UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ECOLOGIA

TESE DE DOUTORADO

Efeitos do solo, geografia, demografia e uso e manejo da terra na diversidade e estrutura genética e estratégia reprodutiva de populações naturais de *Caryocar brasiliense* e *Dipteryx alata*

NATASHA BRIANEZ RODRIGUES

Orientador: Dr. Aldicir Osni Scariot

Co-orientadora: Drª. Vânia Cristina Rennó Azevedo

Brasília – DF
Março de 2017

UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ECOLOGIA

TESE DE DOUTORADO

Efeitos do solo, geografia, demografia e uso e manejo da terra na diversidade e estrutura genética e estratégia reprodutiva de populações naturais de *Caryocar brasiliense* e *Dipteryx alata*

NATASHA BRIANEZ RODRIGUES

Orientador: Dr. Aldicir Osni Scariot

Co-orientadora: Drª. Vânia Cristina Rennó Azevedo

Tese apresentada ao Programa de Pós-Graduação em Ecologia da Universidade de Brasília, como pré-requisito para a obtenção do título de Doutora em Ecologia.

Brasília – DF
Março de 2017

Tese de Doutorado

NATASHA BRIANEZ RODRIGUES

Efeitos do solo, geografia, demografia e uso e manejo da terra na diversidade e estrutura genética e estratégia reprodutiva de populações naturais de *Caryocar brasiliense* e *Dipteryx alata*

Tese aprovada junto ao Programa de Pós-Graduação em Ecologia da Universidade de Brasília como requisito parcial para obtenção do título de Doutora em Ecologia.

Banca examinadora:

Dr. Aldicir O. Scariot
Presidente/Orientador
Depto. de Ecologia – UnB/Embrapa Cenargen

Dr. Alexandre S. G. Coelho
Membro Titular
Setor de Melhoramento de Plantas - UFG

Drª. Christina C. Vinson Williams
Membro Titular
Embrapa Cenargen

Drª. Fernanda A. Gaiotto
Membro Titular
Depto. de Ciênc. Biol., Genét. e Biol. Mol. - UESC

Dr. Ludgero C. G. Vieira
Membro Titular
UnB - FUP

Brasília – DF
Março de 2017

“Thanks to impermanence, everything is possible”

Thich Nhat Hanh

Ao meu marido, Stephano.

Agradecimentos

Agradeço:

Meu marido querido, Stephano, pela força, suporte, carinho, paciência, por acreditar em mim, por existir na minha vida.

Meus pais, Zenaide e Manoel, por me apoiarem nessa jornada.

Bhumi e Mita, por colorirem minha vida em momentos sombrios.

Meus orientadores, Vânia Rennó e Aldicir Scariot, por acreditarem em mim, me darem essa oportunidade e pelo todo apoio ao longo destes vários anos.

Nilton Barbosa, Juarez Amaral e Juliana Benez, pelo auxílio na coleta de campo.

Carolina Pamplona e Flávio Barreto, pelo auxílio nas análises de laboratório.

Aelton Giroldo e Juliana Benez, pela ajuda quanto à caracterização das áreas de estudo.

Joseane Padilha, pelo auxílio e sugestões com análises estatísticas.

Sérgio Noronha, pela assistência no desenvolvimento de imagens.

Todos os amigos do laboratório: Vânia, Neide, Lorena, Marília, Bruna, Lucileide, Peter, Anádria, Nayara e demais colegas. Obrigada pela amizade, por fazerem parte dessa jornada e fornecerem ajuda no dia a dia quanto aos mais diversos problemas, pessoais e de trabalho.

Departamento de Ecologia da UnB, pela oportunidade.

CAPES e CNPq, pelas bolsas concedidas.

Embrapa Cenargen, pelo apoio estrutural e logístico.

Funbio, pelo financiamento do projeto.

Proprietários de terras, por permitir o acesso às áreas de estudo.

Sumário

Índice de figuras

Capítulo I. Ecological and anthropic associations with genetic diversity of populations of *Caryocar brasiliense*, a neotropical savanna tree

Capítulo II. Cattle ranching affects the genetic diversity of natural populations of *Dipteryx alata*, a vulnerable tree from the Brazilian savanna

Capítulo III. Do demographic persistence strategies of two neotropical savanna trees include clonal reproduction?

Índice de tabelas

Resumo

Para embasar decisões de manejo para a conservação da biodiversidade, é necessário não apenas descrever a variação genética, mas também identificar os fatores específicos que a afetam. É também essencial conhecer a habilidade de espécies de se reproduzirem vegetativamente, uma estratégia que permite a persistência demográfica frente a distúrbios. Foram estudadas associações entre variáveis de solo, geográficas, demográficas e de uso e manejo da terra e a diversidade e estrutura genética de 20 populações de duas árvores do Cerrado (*Dipteryx alata* Vog. e *Caryocar brasiliense* Camb.), através de modelos lineares generalizados. Foi investigada a magnitude de reprodução clonal em populações destas espécies em diferentes condições de uso e manejo da terra, através da identificação de genótipos multilocos. Aproximadamente 2.300 indivíduos foram genotipados por microssatélites. Encontramos gradientes latitudinais de riqueza alélica (*Ar*) e heterozigosidade esperada (*He*) para *C. brasiliense* e altitudinais de *Ar* para *D. alata*. Detectamos associações negativas e significativas entre criação de gado e *Ar* para ambas as espécies e *He* para *C. brasiliense.* A frequência de fogo apresentou associação negativa e significativa com *Ar* e *He* para *C. brasiliense*. O extrativismo de frutos não apresentou associação significativa com a diversidade genética de nenhuma espécie, enquanto o conteúdo de areia do solo a apresentou para ambas as espécies. Não foram encontrados clones para nenhuma das espécie. Apenas 8 de 15 locos para *D. alata* e 4 de 9 para *C. brasiliense* foram capazes de distinguir todos os genótipos multilocos eficientemente. Nós discutimos possíveis mecanismos que possam explicar as associações encontradas, propondo estratégias para auxiliar na conservação de diversidade genética das espécies em paisagens de uso-múltiplo. A inabilidade de reprodução clonal pelas espécies avaliadas é discutida à luz de suas características ecológicas e histórias demográficas. Possíveis implicações para a conservação genética são apresentadas.

Palavras-chave: genética de populações; criação de gado; produtos florestais não-madeireiros; fogo; reprodução clonal; Cerrado.

Abstract

To support management decisions concerning biodiversity conservation, it is necessary not only to describe genetic variation, but also to identify the specific factors affecting it. It is also essential to know the ability of species to reproduce vegetatively, a strategy that enables demographic persistence in face of disturbances. We studied the associations of soil, geographic, demographic and land use and management variables with the genetic diversity and structure of 20 populations of two trees from the Cerrado (*Dipteryx alata* Vog. and *Caryocar brasiliense* Camb.), using generalized linear models (GLMs). We investigated the magnitude of clonal reproduction in populations of these species in different conditions of land use and management, through identification of multilocus genotypes (MLGs). Approximately 2,300 individuals were genotyped with microssatelites. We found latitudinal gradients of allelic richness (*Ar*) and expected heterozygosity (*He*) for *C. brasiliense* and altitudinal gradients of *Ar* for *D. alata*. There were negative and significant associations between cattle ranching and *Ar* of populations of both species and *He* for *C. brasiliense*. Fire frequency was negatively and significantly associated with *Ar* and *He* of populations of *C. brasiliense*. Fruit harvesting did not present significant associations with the genetic diversity of neither species, while soil sand content did for both species. No clones were encountered for either species. Merely 8 out of 15 loci for *D. alata* and 4 out of 9 loci for *C. brasiliense* were able to differentiate all MLGs efficiently. We discuss possible mechanisms that might explain the encountered associations, proposing strategies to aid in the conservation of genetic diversity of the studied species in multiple-use landscapes. The inability of reproducing clonally by both species is discussed in light of their ecological characteristics and demographic histories. Possible implications for conservation of their genetic diversity are presented.

Keywords: population genetics; cattle ranching; non-timber forest products; fire; clonal reproduction; Cerrado.

Introdução geral

A genética da conservação é uma área aplicada de conhecimento que vem crescendo de maneira contínua ao longo das últimas quatro décadas, subsidiando a conservação não apenas em nível de genes, mas também em nível de espécies e ecossistemas (Avise 2008; Frankham 2010). Dentre as áreas emergentes e os desafios para a conservação genética está a integração de informações genéticas com outros fatores biológicos e não-biológicos, combinando metodologias genéticas com abordagens ecológicas, o que é capaz de fornecer uma representação mais acurada de padrões e processos em sistemas complexos (DeSalle & Amato 2004). Para subsidiar estratégias efetivas de conservação e manejo genético de populações naturais, é necessário, além da simples descrição da diversidade e estrutura genética, identificar os fatores que afetam tal diversidade, quantificando seus efeitos e elucidando seus mecanismos subjacentes (Rao & Hodgkin 2002). Nesse sentido, é necessário integrar a genética de populações em um contexto maior que inclui a presença de diversos fatores, como de solo, geográficos, demográficos e antropogênicos (Frankham 2010), de maneira a elucidar padrões genéticos em diferentes escalas e identificar os fatores de maior importância para o manejo e conservação de espécies de maneira específica.

Dentre as informações ecológicas imprescindíveis para o planejamento da conservação e manejo de espécies e populações está o conhecimento sobre mecanismos reprodutivos que assegurem a viabilidade demográfica em condições de estresse ambiental. A maioria das plantas perenes possui a aptidão de combinar dois tipos de reprodução: sexuada e vegetativa (Eckert 2001; Arnaud-Haond *et al.* 2007). O investimento relativo em cada um destes modos reprodutivos geralmente é função de fatores ecológicos, sendo a reprodução clonal favorecida em situações envolvendo distúrbios ambientais (Bond & Midgley 2001, 2003; Eckert 2001). Nestas situações, a reprodução vegetativa possui alta importância para a persistência demográfica de populações, não passando por atribulações relacionadas à regeneração por sementes e permitindo amortecimento de impactos genéticos devido a atrasos de respostas (Bond & Midgley 2001, 2003; Eckert 2001). Considerando que a regeneração é o estágio de vida mais passível de mudanças genético-estruturais em populações de espécies arbóreas (Finkeldey & Ziehe 2004; Ratnam *et al.* 2014), quantificar a ocorrência e detectar fatores que acarretam em

reprodução clonal é grande relevância para o correto manejo e conservação de populações e espécies.

O Cerrado, a maior savana neotropical, é uma das áreas mais biologicamente ricas e ameaçadas do mundo (Ratter *et al.* 1997). Devido a agricultura, criação de gado e ocupação urbana, essa formação existe em apenas cerca de 50% de sua área original (MMA 2011), sendo apenas cerca de 3% do seu território protegido (Françoso *et al.* 2015). Dentre as espécies arbóreas mais comuns desta região estão *Dipteryx alata* Vog. (Fabaceae), o Baru, e *Caryocar brasiliense* Camb. (Caryocaraceae), o Pequi (Ratter *et al.* 2003), árvores socioeconomicamente importantes devido a suas sementes e frutos, respectivamente, os quais fornecem renda familiar e alimento para diversas comunidades extrativistas (Araújo 1995; Almeida *et al.* 1998; Sano *et al.* 2004).

Considerando as ideias apresentadas, os objetivos deste trabalho são:

1) Identificar e avaliar possíveis causas e mecanismos que possam interferir na diversidade e estrutura genética de populações nativas de *Caryocar brasiliense* e *Dipteryx alata* no Cerrado brasileiro, quantificando efeitos de associações entre variáveis de solo, geográficas, demográficas e de uso e manejo da terra e variáveis de diversidade e estrutura genética, através de marcadores microssatélites e modelos lineares generalizados, de maneira a subsidiar estratégias de conservação e manejo para estas espécies (capítulos I e II).

2) Verificar e quantificar a existência de reprodução clonal por brotação de raízes em populações naturais de *Caryocar brasiliense* e *Dipteryx alata* no Cerrado em diferentes contextos de uso e manejo da terra, através da identificação de genótipos multilocos por marcadores microssatélites, de maneira a compreender os efeitos de distúrbios antropogênicos na magnitude de ocorrência deste mecanismo de persistência demográfica, avaliando possíveis consequências para a conservação destas espécies (capítulo III).

I. Ecological and anthropic associations with genetic diversity of populations of *Caryocar brasiliense*, a neotropical savanna tree

1. Introduction

What determines and influences genetic variance? According to the Hardy-Weinberg principle, genetic drift, non-random mating, selection, mutation and gene flow are the forces that drive evolution in populations, through changes in allelic and genotypic frequencies (Hedrick 2011). These forces are influenced by particular factors, such as environmental, ecological, geographical and anthropogenic patterns. In order to support management decisions concerning conservation of biodiversity and assist in planning of sustainable use of natural resources, it is necessary not only to describe genetic variation, but also to identify the specific factors affecting it, quantifying their effects and elucidating their underlying mechanisms (Rao & Hodgkin 2002; Allendorf & Luikart 2007).

Geographic features, such as latitudinal and altitudinal gradients, can have important roles in the distribution and structuring of genetic diversity. Latitudinal gradients of genetic diversity can be explained by two models: the latitudinal and the species diversity model (Schrey *et al.* 2011a). The first model predicts an increase of genetic diversity with latitudes closer to zero, which can be explained by different ecological and evolutionary processes that operate simultaneously (Martin & McKay 2004; Adams & Hadly 2013). The second model postulates that the factors that contribute to species diversity might also contribute to genetic diversity, such as locality characteristics and ecological and demographic processes (Vellend & Geber 2005). Altitudinal gradients, in turn, can be related to genetic diversity in complex ways; even small gradients can encompass many different environmental and ecological variables that can constrain the genetic diversity of populations (Ohsawa & Ide 2008; Thiel-Egenter *et al.* 2009; Yan *et al.* 2009; Shen *et al.* 2014).

The environment is one of the main factors that structure genetic diversity in space and time, through complex interactions and natural selection that leads to local adaptation (Rao & Hodgkin 2002). For plants, specifically, soil conditions can limit survival, growth and reproduction, as it can constrain the availability of nutrients and water (Gurevitch *et al.* 2006; Stein *et al.* 2016). Environmental stress through extreme soil features can be an important driver of evolution for species (Li *et al.*

2016), which is particularly important to this case, since our study area, the Cerrado, presents harsh soil conditions.

The genetic diversity and structure of populations can also be affected by anthropogenic activities, which can lead to changes such as alteration of population subdivision, loss of genetic diversity and changes due to selection (Allendorf *et al.* 2008). Disturbance might be the main driver that shapes the genetic diversity in several populations (Banks *et al.* 2013; Davies *et al.* 2016). The effects associated with forest fragmentation (Young *et al.* 1996; Lowe *et al.* 2005; Kramer *et al.* 2008; Aguilar *et al.* 2008; Vranckx *et al.* 2012) and timber logging (Finkeldey & Ziehe 2004; Wernsdörfer *et al.* 2011; Ratnam *et al.* 2014) have been reasonably studied. However, studies still lack on the genetic consequences of other types of disturbances, such as livestock farming (but see Mengli *et al.* 2005; Shan *et al.* 2006), non-timber forest products (NTFP) harvesting (but see Wang *et al.* 2013; Xu *et al.* 2013; Gaoue *et al.* 2014; Shaanker *et al.* 2016) and use of fire (but see Rajora & Pluhar 2003; Uchiyama *et al.* 2006; Neville *et al.* 2009; Schrey *et al.* 2010, 2011a; Suárez *et al.* 2012; Smith *et al.* 2014). These activities not only are very common, but also of great socio-economic importance in many cultures and countries; thus, understanding its association with genetic diversity patterns is essential.

Natural populations can present significant interactions between demography and genetic diversity (Goodell *et al.* 1997; Gibbs 2001). Genetic diversity is affected by disturbance through both selective and neutral, demographically driven, processes (Banks *et al.* 2013). The great portion of genetic diversity that is selectively neutral is strongly influenced by neutral demographic processes, such as mortality, reproduction, movement and social behavior. Thus, the quantification of demographic parameters can help to clarify mechanisms underlying effects of geography, environment and disturbances on genetic variance. In addition, integrating demography into population genetics usually leads to more useful recommendations for conservation (DeSalle & Amato 2004).

