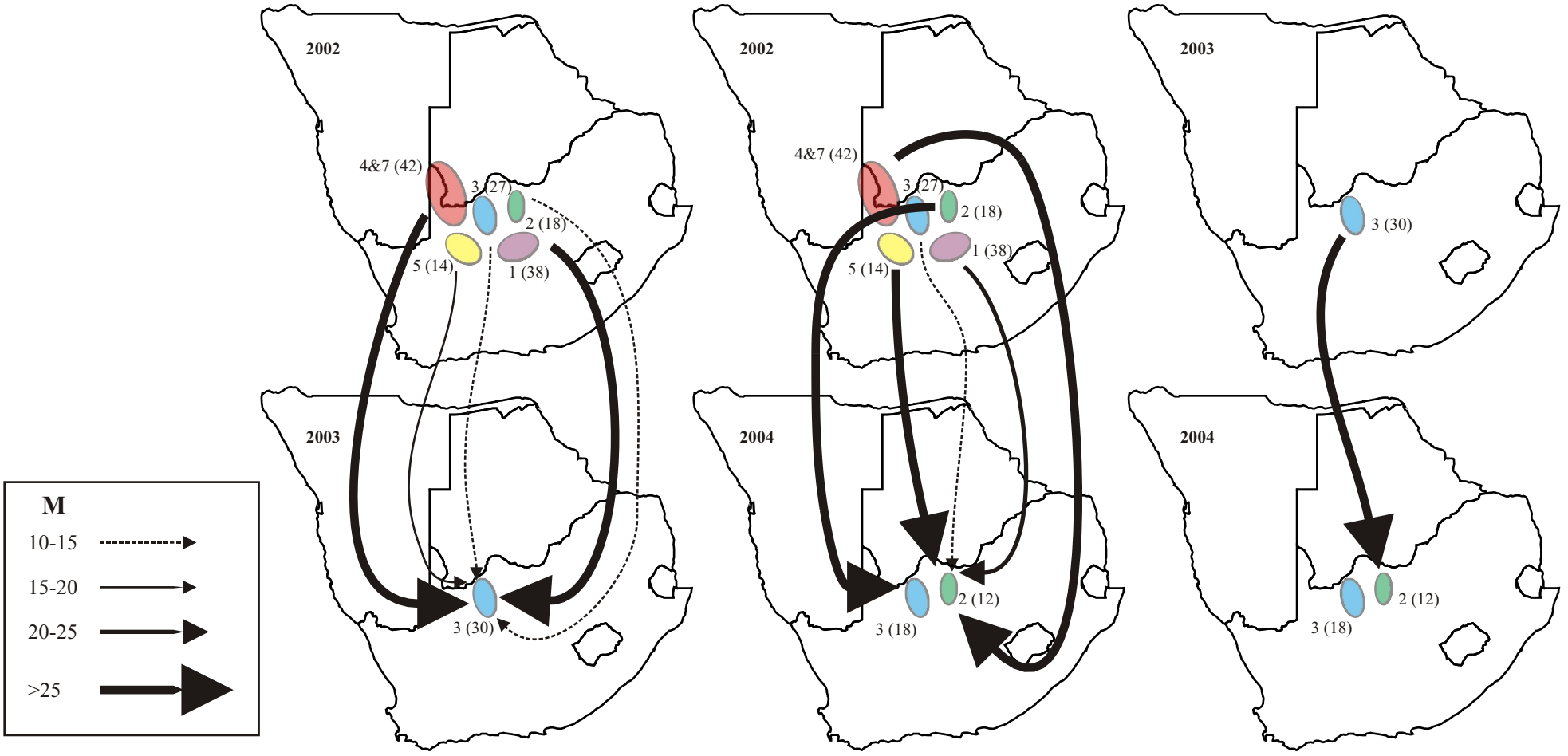
