ANNALS OF FOREST RESEARCH
www.e-afr.org

A simulation study on the behavior of allelic richness and inbreeding coefficient over generations in fragmented populations of tree species

V. M. Stefenon, L. S. Costa

Stefenon V. M., Costa L. S. 2012. A simulation study on the behavior of allelic richness and inbreeding coefficient over generations in fragmented populations of tree species. Ann. For. Res. 55(1): 3-10, 2012.

Abstract. Computer simulations were employed in this study aiming to understand the effects of repeated cycles of inbred mating in isolated populations of tree species with different effective sizes and over up to 1000 generations. The results revealed a susceptibility of allelic richness to both, population size and repeated generations under inbred mating and a low but significant increase of the inbreeding coefficient over generations in populations with 50 and 100 plants, but not in populations with 500 and 1000 individuals. The loss of alleles occurred throughout all generations and was largely influenced by the population size. The most outstanding increase in the inbreeding coefficient occurred from the initial generation to the 5th generation, independent of the population size. The comparison of simulated results with data obtained from a field studie corroborated the hypothesis that isolated populations tend to more drastically suffer with loss of alleles and increase of inbreeding coefficient, while continuous forests, with effective production of fertile seeds and regeneration of seedlings, are inclined to preserve comparatively higher allelic richness and lower inbreeding coefficient over generations. In general, the results obtained highlight the importance of special care in selecting ESUs and preserving isolated populations, towards the conservation of forest genetic resources and adapatedness preservation.
Keywords adaptability, allelic richness, G_{IS}, heterozygosity, inbred mating, inbreeding coefficient.

Authors. Valdir Marcos Stefenon (valdirstefenon@unipampa.edu.br), Leonardo Severo da Costa - Universidade Federal do Pampa, Campus São Gabriel, Av. Antonio Trilha 1847, 97300-000, São Gabriel, RS, Brazil.
Manuscript received September 19, 2011; revised October 22, 2011; accepted November 21, 2011; online first January 11, 2012.

Introduction

The long life span and the sessile nature of tree species imply that they need to survive under different environmental conditions during their lifetime. Consequently, adaptation of trees to the environment is ubiquitous and required for survival (Finkeldey & Hattemer 2007). Adap-

tation is directly related to the species adaptedness – the degree to which an organism is able to live and reproduce in a specific environment (Erikson 2005) – promptly responding to environmental inconsistencies. In times of emergent climatic instability, such adaptedness is even more important and maintenance of genetic diversity is essential to uphold such competence, towards conservation of species genetic resources.

As result of forest fragmentation, large continuous forests have been converted to isolated fragments of different effective population sizes, becoming a dilemma to conservation geneticists. Conservation of genetic resources usually has been centered in determining evolutionary significant units (ESUs), which may be defined as populations reproductively separate from other populations and having unique or different adaptations (Waples 1991), or as populations that show significant divergence of allele frequencies at nuclear loci, in comparison to other populations (Moritz 1994). Operationally, efforts to determine ESUs have emphasized reproductive isolation of the populations (Crandal et al. 2000). Usually, isolated populations tend to present high genetic divergence of evolutionary lineages by local adaptation, in addition to high levels of inbreeding and lower allelic richness (Prakash et al. 1969, Newmann & Talmonn 2001, Stefenon et al. 2007).

Concerning conservation of forest tree genetic resources, the effects of such isolation in the levels of inbreeding and allelic richness over generations of such long-lived species is still a query.

Studies on population genetics of forest tree species usually report data from a single generation and are the only source of information towards planning management, conservation and restoration of forest environments. Given the long life span of tree species, a field survey aiming to understand the effect of repeated generations of inbred mating inside isolated populations is impracticable, since many hundred years will be necessary in order to study the population during tens of generations.

In this study, we report the results of a simulation-based survey employed to explore the levels of allelic richness and inbreeding coefficient in isolated populations of different sizes, over several generations of inbred mating. These simulations were employed to test the hypothesis that isolated populations tend to more drastically suffer with loss of alleles and increase of inbreeding coefficient, while continuous forests with effective production of fertile seeds and regeneration of seedlings are inclined to preserve comparatively higher allelic richness and lower inbreeding coefficient over generations. In order to test to which extend the simulated data fits the events occurring in real populations, we evaluated the correspondence of the simulated data with microsatellite records reported from two generations of three different conifer species.