The Cerrado, the largest neotropical savanna, located in the central area of Brazil, occupies ca. 2 million km^2, corresponding to 23% of the country area (Ratter *et al.* 1997; Furley 1999). It is one of the richest and most threatened areas of the world (Ratter *et al.* 1997), presenting approximately 12,000 plant species, of which 4,200 are endemic (Forzza *et al.* 2010). This savanna presents strong climate seasonality, with marked dry and rainy seasons (Silva *et al.* 2008). Most of its soils

are old, weathered, acid, poor in nutrients and with high concentrations of aluminum (Furley & Ratter 1988; Haridasan 2000). An important feature of the Cerrado is its fire regime; fires occurs typically at intervals of 1-3 years, a rate that is significantly associated to anthropogenic causes, although natural fires have existed in this region long before the arrival of humans (Hoffmann 1998; Furley 1999; Miranda *et al.* 2002). Agriculture and cattle ranching activities have led to the massive devastation of this savanna, that today is only present in 50% of its original area (MMA 2011). In spite of great need for conservation, it was not until recently that some attention has turned to the Cerrado, and still only approximately 3% of its territory is protected (Françoso *et al.* 2015).

Caryocar brasiliense Camb. (Caryocaraceae), commonly known as “Pequi”, is a tree species that is common and well distributed throughout the Brazilian savanna, occurring from low densities to highly aggregate groups in typical Cerrado vegetation (Almeida *et al.* 1998). This species is highly important socio-economically due to its nutritional fruit, used in various forms by local people and harvested manually by traditional communities on natural populations in native vegetation, crops and pasture areas, where they are commonly maintained and sometimes nursed after (Araújo 1995; Almeida *et al.* 1998). *C. brasiliense* is diploid, with hermaphrodite flowers (Araújo 1995); it is pollinated by small nectarivorous bats and its fruits and seeds are dispersed through gravity and the action of animals, mostly by the marsupial *Didelphis albiventris* and the corvid *Cyanocorax cristatellus* (Gribel 1986; Gribel & Hay 1993); it is preferentially allogamous (Collevatti *et al.* 2001a, 2010). Due to its ecological, social and economical importance, it can be considered a strategic species for studies in the Cerrado.

We aimed to quantify and understand associations of geography, soil, demography and land use and management with genetic diversity and structure of populations of *Caryocar brasiliense*, an important tree species of the Brazilian savanna. The following question was asked: Do geographic, soil, demographic and land use and management variables affect the genetic diversity and structure of populations of *C. brasiliense*? To answer this, we analyzed naturally occurring populations with the use of microsatellite markers and generalized linear models.

2. Methodology

Study areas characterization and sampling

We sampled 20 areas containing populations of *C. brasiliense* in the Cerrado (Figure 1), which presents tropical climate with dry winter and rainy summer seasons (Aw - Köppen). These areas are subjected to different land use and management, together composing a gradient of conservation status. Six of these 20 areas were selected to form three geographical units composed of paired populations in two contrasting land use types: native Cerrado vegetation and pastureland (Figure 1). This strategy was used to control for history and geography on genotypic diversity in each geographical unit. The land use and management impacts on the demographic structure of these populations were previously studied (Giroldo & Scariot 2015).

Although the study areas present a wide range of land use types and management, we focused on the impacts associated with *C. brasiliense* fruit harvesting, cattle ranching, vegetation thinning and occurrence of fire, which can all occur at different intensity levels. We used a modified interaction matrix (Leopold *et al.* 1971) to estimate the disturbance caused by cattle ranching, vegetation thinning and *C. brasiliense* fruit harvesting. Based on information from landowners and field observations, we classified areas according to the severity, duration and area of disturbance in a rank of 0-3 for these variables. Based on the sum of these three components, we constructed a 0-9 rank index for each of the three variables. We determined the frequency of fire by counting fire scars in the 11 years prior to the field sampling using LandSat 5.0 images with false-color composition R4G5B3. No area presented fire occurrence more than once a year and we used the sum of fire events in each area as fire frequency.

To characterize the demographic structure of populations, we sampled in each area a contiguous plot varying from 1 to 7 ha according to fragment size and disturbance level. The stem diameter at 30 cm above soil (D_{30cm}) and height of every individual with $D_{30cm} \geq 10$ cm (hereinafter referred to as adults) inside plots were measured. For individuals with more than one stem, we converted each measurement into basal area, summed and transformed back to a single diameter. To characterize the soil, three samples were collected randomly at a depth of 0-30 cm in each area. These were homogenized and analyzed as to particle size (sand,

silt and clay), pH-active acidity, percent base saturation - V%, aluminum saturation - m% and cation exchange capacity - $cmol_c/dm^3$ (Reatto *et al.* 2008).

In each area, ~ 30 adults, as defined above, were sampled randomly with a minimum of 30 m distance from each other. To compare effects on different size classes, additional ~ 70 young individuals with $D_{30cm} \leq 5$ cm (hereinafter referred to as juveniles), were sampled in the six populations that compose the three geographical pairwise units (Figure 1). Adult individuals were georeferenced. Fresh leaves for genetic analysis were collected from the 1,037 sampled individuals.

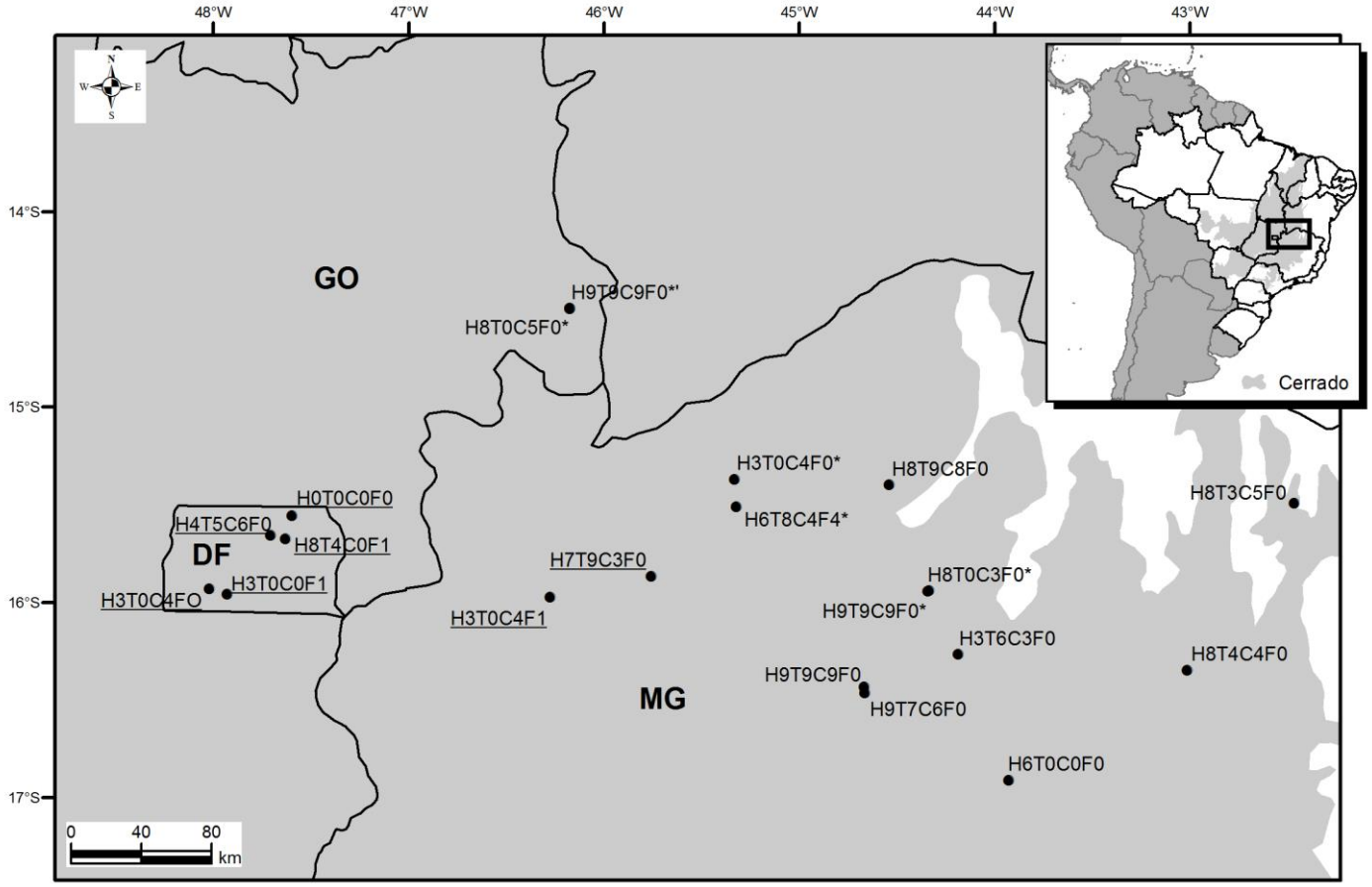


Figure 1. Location of the 20 *C. brasiliense* populations in the Cerrado. Due to the scale used, some points represent more than one population. Areas are named after fruit harvesting (H), vegetation thinning (T) and cattle ranching (C) indexes and fire frequency (F). Indexes summarize severity, duration and area of disturbances. ` differentiates populations with equal indexes and fire frequency. * discriminates populations that compose the three geographical pairwise units. Underlining of population names segregates populations in two groups based on Bayesian cluster analysis, assuming K = 2, as presented in Supplementary Material 1.

Laboratory analyses

Laboratory analyses were performed in *Laboratório de Genética Vegetal* in *Empresa Brasileira de Pesquisa Agropecuária - Recursos Genéticos e Biotecnologia (Embrapa - Cenargen).* Total DNA was extracted from leaf samples using the CTAB protocol (Doyle & Doyle 1987). Extracted DNA was quantified with NanoDrop 2000 (Thermo Scientific).

Microsatellite fragments were amplified with 9 previously developed primers for *C. brasiliense* (Collevatti *et al.* 1999). PCR reactions had a total volume of 8 µl and were composed of 1X PCA buffer (10 mM of Tris-HCl; 8.3 pH; 50 mM of KCl); 0.25 mM of dNTP; 10 ηg of DNA; 0.25 µM of reverse fluorescent-labeled primer; 0.25 µM of forward primer; 1 U of Taq DNA Polimerase; 0.25 mg/mL of BSA and ultrapure water. The PCR protocol consisted of a prior initial denaturation at 95°C for 5 min; 30 cycles composed of 95°C for 1 min, annealing temperature for 1 min and 72°C for 1 min; and a final extension step at 72°C for 20 min. Annealing temperatures followed Collevatti *et al.* (1999).

Fluorescent peaks of the amplified fragments were detected through automated DNA analyzer ABI 3730 (Applied Biosystems), using for this a mix with 1 µl of the amplicon, 1 µl of an internal size standard (ROX – synthetized in lab), and 18 µl of formamide. The peaks were then genotyped with Genemapper 4.1 software (Applied Biosystems).

Data analyses

Alleles were rounded off with the use of AlleloBin software (Prasanth *et al.* 2006). Test for null alleles was performed through Micro-Checker software (Oosterhout *et al.* 2004), using for this a 95% confidence interval and 1,000 Monte Carlo simulations. Test for linkage disequilibrium was performed through Fisher exact test (Weir 1996a) with GDA (Genetic Data Analysis) software (Weir 1996b). The number of alleles per locus, expected and observed heterozygosity and the inbreeding coefficient of adult and juvenile populations were estimated according to Weir (1996a), also with the GDA software. Since adult and juvenile populations sample sizes varied significantly, we estimated allelic richness through rarefaction (Leberg 2008), using for this the minimal number of genotyped individuals per population times two, as *C. brasiliense* is diploid. Genetic divergence between

pairwise adult and juvenile populations was estimated according to Weir & Cockerham (1984) with 10,000 bootstraps. Both allelic richness and genetic divergence were calculated through the hierfstat package (Goudet 2005) in R (R Core Team 2016).

In order to test for occurrence of recent decrease in effective sizes of adult and juvenile populations, we used the Bottleneck software (Piry *et al.* 1999). This software is based on the fact that populations that have gone through recent bottlenecks present a slower decrease of expected heterozygosity than allele number and, consequently, than heterozygosity under mutation-drift equilibrium, since this is calculated from the allele number. We used the TPM (two-phase model) with 95% of SMM (stepwise mutation model) and 5% of IAM (infinite allele model), a variance among IAM steps of 12 and the Wilcoxon test with 1,000 iterations, as recommended by Piry *et al.* (1999) for microsatellite studies with less than 20 loci.

The spatial genetic structure (SGS) of individual adult populations and of all adult populations combined were analyzed through the kinship coefficient (Loiselle *et al.* 1995) for pairwise individuals in distance classes for the loci set. We used eight and 20 distance classes for individual and combined populations, respectively. We opted for using equal numbers of pairs of individuals per distance class for each analysis, which results in different sizes of distance classes but controls the bias associated with different number of pairs per class. The standard error and 95% probability confidence intervals of kinship coefficient estimates were calculated by 10,000 jackknife resampling of individuals among loci. To test for absence of SGS, we constructed a region of acceptance of the null hypothesis based on 10,000 Monte Carlo permutations of individuals among distance classes. All estimates were calculated through the SPAGeDI (Spatial Pattern Analysis of Genetic Diversity) software (Hardy & Vekemans 2002).

In order to improve our understanding of the adult populations delimitation, we performed a Bayesian cluster analysis, allocating adult individuals to gene pools. To do this, we used the Structure software (Pritchard *et al.* 2000), which uses a Bayesian method and a MCMC (Markov Chain Monte Carlo) procedure to probabilistically designate individuals to one or more populations. We used a total run length of 1,000,000 iterations, with a burn-in length of 250,000 iterations, the ancestry model that allows for genome mixture and the frequency model in which allele frequencies are correlated between populations (Pritchard *et al.* 2010).

Simulations were carried out with a K (number of gene pools) interval of one to 23, and 20 independent repetitions (Evanno *et al.* 2005). To choose the most likely number of K, we used the statistic method described by Evanno *et al.* (2005), which enables the selection of the most verisimilar and homogeneous groups of individuals.

To identify if soil, geographic, demographic and land use and management variables are associated with the genetic diversity of adult populations, we used generalized linear models (GLMs) (Lindsey 1997). We used allelic richness and expected heterozygosity as genetic diversity response variables in distinct analysis. Four sets of explanatory variables were initially considered: soil (clay, sand and silt content, pH, cation exchange capacity, base and aluminum saturation), geographic (latitude and altitude), demographic (adult density and adult basal area) and land use and management (cattle ranching, fruit harvesting and vegetation thinning indexes and fire frequency). To detect collinearity between variables we used Pearson correlation analysis inside sets and removed the most difficult variables to interpret when paired variables were collinear ($r \geq 0.7$). All variables cited above were used as explanatory variables in the initial models, except for soil variables, which were summarized by sand content, pH and base saturation (Supplementary Material 2).

Latitude and altitude were selected because there is evidence of latitudinal and altitudinal gradients of genetic diversity (Martin & McKay 2004; Adams & Hadly 2013; Vellend & Geber 2005; Ohsawa & Ide 2008; Thiel-Egenter *et al.* 2009; Yan *et al.* 2009; Shen *et al.* 2014). Soil sand content, pH and percent base saturation summarize soil conditions (Reatto *et al.* 2008). Fruit harvesting, cattle ranching and vegetation thinning indexes and fire frequency summarize land use and management; fire frequency is also one of the major determinants of the Cerrado (Miranda *et al.* 2002). Adult density was selected since it can affect the dynamics of pollination (Loveless & Hamrick 1984; Ward *et al.* 2005; Dick *et al.* 2008) and inbreeding in populations of the species (Collevatti & Hay 2011). Adult basal area was selected since it is related to fruit production of *C. brasiliense* (Oliveira & Scariot 2010), which might affect genetic diversity.

We used the gaussian distribution as the variance function, which resulted in good fit, and identity as the link function. In order to select the models that best fit the data for each response variable, we performed a stepwise model selection by Akaike Information Criterion (AIC) (Akaike 1998), with backward direction and a maximum of

1,000 steps. The selected models were used as global models and were tested against null models through Chi-square test. We then performed the same model selection for multiple combinations of explanatory variables from the global models in order to identify other possible good models. We tested 52 and 24 combinations for the allelic richness and expected heterozygosity, respectively. The selected models were tested against the global models through Chi-square test. The models with close AIC values to the global models were considered for further analyses.

We performed an Analysis of Variance (ANOVA) with the genetic parameters of the adult and juvenile populations that constitute the three geographical pairwise units (Figure 1). For these analyses, we used a factorial design with randomized blocks. Normality of data was confirmed through the Shapiro-Wilk test. Two analyses were carried out, each one with one of the following genetic diversity parameters as response variable: allelic richness and expected heterozygosity. The blocks, which correspond to the three geographical regions of the pairwise units (1 = H3T0C4F0*/H6T8C4F4*, 2 = H8T0C5F0*/H9T9C9F0*` and 3 = H8T0C3F0*/H9T9C9F0*), were used in order to control for history and geography on the genetic diversity estimates. Land use (Cerrado and pastureland) and size class (adult and juvenile) were used as factors, with two levels each. GLMs and ANOVA were carried out with the vegan (Oksanen *et al.* 2008), boot (Davison & Hinkley 1997) and MASS (Venables & Ripley 2002) packages in R (R Core Team 2016).