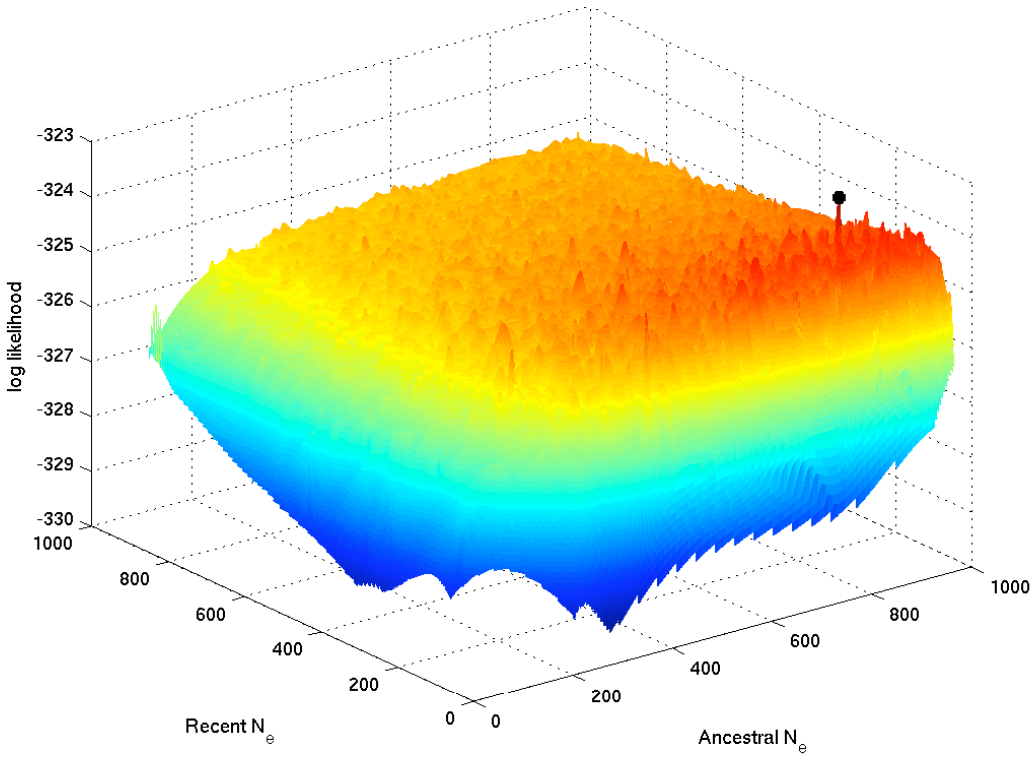