Methods

This study relied on individual-based simulations of microsatellite genotypes of mature hermaphrodite diploid individuals, belonging to populations of four different effective sizes: 50, 100, 500 and 1000 individuals. These populations were simulated so that they replicate the behavior of completely isolated populations, lacking any level of gene influx. Data on number of alleles (A) and inbreeding coefficient (G_{IS}) were recorded for each population after 5, 45, 100, 500 and 1000 generations of mating among individuals, with a proportion of 50% of self pollination and 50% out-crossing. A preliminary round of five generations was run, in order to obtain the initial data on A and G_{IS} (generation 0). In order to avoid stochastic variations of the simulation process, all simulations were repeated 1000 times and the mean value over these simulations are reported. The most common assumption employed in genetic models was assumed for these simu-

lations: same mutation rate over all individuals and generations, non-overlapping generations, same mutation model and constant population size over generations. Simulations were performed using the software EASYPOP version 1.0 (Balloux 2001). Individuals were simulated in order to display information on diploid genotypes at 20 microsatellite loci, each reaching a maximal of 80 alleles. The stepwise mutation model (SMM, Kimura & Ohta 1975) was chosen for the evolutionary development of the genotypes. The mutation rate (mutations/locus/generation) was set to $\mu = 0.003$. Given that tree species are usually characterized by high levels of within population genetic diversity and relatively low levels of differentiation among populations (Loveless & Hamrick 1984, Finkeldey & Hattemer 2007), all simulations started from a population with maximal diversity (all individuals having the same probability of display any of the 80 alleles). Estimations of number of alleles (direct count) and inbreeding coefficient ($G_{IS} = 1 - H_o / H_S$, where H_o is the observed heterozygosity and H_S is the expected heterozygosity under Hardy-Weinberg equilibrium) for each simulated population were obtained using the software FSTAT (Goudet 1995) version 2.9.3. The statistical significance of the difference of genetic diversity measures over generations and over populations were tested through an analysis of variance (ANOVA), using the software StatPlus 2009.

The matching of simulated and empirical data was obtained by comparing the computer-generated data with number of alleles and inbreeding coefficient recorded for microsatellite markers in two generations (adults and juveniles) of three species of the genus *Araucaria: A. angustifolia* (Sant'Anna 2011), *A. nemorosa* and *A. columnaris* (Kettle et al. 2007). The statistical significance of the differences obtained across generations for number of alleles and inbreeding coefficient were determined using an uni-modal *t*-test.

Results

The recorded number of alleles was lower in smaller populations and reduced significantly over generations of inbred mating (Figure 1a). The reduction in number of alleles from the initial to the 1000th generation was about 5-fold for the populations with 50 and with 100 individuals, 3-fold for the populations with 500 individuals and 2-fold for the populations with 1000 individuals, confirming that the effective population size is very important for the maintenance of the allelic richness in isolated populations, which will develop under repeated generations of inbred mating. In all pair-wise comparisons of populations with different size for the same number of generations (Table 1), the differences were statistically significant at 1% level, according to the ANOVA test ($p < 0.001$). In the pair-wise comparisons of generations for each population size (Table 2), just the pair 500 vs 1000 generations revealed a non-significant difference ($p = 0.145$) for the population with 50 individuals. All the other pairs revealed a significant difference in number of alleles ($p < 0.001$).

The inbreeding coefficient (G_{IS}), measured as the loss of heterozygote individuals in relation to expected under Hardy-Weinberg equilibrium, was slightly higher in larger populations and increased over generations. The enhancement over generations was proportional in all populations, independent of the effective population size, but with an outstanding increase in the five initial generations and a relative constancy over the following generations (Figure 1b). Despite the overall small increase, the pair-wise comparisons of populations with different size for the same number of generations (Table 1) revealed a statistically significant difference for all pair of populations ($p < 0.014$), except for the pair 100 vs 500 individuals after 100 generations ($p = 0.154$) and for the pair 500 vs 1000 individuals ($p > 0.459$), independent of the number of generations. Concerning the pair-wise comparisons of generations for