3. Results

Genetic diversity and spatial structure

Land use and management indexes and fire frequencies are presented in Figure 1; demographic and soil parameters are presented in Supplementary Material 3. No locus presented evidence for null alleles in more than 55% of populations analyzed. Since this analysis is based on Hardy-Weinberg equilibrium, these results could indicate that null allele detection was actually enabled by homozygosity excess in specific populations and, consequently, that null allele frequencies were overestimated. Linkage disequilibrium was observed in a higher proportion of populations (80%) for only one pair of loci, which we opted to maintain in the analyses.

Allelic richness varied from 8.22 to 12.33, with a mean of 10.02; expected

heterozygosity varied from from 0.70 to 0.84, with a mean of 0.77; observed heterozygosity varied from 0.65 to 0.84, with a mean of 0.75; all except one inbreeding coefficients were not statistically different from zero (Supplementary Material 4). The inbreeding coefficient for the entire set of populations also did not statistically differ from zero (-0.0287; confidence interval: -0.0336 to 0.0867). Genetic divergence between pairs of populations varied from 0.0000 to 0.1581, with an average of 0.0678 (Supplementary Material 5). Genetic divergence for the entire set of populations was 0.0745, with a confidence interval of 0.0463 to 0.1174.

Bottleneck analysis showed that no population, adult or juvenile, went through a recent genetic bottleneck (Supplementary Material 6). SGS was essentially absent at the intrapopulation level, but was significant for all populations combined, with a maximum kinship coefficient of 0.0998 and SGS for up to 416.6 Km (Supplementary Material 7). As for the Bayesian cluster analysis, the optimal K according to Evanno *et al.* (2005) was two, which divided the combined populations in two geographically distinct clusters: East (H6T0C0F0, H8T4C4F0, H8T3C5F0, H3T6C3F0, H9T7C6F0, H9T9C9F0, H9T9C9F0*, H8T0C3F0*, H8T9C8F0, H3T0C4F0*, H6T8C4F4*, H8T0C5F0*, H9T9C9F0*`) and West (H0T0C0F0, H8T4C0F1, H4T5C6F0, H3T0C0F1, H3T0C4F0, H3T0C4F1, H7T9C3F0) (Supplementary Material 1, Figure 1).

Associations of geography, soil, demography and land use and management with genetic diversity

Based on AIC, we selected eight models for allelic richness and expected heterozygosity of 20 *C. brasiliense* adult populations, respectively (Table 1). Both global models were statistically different from null models (Table 2). The global model for allelic richness was built with seven explanatory variables: cattle ranching index, vegetation thinning index, fire frequency index, adult density, adult basal area, latitude and soil sand content, of which all but adult basal area were statistically significant (Table 2). The expected heterozygosity global model was built with eight explanatory variables, five of which were significant (cattle ranching index, vegetation thinning index, adult density, latitude and soil sand content), and three of which were not (fruit harvesting index, fire frequency and altitude) (Table 2).

Table 1. Best models and global models for genetic diversity (Ar = allelic richness, He = expected heterozygosity) of 20 *C. brasiliense* adult populations. N is the number of variables. Global models are the first models (1).

Response variable	Model number	Model	N	AIC
Ar	Ar(1)	~ CAT** + THIN* + FIR* + DENS* + ABA + LAT*** + SAND***	9	36.474
	Ar(2)	~ CAT* + THIN* + FIR* + DENS + LAT*** + SAND***	8	38.424
	Ar(3)	~ CAT* + THIN* + FIR* + LAT*** + SAND***	7	40.815*
	Ar(4)	~ CAT* + FIR + DENS* + ABA + LAT*** + SAND***	8	41.364*
	Ar(5)	~ CAT + FIR + DENS + LAT*** + SAND***	7	43.203*
	Ar(6)	~ CAT + DENS + ABA + LAT*** + SAND**	7	43.669*
	Ar(7)	~ CAT + DENS + LAT*** + SAND**	6	44.401**
	Ar(8)	~ LAT*** + SAND***	4	44.643**
He	He(1)	~ CAT* + FRU + THIN* + FIR + DENS* + LAT** + ALT + SAND*	10	-83.452
	He(2)	~ CAT + FRU + THIN* + FIR + DENS* + LAT** + SAND	9	-81.881
	He(3)	~ CAT* + THIN* + FIR* + DENS* + LAT** + ALT + SAND**	9	-81.493
	He(4)	~ CAT + FRU + THIN + DENS + LAT** + SAND	8	-80.434
	He(5)	~ FRU* + THIN + DENS + LAT* + SAND	7	-80.159
	He(6)	~ CAT + FRU* + THIN + DENS + LAT*	7	-79.937
	He(7)	~ FRU* + THIN + DENS + LAT*	6	-79.540
	He(8)	~ FRU* + DENS* + LAT*	5	-79.239*

Variables: CAT = cattle ranching index; THIN = vegetation thinning index; FIR = fire frequency; FRU = fruit harvesting index; DENS = adult density/ha; ABA = adult basal area (cm^2/ha); LAT = latitude; ALT = altitude; SAND = soil sand content (g/Kg). Significance codes: *** = 0.001; ** = 0.01; * = 0.05. Codes after variables in models point to the significance of the variable for the model. Codes after AIC values point to statistically different models from the global model for each response variable.

Table 2. Global models estimates of explanatory variables and their respective standard deviation for genetic diversity (Ar = allelic richness; He = expected heterozygosity) of 20 adult populations.

Explanatory Variable	Response Variable	
	Ar (1) ***	He (1) ***
	β ± SD	β ± SD
Intercept	35.6400 ± 3.6380 ***	1.5570 ± 0.1823 ***
CAT	-0.2050 ± 0.0604 **	-0.0070 ± 0.0032 *
THIN	0.1034 ± 0.0466 *	0.0061 ± 0.0023 *
FIR	-0.4258 ± 0.1523 *	-0.0176 ± 0.0085
FRU		-0.0045 ± 0.0029
DENS	-0.0125 ± 0.0052 *	-0.0004 ± 0.0002 *
ABA	0.0000 ± 0.0000	
LAT	1.3450 ± 0.2224 ***	0.0374 ± 0.0112 **
ALT		-0.0001 ± 0.0000
SAND	-0.0059 ± 0.0011 ***	-0.0001 ± 0.0000 *

Variables: CAT = cattle ranching index; THIN = vegetation thinning index; FIR = fire frequency; FRU = fruit harvesting index; DENS = adult density/ha; ABA = adult basal area (cm^2/ha); LAT = latitude; ALT = altitude; SAND = soil sand content (g/Kg); Significance codes: *** = 0.001; ** = 0.01; * = 0.05. Codes after values for variables in models point to the significance of the variable for the model. Codes after model number point to statistically different models from null models (response variable ~ 1).

ANOVA indicated that the main effects of land use and size class were not statistically significant for none of the response variables. However, the effect of region (block) and the effect of interaction between land use and size class were both statistically significant for the allelic richness (p-value = 0.0048 and 0.0303,

respectively) (Figures 2 and 3). Allelic richness was higher for the H8T0C5F0*/H9T9C9F0*` pair (10.9880), followed by H3T0C4F0*/H6T8C4F4* (9.7119) and H8T0C3F0*/H9T9C9F0* (9.3602). While allelic richness was lower in pasturelands than areas with Cerrado vegetation for the juvenile size class, the opposite was true for the adult size class, which resulted in a interaction between land use and size class for allelic richness (Figures 2 and 3).

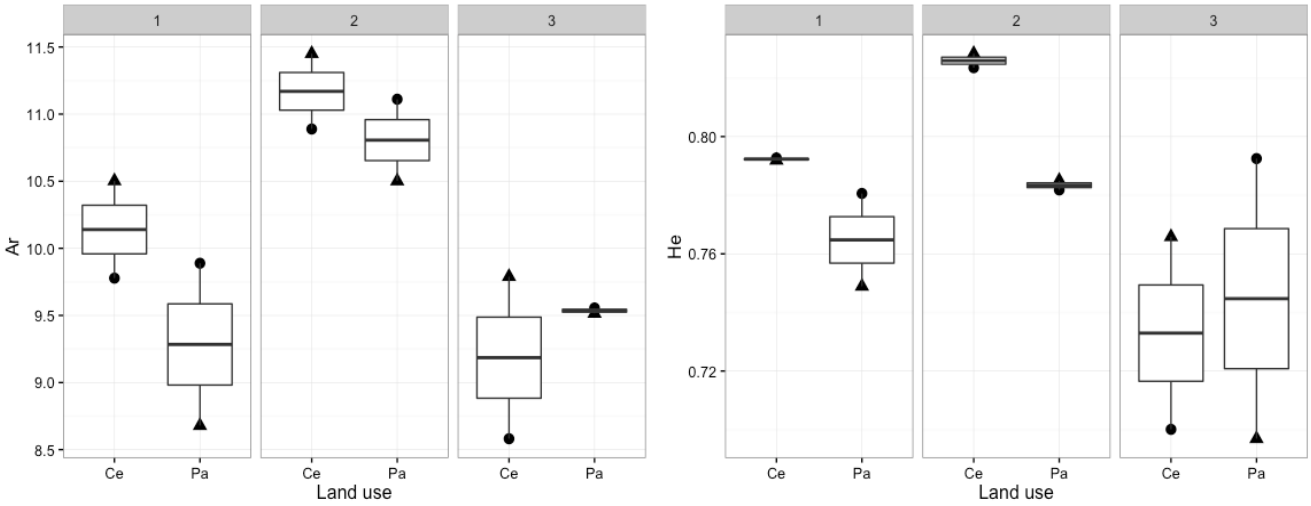


Figure 2. Boxplot of data used for ANOVA with a factorial design and randomized blocks for six populations in three geographical pairwise units. Ar = allelic richness, He = expected heterozygosity. Blocks = regions (1 = H3T0C4F0*/H6T8C4F4*, 2 = H8T0C5F0*/H9T9C9F0*` and 3 = H8T0C3F0*/H9T9C9F0*); factors = land use (Ce = Cerrado, Pa = Pasture) and size class (● = adult, ▲= juvenile).

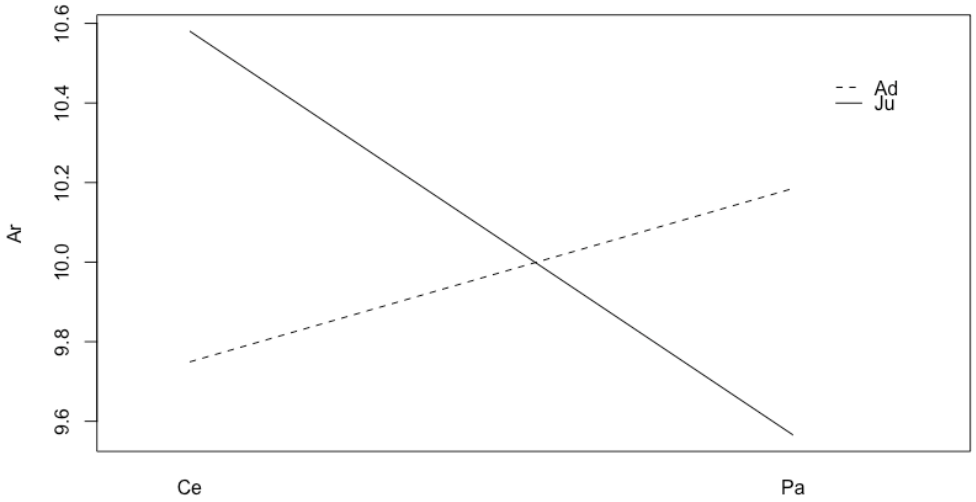


Figure 3. Allelic richness (Ar) interaction between land use (Ce = Cerrado, Pa = pasture) and size class (Ad = adult, Ju = juvenile) for six populations in three geographical pairwise units.

4. Discussion

Genetic diversity and spatial structure

Genetic diversity was similar to that encountered by other studies on the species (Collevatti *et al.* 2001b; c; Collevatti & Hay 2011). Allelic richness presented higher differentiation among populations than expected heterozygosity (coefficient of variation = 0.1148 and 0.0538 for Ar and He, respectively). Alleles are lost more rapidly than heterozygosity in face of disturbances, mainly because rare alleles are especially susceptible to loss (Allendorf & Luikart 2007), which would permit higher differentiation of allele number estimates among populations in different conditions of stress. We believe that such differentiation allowed for higher detection of associations for the allelic richness global model in comparison to the expected heterozygosity global model. Rapid loss of alleles in comparison to heterozygosity is also postulated by the theory on which recent bottleneck analysis is based on (Cornuet & Luikart 1996); however no population presented evidence for having gone through a recent reduction in effective size.

Our results concerning inbreeding coefficients indicate that there are no evidences of non-random reproduction inside populations and, thus, that pollination probably occurred without spatial constraint. Estimates of genetic divergence between pairwise populations and for all populations are considered to be low to moderate (Hartl & Clark 1997), which points to historical occurrence of moderate levels of gene flow among populations. The absence of significant SGS for individual populations could be related to the lack of sampled trees that are apart by less than 30 m, a distance that might have presented SGS, considering the demographic clustering presented by the species (Collevatti & Hay 2011). The existence of SGS for all populations combined, in turn, indicates that *C. brasiliense* presents higher genetic differentiation with increasing distances when considering larger scales,

In spite of our results indicating past existence of gene flow in moderate distances, other studies have encountered evidence of restricted gene flow for populations of *C. brasiliense* (Collevatti *et al.* 2001a, 2010; Collevatti & Hay 2011), associating it to the seed dispersal and pollination syndromes of the species, as well as its demographic clustering pattern. Although some animals can disperse fruits of this species, these, which are large and heavy, are dispersed mainly by gravity (Gribel 1986), leading to commonly found concentrated aggregates of individuals. *C.*

brasiliense is pollinated mainly by bats, specially *Glossophaga soricina*, which, in face of higher concentration of flowers, presents territorial foraging behavior, diminishing flight range (Gribel & Hay 1993). Although *G. soricina* does present territorial foraging behavior, it can also forage several plants along regularly used routes when facing diminished resources (Lemke 1984), which would allow for larger gene flow distances. Since a considerable portion of the populations studied here are present in pasturelands, presenting lower densities and, thus, lower availability of resources for pollinators, it seems reasonable that pollinators might have foraged plants along larger distances, increasing gene flow.

Associations of geography, soil, demography and land use and management with genetic diversity

Considering global models, soil sand content presented negative and significant associations with allelic richness and expected heterozygosity. Soil texture is a very important element for the survival, growth and reproduction of plants: soils with greater proportions of sand drain more rapidly and do not present a strong electrochemical charge on their surface, not holding water or adsorbing nutrient cations satisfactorily, which can lead to leaching of nutrients (Gurevitch *et al.* 2006). Although *C. brasiliense* is adapted to sandy and nutrient-poor soils, less fit individuals can underperform or even die in extreme situations (Leite *et al.* 2012), which could lead to loss of rare alleles and decrease expected heterozygosity consequently. This demonstrates how environmental conditions can constrain genetic diversity.

There were positive and significant associations of latitude with allelic richness and expected heterozygosity, meaning that areas closer to the Equator line present populations of *C. brasiliense* with higher genetic diversity. Such importance of geographical variables for the explanation of genetic diversity is corroborated by the statistical significance of region for the ANOVA based on the six populations in the three geographical pairwise units. The increase of allelic richness with latitude can be observed for the average of each region in Figure 1: it was higher for H8T0C5F0*/H9T9C9F0*` pair (10.9880), followed by H3T0C4F0*/H6T8C4F4* (9.7119) and H8T0C3F0*/H9T9C9F0* (9.3602). In addition, Bayesian clustering results also points to the importance of geography in differentiating gene pools.

Although our data is limited, it is possible that genetic diversity of *C. brasiliense* responds to a latitudinal gradient in a similar way that species diversity does. This is corroborated by two models of distribution of genetic diversity that most likely act simultaneously: the latitudinal model (Martin & McKay 2004; Adams & Hadly 2013) and the species diversity model (Vellend & Geber 2005), as classified by Schrey *et al.* (2011a). Two important processes that explain the latitudinal model are: increased evolutionary rates due to higher mutation rates and selection in lower latitudes; more intense historical distributions changes in higher latitudes due to climatic changes in the last glaciation period (Adams & Hadly 2013). Studies show that northern as well as southern portions of Brazil acted as refuge to *C. brasiliense* during the last glaciation period (Collevatti *et al.* 2002, 2012). This shows that historical range shifts from this period most likely do not explain the latitudinal gradient found for genetic diversity of the species, which might be explained by higher mutation rates and selection and by local, ecological and demographic processes that act together on genetic and species diversity. Another possible and non-exclusive explanation is related to the high fragmentation and increased economical activities present towards the south of the Cerrado, as showed by Diniz-Filho *et al.* (2009). These authors presented evidence for latitudinal gradient of the inbreeding coefficient for the species, with higher inbreeding towards the south of the Cerrado, corroborating with our results.