Figure 3: Temporal estimates of migration, $M = m/\mu$ (m = immigration rate, μ = mutation rate), between demes sampled in successive years. Only values greater than 10 are presented, whereas remaining immigration levels are shown in Table4.

Table 4: Maximum likelihood estimates of temporal migration and θ (boxed), as estimated with migrate-n (Beerli & Felsenstein 1999, 2001). Only migration between demes (d) collected in successive years (Y) was estimated so as to reduce the number of parameters estimated. Both *M*, the relative contribution of immigration versus mutation in generating new alleles ($M = m/\mu$, where m = immigration rate, μ = mutation rate) and *4Nm*, the effective number of migrants (N = population size, m = migration rate), are presented. *MLE*'s (maximum likelihood estimates) of *M* are shown for each deme/subpopulation pair, as is *4Nm*. MLE's of θ are boxed. Table is read as immigration from column to row. Migration rates shown on Figure 4 are in bold. *U* = Upper 95% percentile of the estimated likelihood surface, *L* = lower 95% percentile of the estimated likelihood surface, *N* = not estimated.

			Migration from deme:							
	Y	**d**		**2002 1**	**2002 2**	**2002 3**	**2002 4&7**	**2002 5**	**2003 3**	**2004 2**	**2004 3**
Migration to deme:	**2003**	**3**	*MLE*	**41.34**	**10.60**	**10.79**	**35.05**	**19.43**	*0.16*	N	N
			L	38.15	9.03	9.21	32.13	17.27	*0.14*	N	N
			U	42.71	12.34	12.55	38.16	21.77	*0.17*	N	N
			4Nm	10.3	2.6	2.7	8.8	4.9		N	N
	2004	**2**	*MLE*	**22.40**	5.42	**12.93**	**26.31**	**29.27**	**43.22**	*0.09*	N
			L	19.88	4.23	11.05	23.57	26.33	39.69	*0.08*	N
			U	25.19	6.81	15.02	29.24	32.38	46.96	*0.10*	N
			4Nm	2.1	0.5	1.2	2.4	2.6	3.9		N
	2004	**3**	*MLE*	3.23	**28.04**	5.26	5.28	4.53	6.46	N	*0.41*
			L	2.58	25.98	4.41	4.43	3.75	5.51	N	*0.38*
			U	3.98	30.30	6.22	6.24	5.42	7.51	N	*0.47*
			4Nm	0.6	5.6	1.1	1.1	0.9	1.3	N	



b)

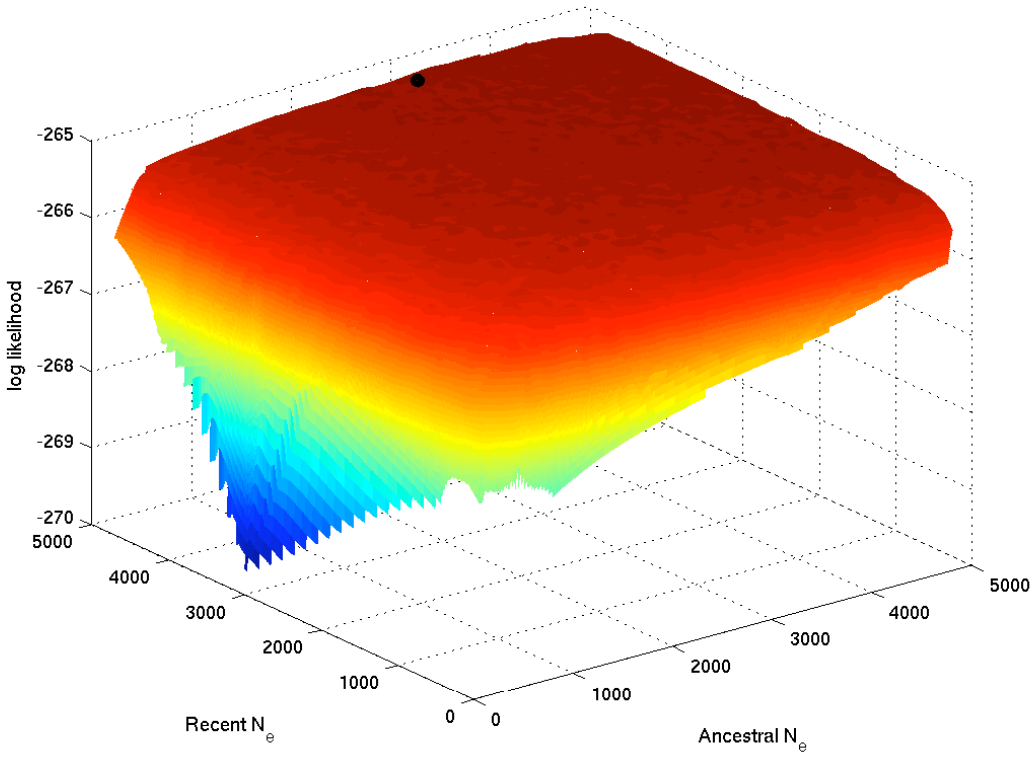


Figure 3: Likelihood surfaces of change in effective population size over the sampling period. The likelihood surfaces were approximated using either a) all the data obtained from 2002-2004, or b) a randomly re-sampled dataset from 2002 of fifty individuals, such that sample sizes in the three sampling periods were equivalent. The maximum likelihood estimate of ancestral and recent Ne is shown as a black point.