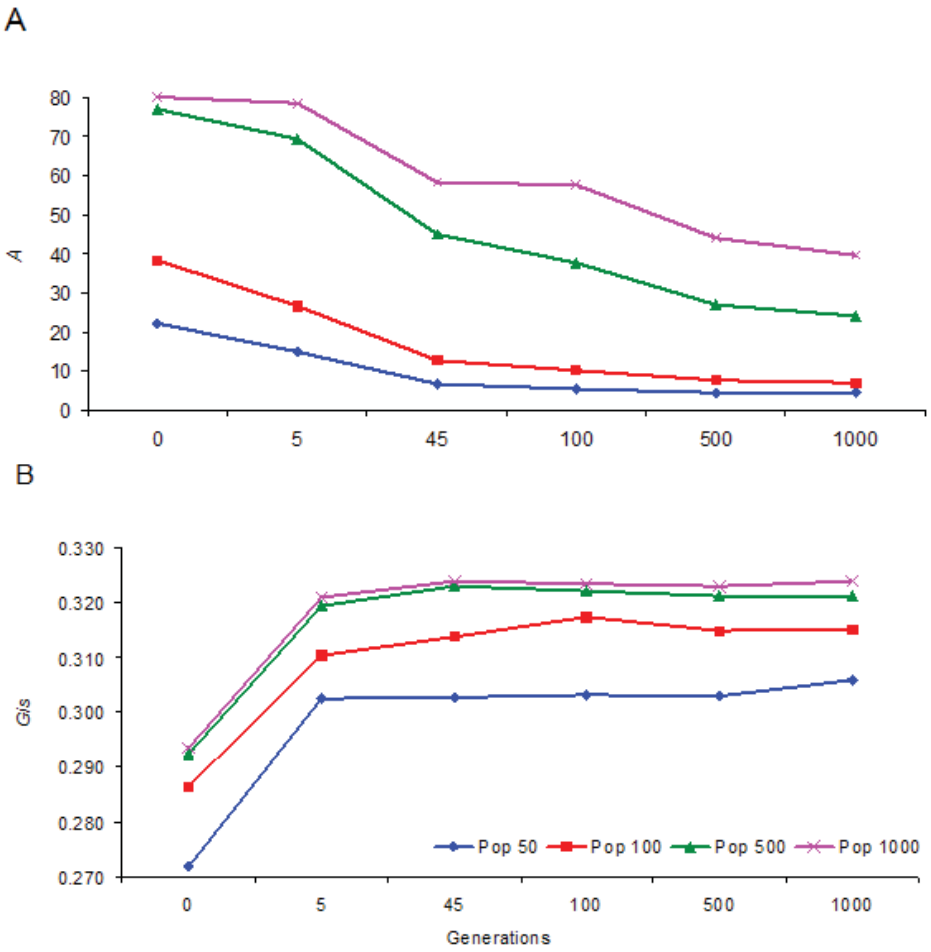

Figure 1 Estimations of genetic diversity measures for four different populations size (50, 100, 500 and 1000 individuals) over 0, 5, 45, 100, 500 and 1000 generations of inbred mating. In (A): estimations of the number of alleles and in (B): estimations of the inbreeding coefficient.

each population size (Table 2), all populations differed ($p < 0.001$) from the generation 0 (0 vs 5, 0 vs 45, 0 vs 100, 0 vs 500, 0 vs 1000). The pairs 5 vs 45 and 5 vs 100 generations for populations with 1000 individuals also revealed a significant difference ($p < 0.05$) in the measure of inbreeding coefficient. The non-significant difference in the measure of inbreeding coefficient in all other pair-wise comparisons suggests a relative stability of the levels of inbreeding, mainly after ten generations, and high influence of the effective population size over this index.

The data from two out of the three field studies matched the behavior revealed by our simulation study. The records from *A. angustifolia* (Sant'Anna 2011) revealed an increase in the inbreeding coefficient from the first to the second generation (G_{IS} = 0.096 and G_{IS} = 0.163, respectively; uni-modal *t*-test p = 0.11) and a decrease in the mean number of alleles over generations (A = 9.33 for the first genera-

tion and A = 8.67 for the second generation; uni-modal t-test p = 0.05).