Cattle ranching index presented a negative and significant association with allelic richness and expected heterozygosity. Grazing and tramping by cattle can influence plant growth and lead to the death of individuals, which, in turn, can lead to the loss of rare alleles and, consequently, decrease of expected heterozygosity. Tramping by animals can also change the physical-chemical properties of the soil, which can lead to modifications in the vegetation (Shan *et al.* 2006). Giroldo & Scariot (2015) encountered negative associations of cattle ranching with the demography of the same populations we studied here, which corroborates with our findings.

Fire also presented a negative and significant association with allelic richness. Although *C. brasiliense* presents morphological and phylogenetic evidences for adaptation to fire (Simon & Pennington 2012), fire events can kill individuals of this species (Medeiros *et al.* 2008), which might result in loss of rare alleles. The response of genetic diversity to fire occurrence can be complicated and species-

specific (Schrey *et al.* 2011b; Smith *et al.* 2014); however other papers have encountered similar results to ours (Uchiyama *et al.* 2006; Schrey *et al.* 2010, 2011a; Smith *et al.* 2014). Whether the death of *C. brasiliense* individuals due to cattle ranching and fire occurrence can be attributed to chance, selection pressure for more fire-tolerant and grazing-tolerant genotypes (Mengli *et al.* 2005) or a synergetic combination of both, which seems more likely, is an issue to be further investigated.

There were negative and significant associations of adult *C. brasiliense* density with allelic richness and expected heterozygosity. Lower densities can lead to an increase of pollination distance, as a consequence of pollinator behavior modification in face of diminished resources, resulting in increased gene flow, which can lead to addition of new alleles, increasing expected heterozygosity as an outcome (Loveless & Hamrick 1984; Ward *et al.* 2005; Dick *et al.* 2008). This could also explain the positive and significant associations between vegetation thinning and allelic richness and expected heterozygosity, since a consequence of vegetation thinning is the decrease of demographic density. However, thinning is applied only on small individuals of *C. brasiliense* and thus, due to time gap, it would be necessary to investigate such association with smaller size classes, which was not possible in this study due to limitation of sample size.

Fruit harvesting did not present a significant association with any of the response variables based on the global models. This is corroborated by Giroldo & Scariot (2015), who did not encounter associations between *C. brasiliense* fruit harvesting and demography of the same populations we analyzed here. In fact, *C. brasiliense* populations can persist through very high fruit harvesting rates, of up to 90 (Oliveira 2009) and 99% (Zardo 2008), not considering, though, harvesting effects on the local fauna. As *C. brasiliense* fruit harvesting does not involve damage to other parts of the trees, it seems reasonable that populations can withstand even high levels of this activity (Ticktin 2004). Other studies also did not detect negative impacts of NTFP harvesting on genetic diversity (Shaanker *et al.* 2004; Wang *et al.* 2013; Xu *et al.* 2013; Gaoue *et al.* 2014). Such consequences, however, will depend on life history, the part of the plant that is harvested, variation in environmental conditions and management practices (Ticktin 2004).

The fact that there was a significant interaction between land use and size class for allelic richness could indicate how a delay in response to the impact associated with pasturelands could be present in this case. Adult populations

presented a higher allelic richness in pasturelands than in areas with Cerrado, which might be interpreted as a lack of response to the impact due to a time lag, considering that juvenile populations did in fact present lower allelic richness in pasturelands, which we consider to be a consequence of such impact. Time lags in loss of genetic diversity due to environmental impacts are common for tree species. This is because of their: long lifespan, which allows for persistence; high levels of phenotypic plasticity, facilitating adaptation; and large size and high level of pollen and seed production, generating gene flow that will lessen the loss of genetic diversity (Vranckx *et al.* 2012).

GLM analysis was able to detect the association of cattle ranching with genetic diversity; however the ANOVA design was not capable to detect significant main effect of land use, which is directly associated to cattle ranching, although a tendency towards this was observed for the juvenile size class. We consider our GLM design to be more robust, since it was based on 11 continuous variables rather than three discrete ones, which signifies an improved representation of ecological reality, in all its complexity. In addition, our GLM design consisted of 20 sample units, while the ANOVA design consisted of six. Thus, we consider that our GLM results are more representative of reality and more reliable.

Implications for conservation and management

Our SGS results imply that it would not be necessary to define minimal distances between sampled trees to delineate seed collection strategies for conservation or restoration projects; however, since we did not sample pairs of trees that are less than 30 m apart and, thus, we cannot infer the SGS at such distances, concentrated aggregates of *C. brasiliense* individuals should be sampled with caution. Based on SGS for all populations combined, genetic diversity would be higher if sampled populations were at least 416 Km apart. Considering Bayesian clustering results, germplasm should be collected from both geographical clusters, East and West, as presented in Figure 1, in order to maximize diversity.

Cattle ranching is one of the most common activity in the Cerrado and our results further supports conservation urges to stop deforestation in this region to implement pasturelands, as there are many unproductive, abandoned or degraded areas available for this (Klink & Moreira 2002). In order to allow for survival of

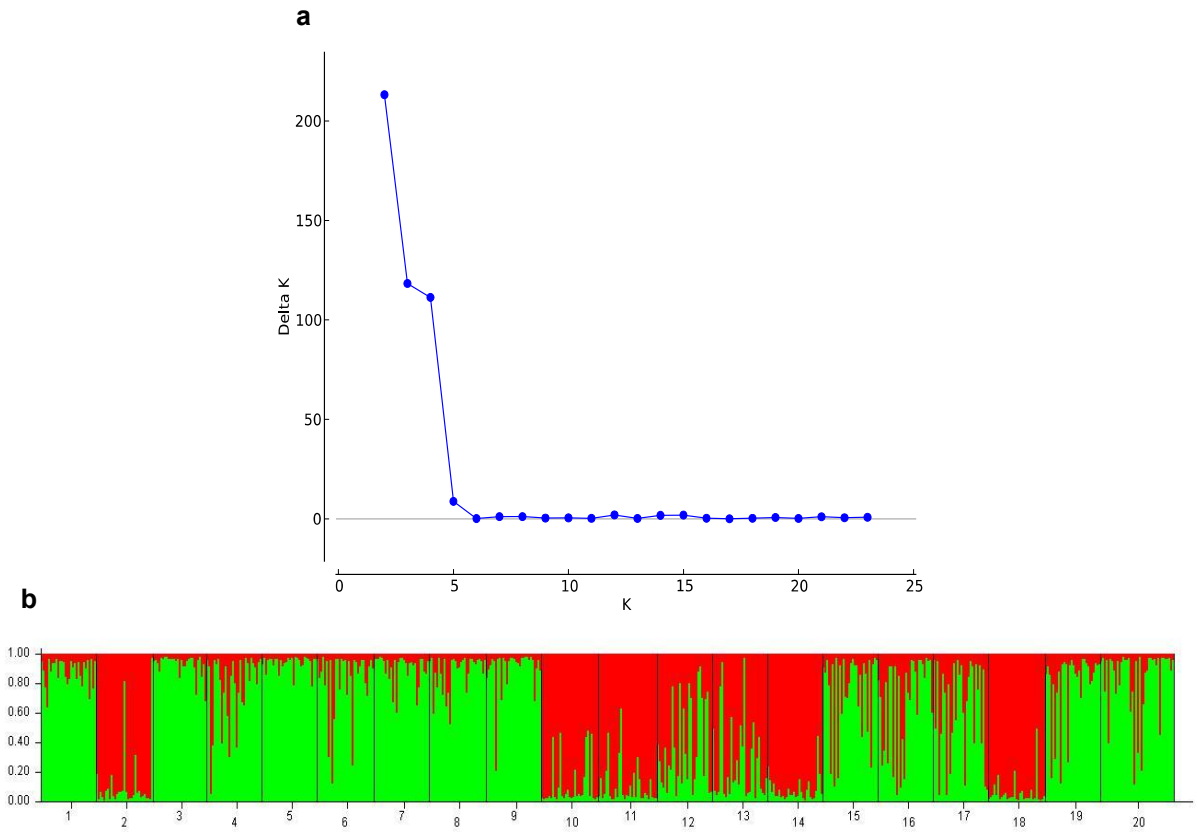
seedlings of *C. brasiliense* and their development into adulthood, we suggest increasing the rotation period between pasture areas and intervals between thinning events, as proposed by Giroldo and Scariot (2015), as well as decreasing thinning intensity and fencing target individuals from grazing and tramping. These strategies would not only improve the conservation of the species, but also provide with additional income through harvesting and selling of its fruits and provide for shadow for the cattle in the future, reducing heat stress levels and productivity losses (Blackshaw & Blackshaw 1994).

Integrating multiple land uses has been advocated as a conservation strategy in landscapes transformed by anthropogenic activity (Moilanen *et al.* 2005). In this sense, different uses should be combined to allow equally for cattle grazing, sustainable use of natural resources and conservation of biodiversity in implemented pasturelands, such as harvesting of NTFP. In fact, our results further supports the sustainability of *C. brasiliense* fruit harvesting (Giroldo & Scariot 2015), which should continue to be promoted as an activity that supports the conservation of the Cerrado and improves social conditions of poor traditional communities, being an important component of the livelihoods of populations. However, bearing in mind the increasing human pressure on natural resources and the possibility of future decrease in production of *C. brasiliense* fruits in face of global climate changes (Nabout *et al.* 2011; Collevatti *et al.* 2011), technical management recommendations for sustainable harvesting should be applied (Oliveira & Scariot 2010).

Although fire occurrence is a common feature of the Cerrado, the majority of fire incidents today are anthropogenic (Miranda *et al.* 2002). We believe our results on fire impact on genetic diversity of populations of *C. brasiliense* further demonstrate the need of fire management in conservation programs. Although controlled fires are not well accepted due to its misuse in agriculture and anti-fire policies predominate in Brazilian protected areas, prescribed burns should be accepted as an appropriate tool for management of fire in the Cerrado, as they are in important ecological reserves worldwide (Pivello & Norton 1996; Ramos-Neto & Pivello 2000). Fire management allows for control of fire frequency and intensity, preventing its most devastating effects due to fire exclusion that lead to uncontrollable incidents, which increases mortality rates of plant and animal species and might lead to competitive exclusion of fire-adapted plants (Pivello & Norton 1996; Ramos-Neto & Pivello 2000). For all these suggested actions to occur in a

significant manner, however, public policies should be developed to promote them through financial and tax incentives.

Supplementary Materials



Supplementary material 1. Bayesian cluster analysis results: a) Selection of optimal number of clusters (K), based on the method of Evanno *et al.* (2005); b) Clustering of 607 *C. brasiliense* adult individuals, collected from 20 *a priori* populations, assuming K = 2. Population identification: 1 = H6T0C0F0, 2 = H0T0C0F0, 3 = H8T4C4F0, 4 = H3T0C4F0*, 5 = H8T3C5F0, 6 = H9T7C6F0, 7 = H8T9C8F0, 8 = H9T9C9F0, 9 = H9T9C9F0*, 10 = H3T0C0F1, 11 = H4T5C6F0, 12 = H7T9C3F0, 13 = H3T0C4F1, 14 = H8T4C0F1, 15 = H3T6C3F0, 16 = H8T0C5F0*, 17 = H9T9C9F0*`, 18 = H3T0C4F0, 19 = H6T8C4F4*, 20 = H8T0C3F0*.

Supplementary Material 2. Pearson correlation for pairwise soil parameters of 20 studied areas.

	Sa	Si	pH	CEC	BS	AS
C	-0.9787***	0.7178***	-0.2379	0.7031***	-0.4440*	0.4111
Sa	-	-0.8454***	0.2615	-0.7694***	0.5096*	-0.4167
Si		-	-0.2679	0.7806***	-0.5734**	0.3437
pH			-	-0.6408***	0.3167	-0.4548*
CEC				-	-0.5998**	0.4748*
BS					-	-0.7322***

C = clay content (g/Kg); Sa = sand content (g/Kg); Si = silt content (g/Kg); CEC = cation exchange capacity ($cmol_c/dm^3$), BS = base saturation (%); AS = aluminum saturation (%). Significance codes: *** = 0.005; ** = 0.01; * = 0.05. Codes points to the significance of correlation between variables.

Supplementary Material 3. Location, physiognomy, demographic and soil parameters of 20 studied areas with populations.

Population	Longitude	Latitude	Altitude (m)	Physiognomy	AD	ABA	Sa	pH	BS
H6T0C0F0	-43.930	-16.911	1017.0	DC	103.00	28168.67	750	4.6	34
H0T0C0F0	-47.599	-15.555	1039.6	TC	37.33	10309.78	525	4.6	13
H8T4C4F0	-43.015	-16.346	893.8	TC	199.00	41340.66	500	4.3	8
H3T0C4F0*	-45.332	-15.370	628.7	DC	38.00	23048.01	825	4.8	22
H8T3C5F0	-42.468	-15.492	987.2	TC	39.14	21519.08	800	4.2	17
H9T7C6F0	-44.668	-16.462	730.7	DC	45.33	41136.15	800	4.9	19
H8T9C8F0	-44.541	-15.399	623.6	PA	12.00	16199.84	800	4.9	25
H9T9C9F0	-44.670	-16.434	734.4	PA	29.67	25498.89	800	4.8	23
H9T9C9F0*	-44.338	-15.941	812.5	PA	13.67	14069.13	775	4.6	28
H3T0C0F1	-47.931	-15.958	1131.6	TC	32.33	7945.21	400	4.8	16
H4T5C6F0	-47.708	-15.657	1012.2	DC	23.00	10782.67	500	4.3	12
H7T9C3F0	-45.758	-15.865	560.0	PA	42.00	25072.10	825	4.5	29
H3T0C4F1	-46.278	-15.972	575.5	DC	57.33	16759.49	325	5.0	22
H8T4C0F1	-47.632	-15.676	989.4	SC	10.50	4617.55	475	4.6	11
H3T6C3F0	-44.189	-16.267	786.9	DC	96.00	38391.04	800	4.8	20
H8T0C5F0*	-46.178	-14.496	702.8	TC	46.67	20715.47	775	4.9	14
H9T9C9F0*`	-46.176	-14.494	704.0	PA	54.29	44062.79	825	5.0	16
H3T0C4F0	-48.023	-15.932	1210.9	TC	29.33	4930.97	350	4.3	7
H6T8C4F4*	-45.325	-15.511	638.0	PA	25.33	12343.59	550	4.7	25
H8T0C3F0*	-44.343	-15.942	807.7	DC	54.22	29173.08	800	4.8	18

AD = adult density/ha; ABA = adult basal area (cm^2/ha); Sa = soil sand content (g/Kg); BS = base saturation (%); DC = dense Cerrado; TC = typical Cerrado; PA = pastureland; SC = sparse Cerrado.

Supplementary Material 4. Genetic parameters for the adult and juvenile *C. brasiliense* populations.