we accounted for sample size differences between the sampling years, in particular 2002 versus that of later years, the analysis of change in N_e yielded different results. The likelihood surface of AN_e and RN_e was essentially flat with a MLE of approximately 3115 and 4793 individuals, for AN_e and RN_e, respectively. We furthermore, conducted the analysis on a broader scale of effective population size (10000 individuals on each axis), and the results were essentially the same, a flat surface with no apparent MLE (results not shown). Again 95% confidence limits indicate that randomly sampling the dataset for equivalent sample sizes removes the low signal contained within the data, as evident in the full-data set.

Discussion

Species that exhibit large population fluctuations have been studied intensively from an ecological and population dynamics perspective (see Turchin 2003 for review). However, the analysis of spatial genetic patterns in these species is seldom addressed, with few exceptions (Fuller *et al.* 1997, Vandewoestine *et al.* 1999, Ibrahim *et al.* 2000, Ibrahim 2001, Burton *et al.* 2002, Ehrich *et al.* 2002). The field of phylogeography, which has a principal role of inferring demographic processes from spatial genetic data, has made considerable contributions to understanding demographic processes of focal species (Avise 2000, Hewitt 2001, Knowles & Maddison 2002). Thus, the examination of spatial genetic patterns in species that exhibit complex population cycles should allow for a better understanding of population connectivity and migration rates. However, an examination of studies of spatial genetic pattern in species with complex population cycles indicates that such inferences might be problematic. Spatial genetic analysis of several cyclical species present high levels of inferred migration (Fuller *et al.* 1997, Vandewoestine *et al.* 1999, Burton *et al.* 2002, Ehrich *et al.* 2002). These results are, however, in contradiction to theory where local population size fluctuations are expected to produce greater spatial genetic structure as a result of local genetic drift within local populations (Wright 1940). The effects of extinction and recolonisation on spatial genetic pattern have been evaluated for species with metapopulation structure (Wade & McCauley 1988, Whitlock & McCauley 1990, Ibrahim *et al.* 2000, Ibrahim 2001). The results from these studies in general indicate that the effect of population turnover on genetic differentiation is dependent on the number of individuals colonizing a

deme relative to the number of recurrent migrants between demes (Whitlock & McCauley 1990, Ibrahim 2001). Fewer colonizers produce greater genetic structure due to founder effects, as do fewer migrants. Whether colonizers are sourced from multiple or single demes would also determine the effects of population cycles. Thus, the observed high levels of gene flow in the examples above, and in *Gonometa postica* considered here, might be the result of multiple colonizers, from multiple demes, combined with many recurrent migrants between demes.

The results presented in this manuscript are in agreement with a previous study (Delport 2005 Chapter 4) that inferred high levels of gene flow between sampled localities of *G. postica*. However, given that *G. postica* only survives for 2-9 days (Hartland-Rowe 1992), has no feeding mouthparts (Hartland-Rowe 1992), and may be regarded as an inefficient flier (Delport 2005 Chapter 4, Veldtman 2005), these results are unexpected. The indirect estimates of dispersal (Delport Chapter 4), based on wing load, the ratio of wing area to body mass, could, however, be problematic. High wing loads are characteristic of poor fliers and low wing loads of good fliers (Casey & Joos 1983). However, some species, such as small-bodied geometrid moths, have very low wing loads but are known to be very poor fliers (Rydell *et al.* 1997). Furthermore, some non-feeding saturniid moths, a family closely related to lasiocampids, have been shown to be good dispersers (Waldbauer & Sternberg 1982), despite large body sizes and high wing loads. Delport (2005 Chapter 4) used simulations to determine the effects of population size fluctuations on spatial genetic structure in continuously distributed species. These simulations supported the notion of increased population structure under low levels of gene flow (Wright 1940), yet it could not be shown how population cycles might result in increased spatial genetic homogeneity and thus gross overestimates of gene flow. Thus it still remains to be explicably demonstrated whether the observed spatial genetic patterns in *G. postica* are the result of high dispersal ability or the effects of large inter-annual population size fluctuations.