Similarly, the data from *A. nemorosa* expressed the same patterns for inbreeding coefficient (G_{IS} = 0.108 for the first generation and G_{IS} = 0.192 for the second generation; uni-modal t-test p = 0.008) and for the mean number of alleles (A = 7.45 for the first generation and A = 6.98 for the second generation; uni-modal t-test p = 0.02). The data from *A. columnaris*, on the other hand, expressed a decrease level of inbreeding coefficient over generations (G_{IS} = 0.123 for the first generation and G_{IS} = 0.093 for the second generation; uni-modal t-test p = 0.45), while the mean number of alleles did not differed (A = 6.0 for both generations; uni-modal t-test p = 0.011).

Table 1 Statistical significance of pair-wise comparisons of the number of alleles (A) and coefficient of endogamy (G_{IS}) for population-pairs of different sizes over 5, 45, 100, 500 and 1000 generations, based on an ANOVA test.

Population size	A (generations)					G_{IS} (generations)				
	5	45	100	500	1000	5	45	100	500	1000
50 vs 100	***	***	***	***	***	***	**	***	***	***
50 vs 500	***	***	***	***	***	***	***	***	***	***
50 vs 1000	***	***	***	***	***	***	***	***	***	***
100 vs 500	***	***	***	***	***	**	**	**	*	**
100 vs 1000	***	***	***	***	***	**	***	***	***	***
500 vs 1000	***	***	***	***	***	n.s.	n.s.	n.s.	n.s.	n.s.

Note: Significance level: *** p < 0.001, ** p < 0.01, * p < 0.05, n.s.: not significant

Table 2 Statistical significance of pair-wise comparison of the number of alleles (A) and coefficient of endogamy (G_{IS}) across generations of inbred mating over populations of different sizes, based on an ANOVA test.

Number of generations	A (population size)				G_{IS} (population size)			
	50	100	500	1000	50	100	500	1000
0 vs 5	***	***	***	***	***	***	***	***
0 vs 45	***	***	***	***	***	***	***	***
0 vs 100	***	***	***	***	***	***	***	***
0 vs 500	***	***	***	***	***	***	***	***
0 vs 1000	***	***	***	***	***	***	***	***
5 vs 45	***	***	***	***	n.s.	n.s.	n.s.	**
5 vs 100	***	***	***	***	n.s.	n.s.	n.s.	*
5 vs 500	***	***	***	***	n.s.	n.s.	n.s.	n.s.
5 vs 1000	***	***	***	***	n.s.	n.s.	n.s.	n.s.
45 vs 100	***	***	***	***	n.s.	n.s.	n.s.	n.s.
45 vs 500	***	***	***	***	n.s.	n.s.	n.s.	n.s.
45 vs 1000	***	***	***	***	n.s.	n.s.	n.s.	n.s.
100 vs 500	***	***	***	***	n.s.	n.s.	n.s.	n.s.
100 vs 1000	***	***	***	***	n.s.	n.s.	n.s.	n.s.
500 vs 1000	n.s.	***	***	***	n.s.	n.s.	n.s.	n.s.

Note: Significance level: *** p < 0.001, ** p < 0.01, * p < 0.05, n.s.: not significant

Discussion

Even if the simulations employed in this study fail to cover all the possible scenarios concerning mating systems and ecological issues, they bring out important considerations related to the behavior of allelic richness and inbreeding coefficient over generations in isolated populations of tree species.

The larger decrease in the number of alleles over generations for smaller populations (50 and 100 individuals) under inbred mating observed in this study is not surprising, but exposes the need of special attention for small forests, which usually display low number of alleles. Low number of alleles can retain high heterozygosity without maintaining allelic richness. However, allelic richness is essential for safeguarding populations' adaptedness, that is vital for forest tree populations which are constantly exposed to different evolutionary factors, changing frequencies of alleles and/or genotypes (Erikson 2005).