	Population	n	A	Ar	He	Ho	f (95% CI)
Adult	H6T0C0F0	29.0	7.4444	8.2222	0.7000	0.6918	0.0118 (-0.0809 to 0.1074)
	H0T0C0F0	27.8	10.7778	11.5556	0.8141	0.8454	-0.0391 (-0.0895 to 0.0194)
	H8T4C4F0	28.0	8.1111	9.0000	0.7013	0.7211	-0.0288 (-0.1013 to 0.0531)
	H3T0C4F0*	28.4	8.8889	9.7778	0.7927	0.7510	0.0535 (-0.0190 to 0.1316)
	H8T3C5F0	28.0	8.5556	9.5556	0.7373	0.7416	-0.0060 (-0.0841 to 0.0716)
	H9T7C6F0	26.4	8.3333	9.1534	0.7515	0.7485	0.0037 (-0.0858 to 0.1012)
	H8T9C8F0	26.2	8.3333	9.1111	0.7809	0.7197	0.0797 (-0.0427 to 0.2055)
	H9T9C9F0	27.1	7.5556	8.2718	0.7298	0.6859	0.0614 (-0.0243 to 0.1529)
	H9T9C9F0*	28.1	8.7778	9.5556	0.7925	0.8191	-0.0342 (-0.1532 to 0.0466)
	H3T0C0F1	28.6	9.8889	10.6118	0.7926	0.7911	0.0020 (-0.0844 to 0.1106)
	H4T5C6F0	30.8	10.7778	11.2575	0.8091	0.7730	0.0452 (-0.0491 to 0.1401)
	H7T9C3F0	28.0	8.8889	9.5556	0.8027	0.7629	0.0501 (-0.0834 to 0.2079)
	H3T0C4F1	29.2	10.3333	11.0000	0.7994	0.7736	0.0328 (-0.0562 to 0.1556)
	H8T4C0F1	28.1	11.5556	12.3333	0.8394	0.7864	0.0642 (-0.0236 to 0.1801)
	H3T6C3F0	24.1	8.7778	9.7778	0.7966	0.7460	0.0647 (-0.0492 to 0.1475)
	H8T0C5F0*	27.9	10.1111	10.8889	0.8235	0.7915	0.0395 (-0.0253 to 0.1047)
	H9T9C9F0*`	28.4	10.3333	11.1111	0.7818	0.7598	0.0285 (-0.0731 to 0.1453)
	H3T0C4F0	30.1	10.3333	10.9270	0.7947	0.7822	0.0159 (-0.0960 to 0.1225)
	H6T8C4F4*	25.9	9.1111	9.8889	0.7806	0.7598	0.0269 (-0.0856 to 0.1364)
	H8T0C3F0*	36.4	8.2222	8.5811	0.7001	0.7307	-0.0443 (-0.1224 to 0.0482)
Juvenile	H3T0C4F0*J	64.4	11.3333	10.5021	0.7919	0.7111	0.1026* (0.0030 to 0.2159)
	H9T9C9F0*J	61.1	10.7778	9.5148	0.6969	0.6465	0.0728 (-0.0169 to 0.1586)
	H8T0C5F0*J	67.2	12.8889	11.4506	0.8283	0.7881	0.0488 (-0.0030 to 0.1025)
	H9T9C9F0*`J	66.1	11.4444	10.5013	0.7850	0.7748	0.0129 (-0.0830 to 0.1235)
	H6T8C4F4*J	66.4	9.5556	8.6787	0.7489	0.7366	0.0165 (-0.0592 to 0.0939)
	H8T0C3F0*J	56.4	10.1111	9.7892	0.7658	0.7553	0.0135 (-0.1271 to 0.1385)
	Average	36.5	9.6624	10.0220	0.7745	0.7536	0.0267

J at the end of population names differentiates juvenile populations. n = sample size; A = average allele number per locus; Ar = average allelic richness per locus; He = expected heterozygosity; Ho = observed heterozygosity; f = inbreeding coefficient. * Significantly different from zero based on 95% confidence intervals estimated by 10,000 bootstraps.

Supplementary Material 5. *Fst* (genetic divergence) estimates for pairwise adult and juvenile populations.

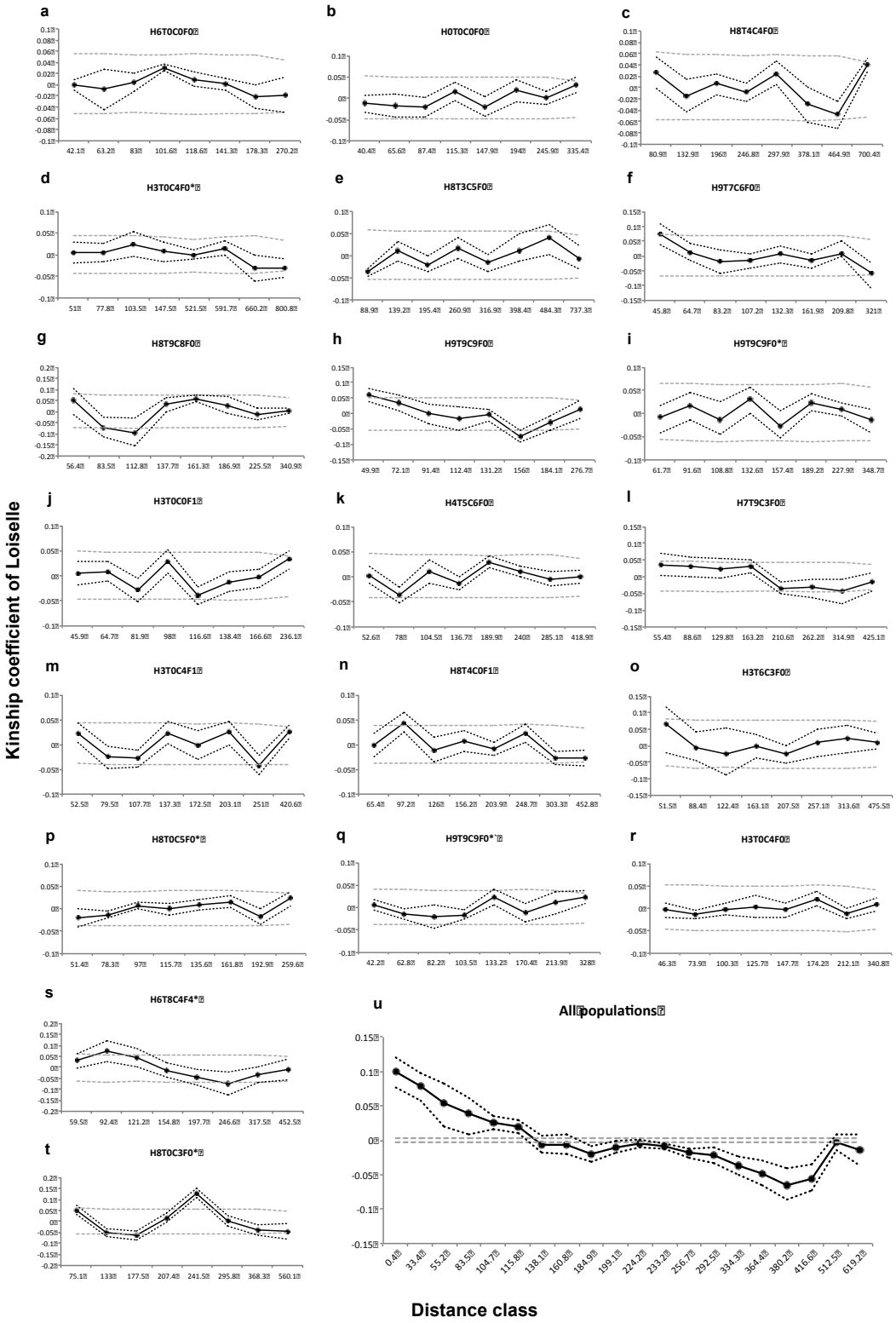
	H0T0C0F0	H8T4C4F0	H3T0C4F0*	H3T0C4F0*J	H8T3C5F0	H9T7C6F0	H8T9C8F0	H9T9C9F0	H9T9C9F0*	H9T9C9F0*J	H3T0C0F1	H4T5C6F0	H7T9C3F0	H3T0C4F1	H8T4C0F1	H3T6C3F0	H8T0C5F0*	H8T0C5F0*J	H9T9C9F0*`	H9T9C9F0*`J	H3T0C4F0	H6T8C4F4*	H6T8C4F4*J	H8T0C3F0*	H8T0C3F0*J
H6T0C0F0	0.0844	0.0488	0.1137	0.1095	0.0539	0.1581	0.1256	0.1422	0.0725	0.1400	0.0920	0.1037	0.0833	0.1070	0.1038	0.0862	0.1007	0.1032	0.1078	0.1041	0.1171	0.0993	0.1258	0.1072	0.0893
H0T0C0F0		0.1101	0.0518	0.0511	0.1019	0.1038	0.0781	0.0873	0.0702	0.0995	0.0131	0.0115	0.0318	0.0469	0.0261	0.0524	0.0489	0.0468	0.0541	0.0589	0.0351	0.0622	0.0733	0.0728	0.0613
H8T4C4F0			0.1277	0.1158	0.0240	0.1556	0.1238	0.1489	0.0654	0.1342	0.1126	0.1109	0.0964	0.1160	0.1098	0.0928	0.0990	0.1024	0.1149	0.1176	0.1377	0.1069	0.1403	0.1129	0.0923
H3T0C4F0*				0.0025	0.1135	0.0608	0.0285	0.0576	0.0635	0.0523	0.0532	0.0515	0.0493	0.0357	0.0409	0.0525	0.0419	0.0384	0.0482	0.0599	0.0659	0.0482	0.0545	0.0433	0.0366
H3T0C4F0*J					0.1051	0.0580	0.0284	0.0508	0.0570	0.0750	0.0563	0.0545	0.0495	0.0515	0.0525	0.0444	0.0459	0.0464	0.0609	0.0672	0.0747	0.0409	0.0589	0.0541	0.0348
H8T3C5F0						0.1381	0.1043	0.1376	0.0631	0.1231	0.1043	0.1068	0.0851	0.1100	0.0959	0.0924	0.0848	0.0898	0.0968	0.1007	0.1266	0.0926	0.1234	0.0999	0.0837
H9T7C6F0							0.0339	0.0181	0.0750	0.0798	0.0887	0.0940	0.0919	0.0889	0.0671	0.0519	0.0744	0.0728	0.0898	0.1018	0.1151	0.0540	0.0858	0.0845	0.0587
H8T9C8F0								0.0422	0.0532	0.0740	0.0746	0.0740	0.0664	0.0613	0.0604	0.0469	0.0479	0.0550	0.0669	0.0732	0.1006	0.0565	0.0790	0.0545	0.0438
H9T9C9F0									0.0705	0.0890	0.0975	0.0943	0.0708	0.0838	0.0660	0.0414	0.0603	0.0679	0.0806	0.0892	0.1196	0.0486	0.0710	0.0645	0.0455
H9T9C9F0*										0.0781	0.0712	0.0743	0.0543	0.0667	0.0655	0.0212	0.0590	0.0589	0.0742	0.0837	0.0935	0.0505	0.0801	0.0458	0.0237
H9T9C9F0*J											0.0937	0.0947	0.0704	0.0600	0.0615	0.0571	0.0650	0.0590	0.0614	0.0873	0.1127	0.0639	0.0799	0.0392	0.0501
H3T0C0F1												0.0073	0.0469	0.0544	0.0308	0.0541	0.0677	0.0575	0.0719	0.0834	0.0140	0.0721	0.0866	0.0865	0.0709
H4T5C6F0													0.0531	0.0489	0.0213	0.0637	0.0624	0.0577	0.0620	0.0787	0.0262	0.0776	0.0826	0.0758	0.0638
H7T9C3F0														0.0253	0.0344	0.0457	0.0203	0.0203	0.0274	0.0304	0.0632	0.0315	0.0443	0.0468	0.0406
H3T0C4F1															0.0282	0.0629	0.0343	0.0250	0.0321	0.0479	0.0538	0.0506	0.0518	0.0554	0.0482
H8T4C0F1																0.0513	0.0396	0.0366	0.0394	0.0598	0.0379	0.0513	0.0525	0.0576	0.0488
H3T6C3F0																	0.0515	0.0538	0.0711	0.0794	0.0835	0.0301	0.0721	0.0473	0.0252
H8T0C5F0*																		0.0000	0.0061	0.0038	0.0766	0.0378	0.0477	0.0482	0.0461
H8T0C5F0*J																			0.0096	0.0129	0.0622	0.0374	0.0462	0.0557	0.0508
H9T9C9F0*`																				0.0023	0.0764	0.0525	0.0572	0.0541	0.0544
H9T9C9F0*`J																					0.0876	0.0516	0.0656	0.0716	0.0686
H3T0C4F0																						0.0882	0.0877	0.1085	0.0900
H6T8C4F4*																							0.0175	0.0481	0.0295
H6T8C4F4*J																								0.0624	0.0548
H8T0C3F0*																									0.0091

J at the end of population names differentiates juvenile populations.

Supplementary Material 6. Bottleneck analysis results for adult and juvenile populations. Number of loci out of nine that presented higher expected heterozygosity than heterozygosity under mutation-drift equilibrium and p-values for Wilcoxon test, significant when ≤ 0.05.

	Population	Loci with He>Heq	Wilcoxon test p-value
Adult	H6T0C0F0	3	0.7871
	H0T0C0F0	2	0.9815
	H8T4C4F0	2	1.0000
	H3T0C4F0*	1	0.7148
	H8T3C5F0	0	0.9902
	H9T7C6F0	4	0.7520
	H8T9C8F0	0	0.9756
	H9T9C9F0	1	0.9356
	H9T9C9F0*	1	0.8984
	H3T0C0F1	2	0.9815
	H4T5C6F0	3	0.9863
	H7T9C3F0	2	0.7148
	H3T0C4F1	2	0.9863
	H8T4C0F1	1	0.6738
	H3T6C3F0	0	0.8203
	H8T0C5F0*	2	0.7520
	H9T9C9F0*`	2	0.9815
	H3T0C4F0	3	0.9971
	H6T8C4F4*	1	0.9971
	H8T0C3F0*	3	0.9990
Juvenile	H3T0C4F0*J	3	0.9990
	H9T9C9F0*J	2	1.0000
	H8T0C5F0*J	3	0.9932
	H9T9C9F0*`J	2	0.9932
	H6T8C4F4*J	2	1.0000
	H8T0C3F0*J	3	0.9971

J at the end of population names differentiates juvenile populations.



Supplementary Material 7. Spatial autocorrelation correlograms of kinship coefficient of Loiselle per distance class for pairwise individuals for 20 individual (a-t) and combined (u) adult populations. 95% confidence interval for the kinship coefficient is shown by black dots (·····) and region of acceptance of the null hypothesis to test for absence of SGS is shown by grey dashes (---). Distance classes are out of scale in axis for presentation purposes.

II. Cattle ranching affects the genetic diversity of natural populations of *Dipteryx alata*, a vulnerable tree from the Brazilian savanna

1. Introduction

The importance of genetic diversity has been traditionally associated to population fitness and viability on both short and long term: on the short term, loss of heterozygosity can result in reduced individual fitness; on the long term, loss of genetic diversity can reduce adaptive and evolutionary potential (Vranckx *et al.* 2012). Genetic diversity can also be of great relevancy due to its effects on ecological processes, such as primary productivity, recovery from disturbances, interspecific competition, community structure and energy and nutrient fluxes, presenting similar importance to that of species diversity (Hughes *et al.* 2008). However, conservation genetics is still overlooked in management and conservation programs and in national and international policies; even with increasing environmental modification and exploration of natural resources, monitoring of possible genetic consequences are still uncommon (Laikre 2010). To simply describe genetic variation is not sufficient to support biodiversity conservation and sustainable use of natural resources; it is also essential to identify particular factors that affect genetic diversity, as well as to quantify their effects and elucidate their underlying mechanisms (Rao & Hodgkin 2002). Integrating genetics into a wider context that includes factors such as geographic, environmental, demographic and anthropogenic factors can assist to clarify what the most necessary and efficient strategies for conservation and management are, especially considering that genetics is only one of the factors that affects population viability (Frankham 2010).

Disturbance might be the main driver shaping genetic diversity in many natural populations, most likely resulting in profound eco-evolutionary consequences, especially considering that new, anthropogenic disturbances are emerging and interacting with natural impacts (Banks *et al.* 2013). Among anthropogenic disturbances, land use change is one of the main drivers of modifications of biodiversity and ecosystem processes at several scales; however, its effects on genetic diversity are mostly unknown (Fischer *et al.* 2010). Natural habitats are being converted to areas with agriculture and pasture in the tropics at unprecedented rates (Lambin *et al.* 2003), resulting in landscapes with multiple uses

and functions (Mander *et al.* 2007). Cattle ranching, one of the main land uses in the tropics, can affect genetic diversity of populations through grazing and tramping, which influence plant growth, survival and reproduction and might change physical-chemical soil characteristics (Shan *et al.* 2006), but studies on this subject are still rare (but see Mengli *et al.* 2005; Shan *et al.* 2006). The harvesting of non-timber forest products (NTFP) is a common practice either in habitat remnants or in disturbed areas, where plants from which parts are harvested remain after disturbances, such as in pasturelands (Giroldo & Scariot 2015). Harvesting of NTFP can result in diverse ecological consequences, which are mostly negative, although they can also be neutral or even positive (Ticktin 2004; Stanley *et al.* 2012). These authors showed that there still is a great deficiency of studies on such impacts, especially on levels other than demographic, such as the gene level (but see Shaanker *et al.* 2004; Wang *et al.* 2013; Xu *et al.* 2013; Gaoue *et al.* 2014).