In this manuscript we determined the levels of gene flow between *G. postica* demes using temporally spaced population genetic data. The results, in general, indicate very high levels of inferred migration between all localities. An anomaly is evident in the greater migration rates between distant, than between adjacent localities (Figure 2). In

contrast, temporal sampling suggested that adjacent localities are likely to seed eruptions in successive years (Figure 3). These migration estimates had narrow confidence limits, however the absence of spatial and temporal patterns in allele frequencies (Figure 1, Table 2) draws question as to which data are contributing to the inference of migrations. Most likely the distribution of rare alleles is contributing to the inference of migration, since the loss of some rare alleles are observed over the sampling period. These reductions in rare allele frequencies can be attributed to either reductions in population size, or insufficient sampling in 2003 and 2004. Delport (2005 Chapter 2) did observe a large decrease in the occurrence of *G. postica* eruptions between 2002 and the subsequent sampling years. Given, however, that the sample sizes in subsequent years were lower it is difficult to determine whether these alleles were actually lost. Similarly, an analysis of the change in effective population size indicated a large reduction in the MLE estimate of effective population size in 2002 versus 2004. Yet, when we corrected for sampling size differences, this result was not apparent. Temporal sampling has been shown to be useful in estimating the change in effective population size of the Mauritius kestrel (Beaumont 2003). However, the distribution of allele frequencies, and whether the population has experienced an eruption or population crash, is likely to affect the success of the method. Given a species with a large skew in the distribution of alleles, and several rare alleles, as in this study, the ability to discriminate population size reduction from insufficient sampling is low, since both would result in the failure to detect rare alleles. However, given several rare alleles and a subsequent population eruption, one would be able to detect whether rare alleles were truly lost or simply not sampled during non-eruptive years. Furthermore, an eruption would allow for rare alleles to become more frequent, and thus allow the tracking of dispersal to neighbouring or distant demes. We believe the inability for temporal genetic data to provide estimates of population size changes and dispersal ability in this manuscript is the result of the decline in eruptions over the sampling period of this study (Delport 2005 Chapter 2). Furthermore, since the demographic estimates in this manuscript are most likely dependent on the frequencies of rare alleles it would be advantageous to increase the number of loci analyzed. This would increase inference power and potentially allow the joint estimation of migration of population size changes using temporally spaced samples.

Acknowledgements

This work was funded by the Mellon Foundation Grant to J. W. H. Ferguson, P. Bloomer and W. Delport, and by a National Research Foundation grant to P. Bloomer (Gun: 2053653). The opinions and views presented in this article are however, not necessarily those of the National Research Foundation. Furthermore, we would like to thank Louis Hauman, Duncan McFadyen and E. O. Oppenheimer & Son for accommodation assistance during sampling. Thanks also to Fourie Joubert and deepthought (http://deepthought.bi.up.ac.za), a 64-node cluster, for computational support.

References

Anderson EC, Williamson EG, Thompson EA (2000) Monte Carlo evaluation of the likelihood for N_e from temporally spaced samples. *Genetics*, **156**, 2109-2118.

Avise JC (2000) Phylogeography: The history and formation of species. Harvard University Press, Harvard.

Beaumont MA (2003) Estimation of population growth or decline in genetically monitored populations. *Genetics*, **164**, 1139-1160.

Beerli P & Felsenstein J. (1999) Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. *Genetics*, **152**, 763-773.

Beerli P & Felsenstein J. (2001) Maximum-likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proceedings of the National Academy of Sciences, USA*, **98**, 4563-4568.