Due to their sessile nature, the production of inbred offspring in plant populations is likely. While some plant species presents the possibility of self pollination, out-crossing species may be pollinated by genetically related plants growing in the near vicinity. Both source of inbreeding are usually more intense in small populations and for species with limited gene flow (e.g. Stefenon et al. 2008a). An important aspect of our study on this subject is the higher increase of the inbreeding coefficient occurred from the generation 0 to the 5th generation, that is, in the initial period after the forest isolation (around 50 to 100 years, considering species starting to reproduce with 10 to 20 years), followed by a scenario of stabilization through the further generations.

The comparatively lower inbreeding coefficient observed in the smaller populations (50 and 100 individuals) in our simulation may be effect of the very high levels of heterozygosity (H = 0.98), emphasizing the importance of the maintenance of high diversity. This is observed when comparing the level of increase in the inbreeding coefficient over populations. The smaller population (50 individuals) revealed the higher loss of alleles over generations, with a decrease of 5-fold in number of alleles from the generation 0 to the 1000th generation. This population revealed a small but comparatively higher (uni-modal t-test $p < 0.001$) increase in the inbreeding coefficient in the same time interval, in comparison to the larger population, likely as effect of the lower allelic richness in the smaller population. This result highlights the importance of maintaining allelic richness towards the conservation of forest genetic resources.

The contrasting result concerning inbreeding coefficient and mean number of alleles for the three species of *Araucaria* evaluated in this study reflects the importance of considering ecological and reproductive/regenerative attributes of each species. Present-day populations of *A. angustifolia* and *A. nemorosa* are highly fragmented and present relatively low seedling regeneration (Stefenon et al. 2007, 2008a, Bittencourt & Sebbenn 2008, Kettle et al. 2007), whereas populations of *A. columnaris* comprise many thousands of individuals which produce abundant fertile seeds and show plentiful seedling regeneration (Kettle et al. 2007).

In addition to the ecological aspects, the kind of genetic marker used in the study and the pressure of selection should also be considered. For instance, a natural population of *A. angustifolia* studied with isozyme markers (Ferreira 2008), which may suffer adaptive pressure, revealed decrease in allelic richness from adults (first generation) to plantlets (A = 1.9 and 1.6, respectively), but an increased number of alleles in the seeds of the first generation (A = 2.2), suggesting that many of the seeds will not further develop and alleles will be lost.

Considering that microsatellite markers are usually neutral, are highly polymorphic and presents elevated number of alleles, they may

better represent the general behavior of allelic richness and inbreeding coefficient over generations, but do not necessarily assess selective patterns.

Considering that heterozygosity, allelic richness and inbreeding coefficient are typically estimated in population genetic surveys, these measures should be the start point for planning forest genetic resources conservation and management. However, care is needed when interpreting data concerning genetic diversity and conservation. Some studies based on computation of *H* and *F*-statistics consider critical populations for conservation based on the estimations of F_{ST} and similar measures. If estimated F_{ST} is low, the subpopulations are usually considered to be not highly differentiated and therefore, the maintenance of different subpopulations of the species is considered not important (Jost 2008). However, F_{ST} and similar metrics are not measures of genetic differentiation and may lead to equivocated decisions concerning genetic resources conservation (for a broad discussion on this topic, see Jost 2008).

Conclusions

Although the levels of heterozygosity are an important indicator for the selection of important ESUs, it should not be the main feature evaluated since high levels of heterozygosity can exist also for loci containing just two alleles. Special attention should be given to the number of alleles of the selected populations. Reduction of the allelic richness and increase of the inbreeding coefficient over generations are expected for isolated populations of forest tree species, which usually display relatively small effective population size. Thus, allelic richness and levels of inbreeding should also be considered when selecting ESUs aiming the conservation of forest species genetic resources.

Moreover, sustainable management of planted forests and forest remnants, as well as forestation/reforestation enterprises should care on population genetic structure and levels of gene flow of the target species, towards maintenance of allelic richness (Stefenon et al. 2008a). Since ecological, demographic and reproductive features of the species are important for the decisions about genetic resource conservation, further studies covering a larger range of options have to be considered, in order to add more information in this issue.

Acknowledgements

This research was made possible by the financial support from FAPERGS (Process n. 0902170). The Authors are sponsored by a PIBIC scholarship (LSC) and a PQ grant (VMS) from CNPq (Process n. 314664/2009-2). The comments of one anonymous reviewer greatly helped to improve the final version of this paper.