The Brazilian Cerrado is the largest neotropical savanna (Ratter *et al.* 1997; Furley 1999) and one the richest and most threatened areas of the world (Ratter *et al.* 1997). This region presents marked dry and rainy seasons (Silva *et al.* 2008), fire regime as one of its major determinants (Miranda *et al.* 2002) and soils that are old, acid and poor in nutrients (Furley & Ratter 1988; Haridasan 2000). The Cerrado has today an alarming conservation status, presenting only 50% of its original area as native vegetation due to due to agriculture and cattle ranching (MMA 2011). Additionally, only 3% of its territory as fully protected (Françoso *et al.* 2015).

Dipteryx alata Vog. (Fabaceae), popularly known as “Baru”, is a common and well-distributed Cerrado tree species, presenting higher density of individuals in more fertile soils and forest physiognomies of this region (Sano *et al.* 2004). It is economically important due to its nutritionally rich and popular nuts, which are harvested from natural populations that exist in native vegetation, crops and pasture areas, where they are commonly maintained and nursed after (Almeida *et al.* 1998; Sano *et al.* 2004). Harvesting of fruits of *D. alata* is an activity with high social importance because traditional communities and families depend on it as a source of income. This species is diploid, with hermaphrodite flowers and preferentially allogamous (Tarazi *et al.* 2010); it is pollinated mainly by bees (Oliveira & Sigrist 2008); its fruits and seeds are dispersed through gravity and by the action of mammals (Sano *et al.* 2004). Due to historical demographic changes in the last

glaciation period, the species presents today a pattern of low genetic diversity and, as a consequence, it is particularly vulnerable to disturbances (Collevatti *et al.* 2013).

The objective of this study was to quantify and comprehend associations between soil, geographic, demographic and land use and management variables with genetic diversity and structure of natural populations of *Dipteryx alata*, a vulnerable neotropical savanna tree species with ecological, social and economical importance. Using microsatellite markers and generalized linear models, we aimed to answer the following question: Do soil, geographic, demographic and land use and management variables affect the genetic diversity and structure of populations of *D. alata*?

2. Methodology

Methodology was as described in chapter I (page 6), with a few differences, which are described below.

Study areas characterization and sampling

Study areas are presented in Figure 1. The land use and management impacts on the demographic structure of these populations were previously studied (Ferreira 2016).

Instead of using the vegetation thinning index, we used the management index, which here includes any procedure performed in order to facilitate cattle ranching, such as vegetation cutting or pruning and plowing. Fire frequency was not used, since areas did not present such disturbance in the 11 years prior to data collection. Thus, land use and management variables considered were indexes for fruit harvesting, cattle ranching and the management associated with this activity.

Plots for characterization of demographic structure of populations varied from 0.05 to 1 ha, according to density of individuals.

Juvenile individuals sampled for genetic analysis were smaller in size (height ≤ 20 cm). In each of the six populations that compose the three geographical pairwise units (Figure 1), 100 juvenile individuals were sampled in the following manner: 10 juveniles were sampled in a 10 m radius around each of 10 randomly

chosen previously sampled adult tree. A total of 1,253 adult and juvenile individuals were sampled.

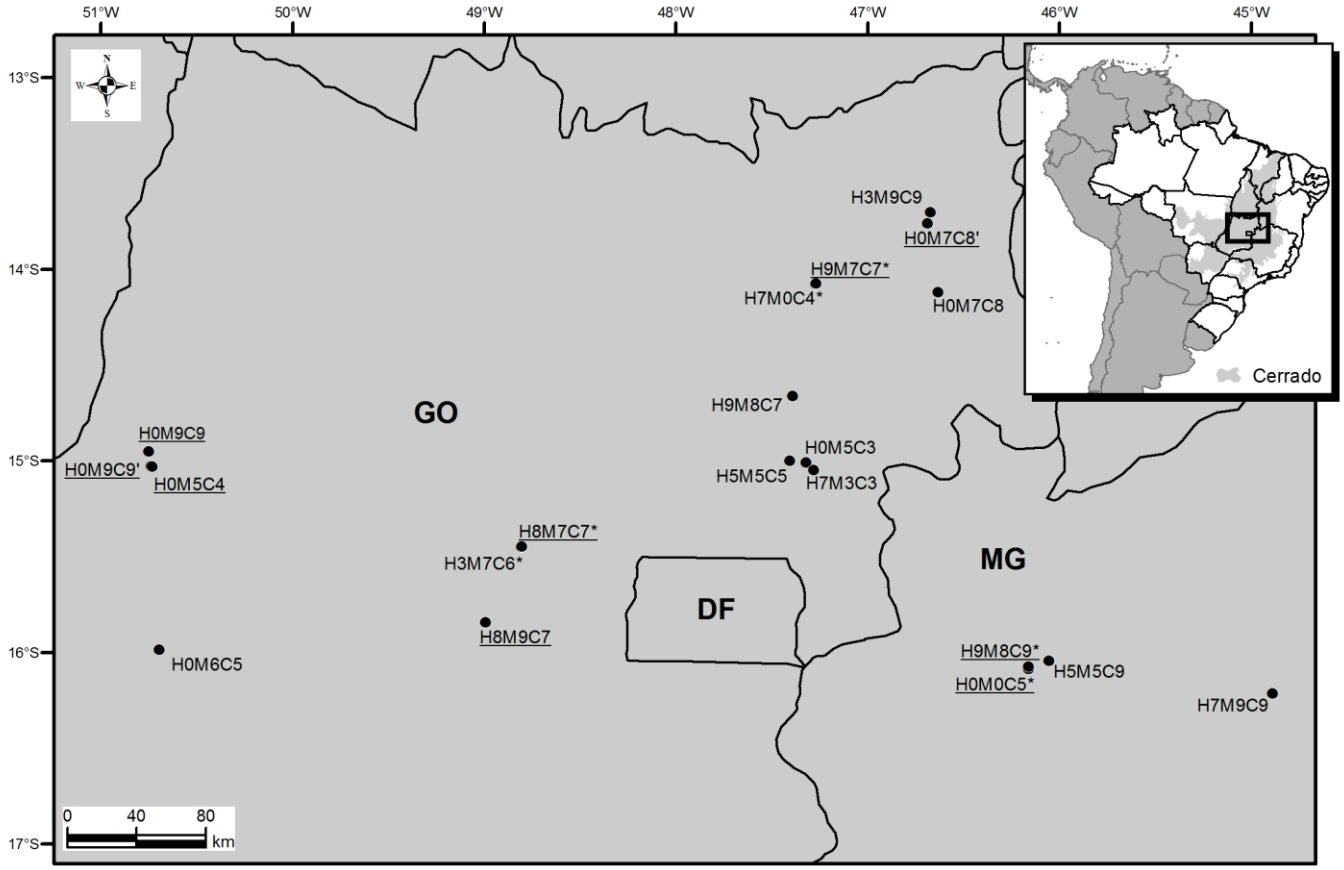


Figure 1. Location of the 20 *D. alata* populations in the Cerrado. Due to the scale used, some points represent more than one population. Areas are named after fruit harvesting (H), management (M) and cattle ranching (C) indexes, which summarize severity, duration and area of disturbances. ` differentiates populations with equal indexes. * discriminates populations that compose the three geographical pairwise units. Underlining of population names segregates populations in two groups based on Bayesian cluster analysis, assuming K = 2, as presented in Supplementary Material 1.

Laboratory analyses

Microsatellite fragments were amplified with 15 previously developed primers: seven were developed for *Dipteryx odorata* (Vinson *et al.* 2009) and then transferred to *D. alata* (Tarazi *et al.* 2010); and eight were developed for *D. alata* (Soares *et al.* 2012). Annealing temperatures followed Tarazi *et al.* (2010) and Soares *et al.* (2012).

Data analyses

Based on the test for absence of SGS in the initial distance classes of spatial autocorrelation correlograms of kinship coefficient (Loiselle *et al.* 1995), we classified individual populations in binary categories: presence (1) or absence (0) of SGS.

In the analysis with GLMs performed to identify associations with the genetic diversity and structure of natural adult populations of *D. alata,* we also used presence of SGS as a response variable, additionally to allelic richness and expected heterozygosity. Four sets of explanatory variables were initially considered: soil (clay content, sand content, silt content, pH, cation exchange capacity, base saturation and aluminum saturation), geographic (latitude and altitude), demographic (adult density) and land use and management (cattle ranching, management and fruit harvesting indexes). After removal of the most difficult variables to interpret when paired variables were collinear ($r \geq 0.7$, through Pearson correlation), all variables cited above were used as explanatory variables in the initial models, except for: soil variables, which were summarized by sand content, pH and cation exchange capacity; and land use and management variables, which were summarized by cattle ranching and fruit harvesting indexes (Supplementary Material 2 and 3), in a total of eight explanatory variables. We used the binomial distribution as the variance function and “logit” as the link function in the GLMs with presence of SGS as a response variable. Fifteen combinations were tested for the allelic richness, while only one combination was tested for expected heterozygosity and presence of SGS, since the global models for these response variables presented one and zero explanatory variables, respectively.

Since the global model for presence of SGS did not differ statistically from the null model, we did not use this variable as a response variable in the ANOVA performed to detect effects of land use and size class in the genetic diversity of the adult and juvenile populations that constitute the three geographical pairwise units (Figure 1). Thus, we used only allelic richness and expected heterozygosity as response variables. The three geographical regions of the pairwise units, which correspond to blocks, are: 1 = H7M0C4*/H9M7C7*, 2 = H0M0C5*/H9M8C9* and 3 = H3M7C6*/H8M7C7* (Figure 1).

3. Results

Genetic diversity and spatial structure

Land use and management indexes are presented in Figure 1; demographic and soil parameters are presented in Supplementary Material 4. No loci combination showed linkage disequilibrium in more than 50% of populations, indicating that such detected associations between loci are most likely related to random allele distribution in specific populations. No locus presented evidence for null alleles in more than 50% of the populations analyzed. Since the frequency of null alleles is estimated based on homozygosity excess in comparison to the Hardy-Weinberg equilibrium, these results might just be a consequence of specific populations presenting lower heterozygosity than expected by this model.

Populations did not differ considerably in genetic diversity estimates. Allelic richness varied from 4.20 to 6.07, with a mean of 5.21; expected heterozygosity varied from from 0.50 to 0.65, with a mean of 0.56; observed heterozygosity varied from 0.43 to 0.67, with a mean of 0.52; inbreeding coefficients were not statistically different from zero in all but four populations, of which three were juvenile (Table 1). The inbreeding coefficient for the entire set of populations also did not statistically differ from zero (0.0818; confidence interval: -0.0064 to 0.1685). Genetic divergence between pairs of populations varied from 0.0000 to 0.2328, with an average of 0.1348 (Supplementary Material 5). Genetic divergence for the entire set of populations was 0.1498, with a confidence interval of 0.1244 to 0.1766. Bottleneck analysis showed that, from all 26 adult and juvenile populations, only one (H7M9C9) went through a recent bottleneck (Supplementary Material 6).

SGS was present for half of all individual adult populations (H7M0C4*, H9M8C9*, H5M5C5, H9M8C7, H0M5C3, H3M9C9, H7M9C9, H0M6C5, H8M9C7, H0M7C8') and for all populations combined (Supplementary Material 7). The maximum kinship coefficient for individual populations was 0.2527, with an average of 0.1358, with SGS in distances of up to 570.8 m and average distance of 337.06 m. The maximum kinship coefficient for all populations combined was 0.2000 and SGS was present for up to 25.4 Km.

The optimal number of K according to the method of Evanno *et al.* (2005) was two, which divided populations in two clusters: 1) H0M9C9', H0M5C4, H0M9C9, H8M9C7, H0M7C8', H9M7C7*, H0M0C5*, H9M8C9*, H8M7C7*; and 2) H5M5C5,

H9M8C7, H0M5C3, H3M9C9, H0M7C8, H7M3C3, H7M9C9, H0M6C5, H5M5C9, H7M0C4*, H3M7C6* (Supplementary Material 1). These clusters, however, were not assembled geographically (Figure 1).

Table 1. Genetic diversity parameters for the adult and juvenile *D. alata* populations.

	Population	n	A	Ar	He	Ho	f (95% CI)
Adult	H0M9C9'	25.7	4.8667	5.2550	0.5347	0.6084	-0.1412 (-0.2716 to 0.0122)
	H0M5C4	28.5	4.8667	5.0680	0.5721	0.6328	-0.1085 (-0.2309 to 0.0103)
	H5M5C5	26.9	4.6667	5.1552	0.6270	0.6004	0.0434 (-0.2128 to 0.2980)
	H9M8C7	25.7	4.9333	5.2745	0.5751	0.6183	-0.0770 (-0.2365 to 0.0675)
	H0M5C3	27.9	5.6000	5.8827	0.5898	0.5705	0.0333 (-0.1402 to 0.2004)
	H3M9C9	21.0	4.2000	4.7632	0.5533	0.4569	0.1790 (-0.0701 to 0.4316)
	H0M9C9	25.6	4.6000	5.0395	0.5648	0.6663	-0.1844 (-0.3936 to 0.0009)
	H0M7C8	26.1	4.2000	4.7747	0.5172	0.5161	0.0020 (-0.1592 to 0.1693)
	H7M3C3	20.7	4.7333	5.4760	0.5477	0.4814	0.1233 (-0.0204 to 0.2800)
	H7M9C9	24.3	4.8000	5.2796	0.6536	0.6014	0.0800 (-0.1564 to 0.3323)
	H0M6C5	25.7	5.3333	5.9024	0.5650	0.4935	0.1284 (-0.0284 to 0.2844)
	H5M5C9	23.9	4.6667	5.3075	0.5761	0.5636	0.0222 (-0.1347 to 0.1947)
	H8M9C7	27.7	3.8667	4.1970	0.5084	0.5480	-0.0797 (-0.2620 to 0.0880)
	H0M7C8'	26.7	4.6667	4.9939	0.5262	0.5788	-0.1019 (-0.2282 to 0.0651)
	H7M0C4*	16.1	4.8000	5.8667	0.5513	0.4495	0.1886* (0.0276 to 0.3673)
	H9M7C7*	23.1	5.1333	5.6475	0.5895	0.5267	0.1054 (-0.0691 to 0.2860)
	H0M0C5*	26.2	4.0000	4.4997	0.5175	0.4389	0.1538 (-0.0061 to 0.3162)
	H9M8C9*	24.6	4.0000	4.6058	0.5171	0.4297	0.1708 (-0.0185 to 0.3550)
	H3M7C6*	26.6	4.7333	4.9626	0.5476	0.4961	0.0945 (-0.1160 to 0.3155)
	H8M7C7*	27.1	4.6000	4.7946	0.5564	0.6211	-0.1188 (-0.2810 to 0.0393)
Juvenile	H7M0C4*J	105.3	6.6667	5.4279	0.5244	0.4309	0.1789* (0.0279 to 0.3360)
	H9M7C7*J	102.8	7.2667	5.8419	0.6102	0.4960	0.1876* (0.0454 to 0.3235)
	H0M0C5*J	94.5	5.1333	4.7687	0.4977	0.4295	0.1373 (-0.0435 to 0.3143)
	H9M8C9*J	101.0	7.0000	5.5304	0.5668	0.4377	0.2284* (0.0753 to 0.3882)
	H3M7C6*J	81.9	5.7333	5.1059	0.5005	0.4704	0.0604 (-0.0966 to 0.2138)
	H8M7C7*J	89.8	6.6000	6.0721	0.5443	0.4868	0.1060 (-0.0489 to 0.2452)
Average		41.4	5.0641	5.2113	0.5552	0.5250	0.0543

J at the end of population names differentiates juvenile populations. n = sample size; A = average allele number per locus; Ar = average allelic richness per locus; He = expected heterozygosity; Ho = observed heterozygosity; f = inbreeding coefficient. * Significantly different from zero based on 95% confidence intervals estimated by 10,000 bootstraps.

Associations of geography, soil, demography and land use and management with genetic diversity

Five and one models were selected based on AIC for allelic richness and expected heterozygosity of 20 *D. alata* adult populations, respectively (Table 2). No model was selected for the presence of SGS because the global model for this response variable presented zero explanatory variables and did not differ statistically from the null model, so results were not presented here. All other global models were statistically different from null models (Table 3). Global model for allelic richness was built with four variables: cattle ranching index, fruit harvesting index, adult density and altitude, of which cattle ranching index and altitude were statistically significant

(Table 3). The expected heterozygosity global model was built with only one statistically significant explanatory variable, sand content (Table 3).

ANOVA results showed that the effect of region (block) was statistically significant for allelic richness and expected heterozygosity (p-values = 0.0186 and 0.0298, respectively) (Figure 2). Pasturelands presented significantly higher expected heterozygosity than areas with Cerrado vegetation (p-value = 0.0069) (Figure 2). The effect of interaction between size class and land use was significant for expected heterozygosity (p-value = 0.0468), since the difference between expected heterozygosity of pasturelands and Cerrado areas is reasonably higher for the juvenile size class than for the adult size class (Figure 2, Supplementary Material 8).