Beerli P (2004) *Migrate Documentation* Version 2.0. School of Computational Science and Department of Biological Sciences, Florida State University, Tallahassee. Florida.

Begon M, Krimbas CB, Loukas M (1980) The genetics of *Drosophila subobscura* populations. XV. Effective size of a natural population estimated by three independent methods. *Heredity*, **45**, 335-350.

Berthier P, Beaumont MA, Cornuet J-M, Luikart G (2002) Likelihood-based estimation of the effective population size using temporal changes in allele frequencies: A genealogical approach. *Genetics*, **160**, 741-751.

Burton C, Krebs CJ, Taylor EB (2002) Population genetic structure of the cyclic snowshoe hare (*Lepus americanus*) in southwestern Yukon, Canada. *Molecular Ecology*, **11**, 1689-1701.

Bjornstad ON, Peltonen M, Liebhold AM & Balstensweiler W (2002) Waves of larch budmoth outbreaks in the European Alps. *Science*, **298**, 1020-1023.

Casey TM, Joos BA (1983) Morphometrics, conductance, thoracic temperature, and flight energetics of Noctuid and Geometrid moths. Physiological Zoology, **56**, 160-173.

Cooper D, Cory JS, Myers JH (2003) Hierarchical spatial structure of genetically variable nucleopolyhedroviruses infecting cyclic populations of western tent caterpillars. *Molecular Ecology*, **12**, 881-890.

Delport W (2005) Temporal and spatial distribution of African Wild Silk Moth, *Gonometa postica*, eruptions in southern Africa. Chapter 2, University of Pretoria thesis, Pretoria, South Africa.

Delport W (2005) Characterisation of six microsatellite loci in the African Wild Silk Moth (*Gonometa postica*, Lasiocampidae). Chapter 3, University of Pretoria thesis, Pretoria, South Africa.

Delport W (2005) The effect of large annual population size fluctuations on spatial genetic pattern in the continuously distributed African Wild Silk Moth (*Gonometa postica*). Chapter 4, University of Pretoria thesis, Pretoria, South Africa.

Drummond AJ, Nicholls GK, Rodrigo AG, Solomon W (2002) Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. *Genetics*, **161**, 1307-1320.

Edwards ACF (1972) Likelihood: expanded edition. The Johns Hopkins University Press, London.

Ehrich D, Jorde PE, Krebs CJ, Kenney AJ, Stacy JE, Stenseth NC (2001) Spatial

structure of lemming populations (*Dicrostonyx groenlandicus*) fluctuating in density. *Molecular Ecology*, **10**, 481-495.

Ewens WJ (1979) Mathematical Population Genetics. Springer-Verlag, New York.

Falush D, Kraft C, Taylor NS, Correa P, Fox JG, Achtman M, Suerbaum S (2001) Recombination and mutation during long-term gastric colonization by Helicobacter pylori: Estimates of clock rates, recombination size, and minimal age. *Proceedings of the National Academy of Sciences, USA*, 98, 15056-16061.

Fu Y-X (2001) Estimating mutation rate and generation time from longitudinal samples of DNA sequences. *Molecular Biology and Evolution*, 18, 620-626.

Fuller SJ, Wilson JC & Mather PB (1997) Patterns of differentiation among wild rabbit populations *Oryctolagus cuniculus* L. in arid and semiarid ecosystems of north-eastern Australia. *Molecular Ecology*, **6**, 145-153.

Hartland-Rowe R (1992) The Biology of the wild silkmoth *Gonometa rufobrunnea* Aurivillius (Lasiocampidae) in northeastern Botswana, with some comments on its potential as a source of wild silk. *Botswana Notes and Records*, **24**, 123-133.

Hewitt G (2001) Speciation, hybrid zones and phylogeography – or seeing genes in space and time. *Molecular Ecology*, **10**, 537-549.