References

Balloux F., 2001. EASYPOP (version 1.7): a computer program for population genetics simulations. The Journal of Heredity 92: 301-302.

Bittencourt J.V.M., Sebbenn A.M., 2007. Patterns of pollen and seed dispersal in a small, fragmented population of the wind-pollinated tree *Araucaria angustifolia* in southern Brazil. Heredity 99: 580-591.

Crandal K.C., Bininda-Emonds O.R.P., Mace G.M., Wayne R.K., 2000. Considering evolutionary processes in conservation biology. Trends in Ecology and Evolution 15: 290-295.

Erikson G., 2005. Evolution and evolutionary factors, adaptations and adaptability. In: Geburek Th., Turok J. (eds.) Conservation and Management of Forest Genetic Resources in Europe. Arbora Publishers, Zvolen, pp. 199-211.

Ferreira D.K., 2008. Caracterização genética e estrutura populacional de diferentes origens de *Araucaria angustifolia* na FLONA de Três Barras [Genetic characterization and population structure of different origins of *Araucaria angustifolia* in the FLONA of Três Barras] Msc Dissertation. Department of Plant Sciences, Federal University of Santa Catarina, Florianópolis, 83 p.

Finkeldey R., Hattemer H.H., 2007. Tropical Forest Genetics. Springer, Berlin, Heidelberg, New York.

Goudet J., 1995. Fstat version 1.2: a computer program to calculate *F*-statistics. The Journal of Heredity 86: 485-486.

Jost L., 2008. G_{ST} and its relatives do not measure differentiation. Molecular Ecology 17: 4015-4026.

Kettle C.J., Hollingsworth P.M., Jaffré T., Moran B., Ennos A., 2007 Identifying the early genetic consequences of habitat degradation in a highly threatened tropical conifer, *Araucaria nemorosa* Laubenfels. Molecular Ecology 16: 3581–3591.

Kimura M., Ohta T., 1975. Distribution of allelic frequencies in a finite population under stepwise production of neutral alleles. Proceedings of the National Academy of Sciences of USA 72: 2761-2764.

Loveless M.D., Hamrick J.L., 1984. Ecological determinants of genetic structure in plant populations. Annula Reviews of Ecology and Systematics 15: 65-95.

Moritz C., 1994. Defining 'Evolutionarily Significant Units' for conservation. Trends in Ecology and Evolution 9: 373-375.

Newman D., Tallmon D.A., 2001. Experimental evidence for beneficial fitness effects of gene flow in recently isolated populations. Conservation Biology 15: 1054-1063.

Prakash S., Lewontin R.C., Hubby J.L., 1969. A Molecular Approach to the Study of Genic Heterozygosity in Natural Populations IV. Patterns of Genic Variation in Central, Marginal and Isolated Populations of *Drosophila pseudoobscura*. Genetics 61: 841-858.

Sant'Anna C., 2011. Diversidade genética, estrutura genética espacial e dispersão realizada de pólen e sementes em uma população contínua de *Araucaria angustifolia* (Bertol.) Kuntze no planalto norte de Santa Catarina. [Genetic diversity, spatial genetic structure and effective dispersion of pollen and seeds in a continuous population of *Araucaria angustifolia* (Bertol.) Kuntze in the north Highlands of Santa Catarina]. M.Sc. Dissertation. Department of Plant Sciences, Federal University of Santa Catarina, Florianópolis, 89 p.

Stefenon V.M., Gailing O., Finkeldey R., 2007. Genetic Structure of *Araucaria angustifolia* (Araucariaceae) Populations in Brazil: Implications for the *in situ* Conservation of Genetic Resources. Plant Biology 9: 516-525.

Stefenon V.M., Gailing O., Finkeldey R., 2008b. The role of gene flow in shaping genetic structures of the subtropical conifer species *Araucaria angustifolia*. Plant Biology 10: 356-364.

Waples R.S., 1991. Pacific salmon, *Oncorynchus* spp., and the definition of 'species' under the endangered species act. Marine Fisheries Review 53: 11-22.