Table 2. Best models and global models for genetic diversity (Ar = allelic richness, He = expected heterozygosity) of 20 adult *D. alata* populations. N is the number of variables. Global models are the first models (1).

Response variable	Model number	Model	N	AIC
Ar	Ar(1)	~ CAT** + FRU + DENS + ALT*	6	24.780
	Ar(2)	~ CAT* + DENS + ALT	5	25.373
	Ar(3)	~ CAT* + ALT	4	26.515
	Ar(4)	~ CAT* + DENS	4	27.228
	Ar(5)	~ CAT	3	27.334*
He	He(1)	~ SAND*	3	-74.919

CAT = cattle ranching index; FRU = fruit harvesting index; DENS = adult density/ha; ALT = altitude; SAND = sand content (g/Kg);
Significance codes: *** = 0.001; ** = 0.01; * = 0.05. Codes after variables in models point to the significance of the variable for the model. Codes after AIC values point to statistically different models from the global model for each response variable.

Table 3. Global models estimates of explanatory variables and their respective standard deviation for genetic diversity (Ar = allelic richness; He = expected heterozygosity) of 20 adult populations.

Explanatory Variable	Response Variable	
	Ar (1) *	He (1) *
	β ± SD	β ± SD
Intercept	7.3201 ± 0.6259 ***	0.4984 ± 0.0287 ***
CAT	-0.1757 ± 0.0533 **	
FRU	0.0411 ± 0.0285	
DENS	-0.0071 ± 0.003	
ALT	-0.0018 ± 0.0007 *	
SAND		0.0001 ± 0.0000 *

CAT = cattle ranching index; FRU = fruit harvesting index; DENS = adult density/ha; ALT = altitude; SAND = sand content (g/Kg);
Significance codes: *** = 0.001; ** = 0.01; * = 0.05. Codes after values for variables in models point to the significance of the variable for the model. Codes after model number point to statistically different models from null models (response variable ~ 1).

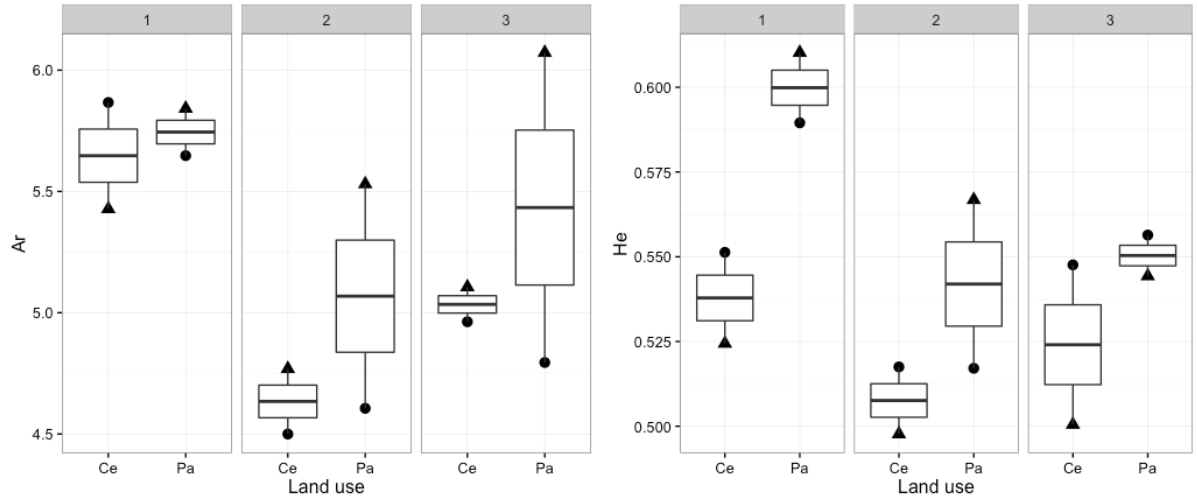


Figure 2. Boxplot of data used for ANOVA with a factorial design and randomized blocks for six populations in three geographical pairwise units. Ar = allelic richness, He = expected heterozygosity. Blocks = regions (1 = H7M0C4*/H9M7C7*, 2 =H0M0C5*/H9M8C9* and 3 = H3M7C6*/H8M7C7*); factors = land use (Ce = Cerrado, Pa = Pasture) and size class (● = adult, ▲= juvenile).

4. Discussion

Genetic diversity and spatial structure

Low genetic diversity is a common feature of *D. alata* (Tarazi *et al.* 2010; Melo *et al.* 2011; Soares *et al.* 2012; Collevatti *et al.* 2013), which is explained by the demographic history of the species. The low genetic diversity of *D. alata* is most likely a consequence of habitat instability during the last glaciation period, which possibly led to range shifts and distribution restriction and subsequent expansion for the species (Collevatti *et al.* 2013). During this period, the climate was drier in the Cerrado region, leading to an expansion of this savanna and restringing distribution of *D. alata,* which is more adapted to moister soil conditions.

Inbreeding coefficients were mostly not statistically different from zero, contrary to findings by Collevatti *et al.* (2013). These authors found positive and significant inbreeding coefficients for populations of the species; however, Tarazi *et al.* (2010) demonstrated that, for the populations analyzed in their study, such values had been overestimated due to a detected existence of Wahlund effect and, after correction, were in fact close to zero. Juvenile populations presented, in general, higher estimates of inbreeding coefficients than adult ones. This difference between size classes might be a consequence of delay in response to impacts, which is

common for tree species, due to their long lifespan, high phenotypic plasticity, large size and high production of pollen and seed (Vranckx *et al.* 2012).

Mean pairwise genetic divergence among populations and genetic divergence for the entire set of populations are considered moderate (Hartl & Clark 1997). These values are very similar to those found for the species (Soares *et al.* 2008a; Tarazi *et al.* 2010; Melo *et al.* 2011; Collevatti *et al.* 2013) and indicate that populations were moderately connected in the past, probably presenting reasonable gene flow among them. *D. alata* is pollinated mainly by bees (Oliveira & Sigrist 2008) and its fruits and seeds are dispersed through gravity and wild mammals (Sano *et al.* 2004), which can explain moderate distances of gene flow. In fact, Tarazi *et al.* (2010) found a mean distance of pollen dispersal of 610 m for the species.

Presence of SGS varied among populations, which is consistent with other studies, that also encountered SGS for part of the studied populations of the species (Soares *et al.* 2008b; Tarazi *et al.* 2010). We were not able to detect associations of geographic, soil, demographic or land use and management variables with SGS, probably due to the existence of many others unaccounted for features that together influence the existence of SGS (Loveless & Hamrick 1984). For the populations in which we did detect SGS, we observed higher kinship coefficients and distances of SGS than Tarazi *et al.* (2010), who evaluated populations of the species in the States of Goiás (GO), Minas Gerais (MG) and Mato Grosso do Sul (MS). Such differences are, again, consequences of many variables that act differently among areas and lead to specific SGS inside populations. The presence of SGS in a significant part of our populations is most likely a result of the main dispersion syndrome presented by the species, barocory, which results in aggregation of individuals around mother trees (Sano *et al.* 2004).

Associations of geography, soil, demography and land use and management with genetic diversity

Considering the negative association of cattle ranching with allelic richness and bearing in mind that alleles are the most sensible genetic parameter being affected by recent impacts (Cornuet & Luikart 1996), our results points to cattle ranching as being an important and recent driver of loss of genetic diversity for populations of *D. alata.* In fact, ranching and crop farming, fragmentation and urban

occupation began to change the Cerrado landscape in a more intense manner approximately 60 years ago (Klink & Moreira 2002), a time gap that does not surpass one or two generations of the species. Bottleneck analysis, however, detected recent reduction of effective size for only one of the populations. This anthropogenic impact encompasses grazing and tramping by cattle, and is associated with management practices such as vegetation cutting and pruning and plowing, which presented high and significant correlation with cattle ranching for our areas. These actions can all influence plant growth and lead to the death of *D. alata* individuals, especially young ones, which can lead to the loss of rare alleles. Also, tramping can lead to physical-chemical changes in the soil, which can alter vegetation structure (Shan *et al.* 2006). Although Ferreira (2016) did not encounter significant associations between cattle ranching and the demographic structure of the same populations we have studied here, this author did find negative and significant effects of management practices, corroborating with our results.

There were no significant associations between *D. alata* fruit harvesting with any of the genetic diversity parameters. Similarly, Ferreira (2016) did not observe negative consequences of such activity on the demography of the same populations we studied here, corroborating with our results. The absence of negative consequences of *D. alata* fruit harvesting to both genetic diversity and demographic structure is related to the fact that harvesting is performed directly from the ground, and, thus, does not cause damage to the tree. A reasonable portion of the few existing studies on the genetic consequences of NTFP harvesting also show inexistence of negative effects (Shaanker *et al.* 2004; Wang *et al.* 2013; Xu *et al.* 2013). However, such results will depend on specific aspects of each case, such as life history, the part of the plant that is harvested, variation in environmental conditions and management practices (Ticktin 2004).

Altitude presented a negative and significant association with allelic richness. Genetic diversity of populations and altitudinal gradient are related in complex ways. This relation can be of different types, one of them being decrease of genetic variability with altitude, which has been demonstrated for various species (Ehinger *et al.* 2002; Ohsawa & Ide 2008; Thiel-Egenter *et al.* 2009; Yan *et al.* 2009; Hahn *et al.* 2012; Jugran *et al.* 2013; Shen *et al.* 2014). Although our altitudes only varied from 277 to 798 m, even small altitudinal gradients can encompass many different environmental and ecological variables that can constrain the genetic diversity of

populations (Yan *et al.* 2009). Climatic conditions associated to higher altitudes, such as lower temperatures and drier soils, can lead to reduction in population effective size, decreasing genetic diversity due to genetic drift (Thiel-Egenter *et al.* 2009; Yan *et al.* 2009; Hahn *et al.* 2012; Shen *et al.* 2014). Although there is no established association between soil and air humidity with altitudinal gradients for the Cerrado, temperatures are generally inversely related to altitude in this region (Cardoso *et al.* 2014). According to the metabolic theory of biodiversity, higher temperatures lead to higher mutation rates; it also leads to shorter generation times and higher selection, increasing evolutionary speed (Adams & Hadly 2012). Thus, it is possible that *D. alata* populations at higher altitudes present lower evolutionary rates, diminishing genetic diversity in comparison to populations at lower altitudes. Furthermore, in higher altitudes, populations can become more isolated, with limitation of gene flow due to lower number and activity of pollinators, leading to inbreeding and loss of diversity (Thiel-Egenter *et al.* 2009; Yan *et al.* 2009; Hahn *et al.* 2012; Shen *et al.* 2014). Although moderate distances of pollen dispersal estimated for a *D. alata* population indicate that distance of gene flow is not critically limited for this species (Tarazi *et al.* 2010), diversity of bees, which are the main pollinators of *D. alata*, might be lower at higher altitudes (Karunaratne & Edirisinghe 2008), restricting the magnitude of gene flow.

There was a positive and significant association of sand content with expected heterozygosity. Plant development and survival are significantly influenced by soil texture. Soils with higher contents of sand do not hold water or adsorb nutrients cations satisfactorily, leading to loss of nutrients by leaching (Gurevitch *et al.* 2006). This is especially relevant to *D. alata*, which is more adapted to moister and richer soils, occurring in higher densities in areas with these characteristics (Sano *et al.* 2004). Thus, we expected that soils with higher contents of sand would lead to decreases of genetic diversity of *D. alata* populations through mortality of less fit individuals, which is the opposite of our results. Since no mechanism could be found to explain such results, we speculate that sand content might be a proxy for unaccounted for ecologically relevant variables that are associated positively with the genetic diversity of the studied populations.

In spite of GLM results showing evidence of negative impacts on allelic richness due to cattle ranching, ANOVA results showed that pasturelands present significantly higher expected heterozygosity than Cerrado areas. Such difference

between both analyses can be attributed to several possibilities. First, only adult populations were considered in the analysis with GLMs, while both adult and juvenile populations were used for ANOVA. Second, higher sample sizes were used for GLM analysis (20 populations, while only six were used for ANOVA). Third, while a categorical binary land use classification was used in ANOVA, we used a continuous gradient to classify cattle ranching for GLM analysis, which better represents reality. Fourth, other explanatory variables were considered in the GLM analysis, possibly refining associations of cattle ranching in the model, since underfitted models can miss important associations (Burnham & Anderson 2002). Thus, considering these differences between analyses, we consider our GLM results to be more representative of reality and more reliable.

Implications for conservation and management

Based on maximum distances up to which SGS is present in individual populations, we recommend that collection of seeds for conservation or restoration projects should be from trees with a minimum distance of 337 m from each other, as to avoid collecting germplasm from close-related individuals. Also, considering SGS for all populations combined, maximization of collected genetic diversity would be achieved if populations were apart by at least 25 Km. In addition, since genetic diversity found in natural populations was relatively low and genetic divergence among populations was moderate, we recommend that seed collection should be from as many different populations and trees as possible, in order to maximize representation (Broadhurst *et al.* 2008).

Our results show that cattle ranching can affect negatively the allelic richness of *D. alata* populations. Allelic richness should be given high priority as a variable in the planning of genetic conservation, since it is highly dependent of effective population size and better indicates past demographic changes (Petit *et al.* 1998). While heterozygosity is a good measure of the capacity of a population to respond immediately to selection in the short term, providing fitness and survival of individuals, allelic richness will limit the response to selection over long periods, and will restrain evolution and survival of populations and species (Allendorf 1986). *D. alata* presents current low genetic diversity due to historic events (Collevatti *et al.* 2013). Thus, this species is vulnerable to any impact that could lead to additional

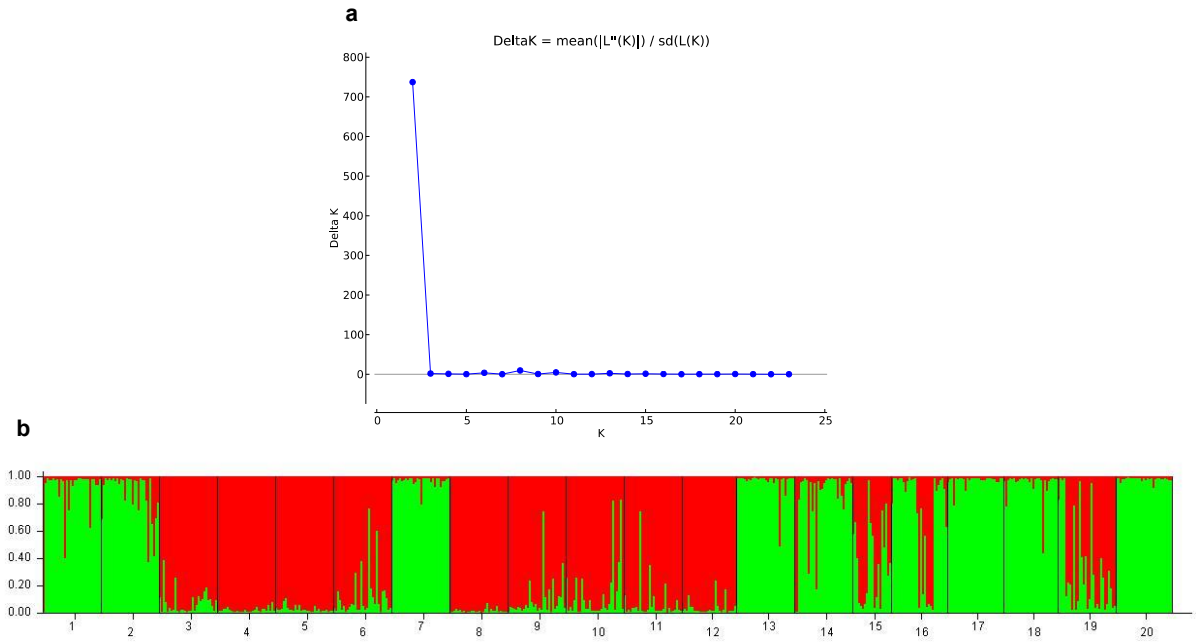
loss of genetic diversity, which, through reduced fitness and higher mortality rates, could lead to local extinction of populations.