Ibrahim KM (2001) Plague dynamics and population genetics of the desert locust: can turnover during recession maintain population genetic structure? *Molecular Ecology*, **10**. 581-591.

Ibrahim KM, Nichols RA, Hewitt GM (1996) Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Heredity*, **77**, 282-291.

Jehle RJW, Arntzen WT, Burke AP, Hodl W (2001) The annual number of breeding adults and the effective population size of synoptic newts (*Triturus cristatus*, *T. marmoratus*). *Molecular Ecology*, **10**, 839-850.

Krimbas CB, Tsakas (1971) The genetics of *Dacus oleae*. V. Changes of esterase polymorphism in a natural population following insecticide control: selection of drift? *Evolution*, 25, 454-460.

Knowles LL, Maddison WP (2002) Statistical phylogeography. *Molecular Ecology*, **11**, 2623-2635.

Luikart G, Cornuet J-M, Allendorf FW (1999) Temporal changes in allele frequencies provide estimates of population bottleneck size. *Conservation Biology*, **13**, 523-530.

Nei M, Tajima F (1981) Genetic drift and estimation of effective population size. *Genetics*, **98**, 625-640.

Pollak E (1983) A new method for estimating the effective population size from allele frequency changes. *Genetics*, **104**, 531-548.

Roff DA & Bentzen P (1989) The statistical analysis of mitochondrial DNA polymorphisms: chi-square and the problem of small samples. *Molecular Biology and Evolution*, **6**, 539-545.

Rydell J, Skals N, Surlykke A & Svensson M (1997) Hearing and bat defence in geometrid winter moths. *Proceedings of the Royal Society of London B*, **264**, 83-88.

Turchin (2003) Complex Population Dynamics: a theoretical/empirical synthesis. *Monographs in Population Biology*, **35**. Princeton University Press, Princeton.

Vandewoestijne S, Neve G, Baguette (1999) Spatial and temporal population genetic structure of the butterfly *Aglais urticae* L. (Lepidoptera, Nymphalidae). *Molecular Ecology*, **8**, 1539-1543.

Veldtman R, McGeoch MA & Scholtz CH (2002) Variability in cocoon size in southern African wild silk moths: implications for sustainable harvesting. *African Entomology*, **10**, 127-136.

Veldtman R (2004) The ecology of southern African wild silk moths (*Gonometa* species, Lepidoptera: Lasiocampidae): consequences for their sustainable use. University of Pretoria PhD thesis.

Wade MJ & McCauley DE (1988) Extinction and recolonisation: their effects on the genetic differentiation of local populations. *Evolution*, **42**, 995-1005.

Wang J, Whitlock MC (2003) Estimating effective population size and migration rates from genetic samples over space and time. *Genetics*, 163, 429-446.

Wang J (2001) A pseudo-likelihood method for estimating effective population size from temporally spaced samples. *Genetical Research*, **78**, 243-257.

Wahlund S (1928) Zusammensetzung von Populationen und Korrelation-

serscheinungen von Standpunkt der Verebungslehre aus betrachtet. *Hereditas*, **11**, 65-106.

Waldbauer GP & Sternburg JG. (1982) Long mating flights by Hyalophora cecropia. *J. Lepid. Soc.*, **36**, 154-155.

Waples RS (1989) A generalized approach for estimating effective population size from temporal changes in allele frequency. *Genetics*, **121**, 379-392.

Whitlock MC & McCauley (1990) Some population genetic consequences of colony formation and extinction: genetic correlations with founding groups. *Evolution*, **44**, 1717-1724.

Williamson EG, Slatkin M (1999) Using maximum likelihood to estimate population size from temporal changes in allele frequencies. *Genetics*, **152**, 755-761.

Wright S (1940) Breeding structure of populations in relation to speciation. *American Naturalist*, **74**, 232-248.

Wright S (1943) Isolation by distance. *Genetics*, **28**, 114-138.