Cattle ranching is extremely common in the Cerrado, representing today one of the biggest threats to the conservation of the biodiversity of this region. Abandoned pasturelands are not uncommon; consequently, further deforestation of Cerrado native areas is not only unnecessary, but also threatening to the Cerrado biodiversity (Klink & Moreira 2002). Management practices such as fencing target individuals from grazing and tramping, increasing the rotation period between pasture areas and intervals between thinning events and decreasing thinning intensity, as proposed by Ferreira (2016), should increase survival of small individuals of *D. alata* and support their development into adulthood. This would not only increase population effective sizes, but also be financially interesting for landowners, since the seeds of *D. alata* are increasingly popular as a nutrition product and could provide for additional income (Sano *et al.* 2004). It would also be beneficial for cattle grazing itself, since shadow from trees reduce heat stress levels and productivity losses (Blackshaw & Blackshaw 1994). Furthermore, *D. alata* fruit pulp is an additional nutrition resource for these animals, especially during dry periods (Sano *et al.* 2004). Considering that divergence among populations is moderate and that the genetic diversity of populations is low, natural areas around these populations should be maintained or restored in order to form natural corridors, as to connect populations through gene flow, avoiding inbreeding and genetic diversity erosion due to genetic drift, which could also be achieved through occasional introduction of local immigrants.

Combining multiple uses, such as cattle grazing and NTFP harvesting, has been advocated as a important conservation strategy for transformed landscapes (Moilanen *et al.* 2005). Our results further supports the sustainability of *D. alata* fruit harvesting (Ferreira 2016), which should continue to be promoted as an activity that conserves the Cerrado, especially considering its socioeconomic importance for people from traditional communities and poor families. For all of these proposed actions to occur in a significant manner, however, public politics should be developed to promote them through financial and tax incentives. Nevertheless, considering that human demands of this resource will continue to grow and that *D. alata* has gone through a historic bottleneck, reducing its genetic diversity and making this species particularly vulnerable to environmental pressure (Collevatti *et*

al. 2013), technical management recommendations for sustainable harvesting should be applied (Carrazza & Ávila 2010).

Supplementary Material



Supplementary material 1. Bayesian cluster analysis results: a) Selection of optimal number of clusters (K), based on the method of Evanno *et al.* (2005); b) Clustering of 583 *D. alata* adult individuals, collected from 20 *a priori* populations, assuming K = 2. Population identification: 1 = H0M9C9'; 2 = H0M5C4; 3 = H5M5C5; 4 = H9M8C7; 5 = H0M5C3; 6 = H3M9C9; 7 = H0M9C9; 8 = H0M7C8; 9 = H7M3C3; 10 = H7M9C9; 11 = H0M6C5; 12 = H5M5C9; 13 = H8M9C7; 14 = H0M7C8'; 15 = H7M0C4*; 16 = H9M7C7*; 17 = H0M0C5*; 18 = H9M8C9*; 19 = H3M7C6*; 20 = H8M7C7*.

Supplementary Material 2. Pearson correlation for pairwise soil parameters of 20 studied areas.

	Si	Sa	pH	CEC	BS	AS
C	0.9057***	-0.9985***	0.2294	0.6663 ***	0.2316	-0.2380
Si	-	-0.9274***	0.0706	0.5679**	-0.0180	-0.0384
Sa		-	-0.2116	-0.6609*	-0.2022	0.2151
pH			-	0.0028	0.9136***	-0.8841***
CEC				-	0.2158	0.0860
BS					-	-0.8612***

C = clay content (g/Kg); Si = silt content (g/Kg); Sa = sand content (g/Kg); CEC = cation exchange capacity ($cmol_c/dm^3$), BS = base saturation (%); AS = aluminum saturation (%). Significance codes: *** = 0.005; ** = 0.01; * = 0.05. Codes points to the significance of correlation between variables.

Supplementary Material 3. Pearson correlation for pairwise land use and management parameters of 20 studied areas.

	MI	FI
CI	0.7200***	0.0923
MI	-	0.0875

CI = cattle ranching index; MI = management index; FI = fruit harvesting index. Significance codes: *** = 0.005; ** = 0.01; * = 0.05. Codes points to the significance of correlation between variables.

Supplementary Material 4. Location, physiognomy, demographic and soil parameters for the 20 studied areas with populations.

Population	Longitude	Latitude	Altitude (m)	Physiognomy	AD	Sa	pH	CEC
H0M9C9'	-50.737	-15.026	310	PA	34	620	4.5	5.2
H0M5C4	-50.732	-15.031	308	WO	46	780	4.2	5.26
H5M5C5	-47.409	-14.999	563	PA /TC	108	690	4.1	7.65
H9M8C7	-47.392	-14.661	563	PA	8	790	4.3	7.17
H0M5C3	-47.321	-15.01	622	DC/WO	48	580	4.1	8.29
H3M9C9	-46.673	-13.702	502	PA	9	280	5	10.34
H0M9C9	-50.752	-14.952	277	PA	4.5	590	4.6	3.81
H0M7C8	-46.633	-14.12	559	PA	5	830	4.3	5.22
H7M3C3	-47.283	-15.047	580	TC	100	700	4.1	7.05
H7M9C9	-44.888	-16.213	594	PA	9	870	4.9	3.96
H0M6C5	-50.695	-15.986	379	WO	26	610	4.7	5.24
H5M5C9	-46.055	-16.045	515	PA	15	820	4.1	5.25
H8M9C7	-48.994	-15.843	746	PA	39	590	4.3	6.42
H0M7C8'	-46.689	-13.76	515	PA	7	300	4.5	8.7
H7M0C4*	-47.273	-14.076	551	DC	12	570	4.5	4.39
H9M7C7*	-47.271	-14.073	526	PA	15	500	3.8	9.11
H0M0C5*	-46.161	-16.085	540	TC	61	420	4.1	9.53
H9M8C9*	-46.162	-16.075	526	PA	43	390	4.5	7.53
H3M7C6*	-48.809	-15.45	790	WO	4	540	5.2	10.03
H8M7C7*	-48.805	-15.447	798	PA	16	640	4.4	5.42

AD = adult density/ha; Sa = sand content (g/Kg); CEC = cation exchange capacity ($cmol_c/dm^3$); PA = pastureland; WO = woodland; TC = typical Cerrado; DC = dense Cerrado.

Supplementary Material 5. *Fst* (genetic divergence) estimates for pairwise adult and juvenile populations.

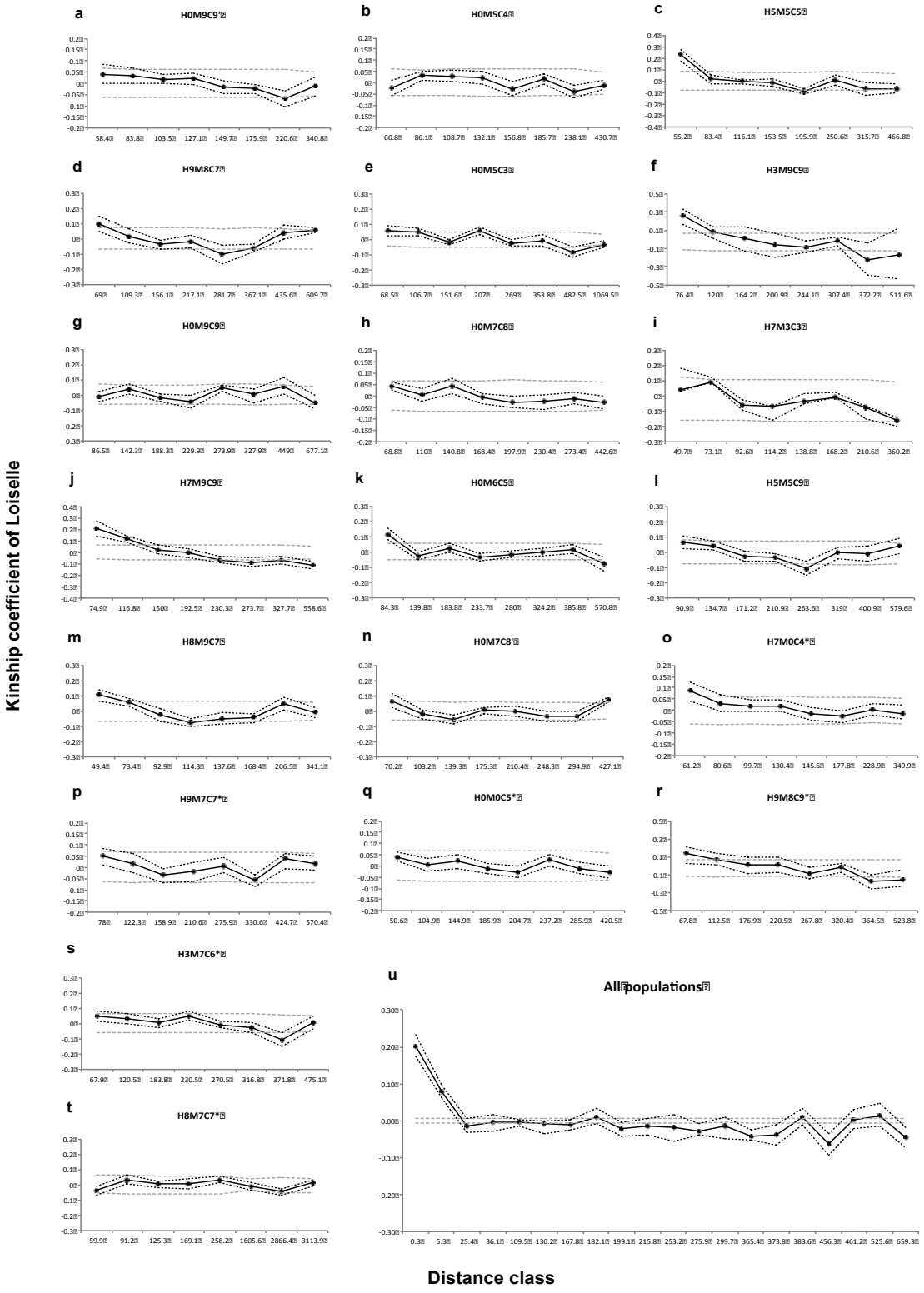
	H0M5C4	H5M5C5	H9M8C7	H0M5C3	H3M9C9	H0M9C9	H0M7C8	H7M3C3	H7M9C9	H0M6C5	H5M5C9	H8M9C7	H0M7C8'	H7M0C4*	H9M7C7*	H0M0C5*	H9M8C9*	H3M7C6*	H8M7C7*	H7M0C4*J	H9M7C7*J	H0M0C5*J	H9M8C9*J	H3M7C6*J	H8M7C7*J
H0M9C9'	0.0505	0.1705	0.1523	0.1420	0.2031	0.0296	0.1679	0.1130	0.1799	0.1288	0.1449	0.1133	0.1102	0.1349	0.1459	0.1666	0.1647	0.1396	0.1206	0.1325	0.1418	0.1563	0.1208	0.1276	0.1339
H0M5C4		0.1217	0.1507	0.1437	0.1967	0.0234	0.1548	0.1555	0.1662	0.1508	0.1346	0.1138	0.0880	0.1200	0.1396	0.1241	0.1196	0.1387	0.0928	0.1090	0.1258	0.1218	0.0959	0.1411	0.1417
H5M5C5			0.1474	0.1206	0.1111	0.1415	0.1534	0.1259	0.0747	0.1350	0.1507	0.1474	0.1336	0.1348	0.1485	0.1606	0.1426	0.1423	0.1475	0.1575	0.1231	0.1563	0.1285	0.1406	0.1289
H9M8C7				0.1277	0.1948	0.1679	0.0482	0.1246	0.1787	0.1124	0.0366	0.2071	0.1409	0.0800	0.1405	0.1591	0.1390	0.1175	0.1930	0.0958	0.1299	0.1478	0.1221	0.1756	0.1174
H0M5C3					0.2002	0.1590	0.1519	0.0647	0.1708	0.0103	0.1337	0.1538	0.1657	0.1471	0.1336	0.1711	0.1625	0.0893	0.1594	0.1589	0.1218	0.1594	0.1325	0.1236	0.0818
H3M9C9						0.1826	0.1994	0.1419	0.0688	0.1958	0.1941	0.2275	0.1727	0.1369	0.1867	0.2257	0.1924	0.2043	0.2252	0.1940	0.1641	0.2078	0.1883	0.1913	0.1875
H0M9C9							0.1689	0.1337	0.1608	0.1550	0.1415	0.1088	0.1041	0.1286	0.1306	0.1455	0.1451	0.1540	0.0940	0.1304	0.1267	0.1443	0.1134	0.1430	0.1529
H0M7C8								0.1394	0.2022	0.1450	0.0515	0.2328	0.1340	0.0930	0.1822	0.1822	0.1741	0.1729	0.2168	0.1014	0.1683	0.1812	0.1642	0.1966	0.1723
H7M3C3									0.1304	0.0489	0.1237	0.1495	0.1468	0.1010	0.0963	0.1802	0.1630	0.1091	0.1523	0.1611	0.0914	0.1630	0.1251	0.1178	0.0968
H7M9C9										0.1710	0.1759	0.1849	0.1586	0.1425	0.1860	0.1976	0.1772	0.1804	0.1829	0.1872	0.1652	0.1984	0.1685	0.1750	0.1700
H0M6C5											0.1188	0.1647	0.1571	0.1348	0.1175	0.1677	0.1553	0.0910	0.1762	0.1540	0.1053	0.1545	0.1182	0.1232	0.0797
H5M5C9												0.2017	0.1439	0.0937	0.1486	0.1439	0.1393	0.1354	0.1774	0.1028	0.1369	0.1540	0.1308	0.1744	0.1388
H8M9C7													0.1405	0.1789	0.1591	0.1410	0.1480	0.1169	0.0436	0.1810	0.1465	0.1371	0.1146	0.0821	0.0947
H0M7C8'														0.0763	0.1269	0.1232	0.1136	0.1381	0.1279	0.0818	0.1124	0.1207	0.0958	0.1520	0.1486
H7M0C4*															0.0900	0.1348	0.1122	0.1417	0.1592	0.0116	0.0751	0.1232	0.0823	0.1663	0.1371
H9M7C7*																0.1538	0.1170	0.1014	0.1402	0.1341	0.0000	0.1407	0.0787	0.1353	0.0934
H0M0C5*																	0.0222	0.1195	0.1452	0.1347	0.1443	0.0105	0.0668	0.1722	0.1300
H9M8C9*																		0.1143	0.1542	0.1221	0.1021	0.0231	0.0327	0.1728	0.1202
H3M7C6*																			0.1040	0.1624	0.1020	0.1117	0.0968	0.0873	0.0168
H8M7C7*																				0.1689	0.1327	0.1435	0.1203	0.0778	0.0970
H7M0C4*J																					0.1121	0.1304	0.0947	0.1803	0.1598
H9M7C7*J																						0.1316	0.0629	0.1262	0.0916
H0M0C5*J																							0.0604	0.1729	0.1243
H9M8C9*J																								0.1322	0.0925
H3M7C6*J																									0.0484

J at the end of population names differentiates juvenile populations.

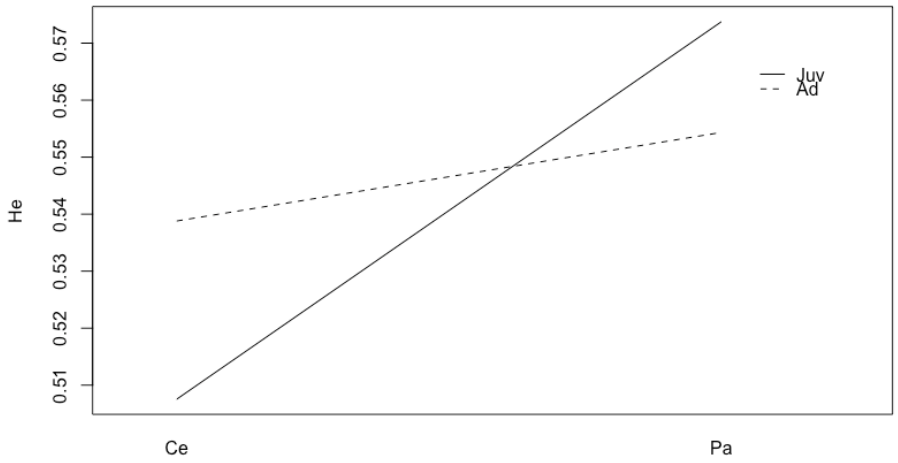
Supplementary Material 6. Bottleneck analysis results for adult and juvenile populations. Number of loci out of 15 that presented higher expected heterozygosity than heterozygosity under mutation-drift equilibrium and p-values for Wilcoxon test, significant when ≤ 0.05 (*).

	Population	Loci with He>Heq	Wilcoxon test p-value
Adult	H0M9C9'	1	0.9993
	H0M5C4	4	0.9584
	H5M5C5	0	0.5980
	H9M8C7	2	0.9243
	H0M5C3	3	0.9849
	H3M9C9	2	0.9465
	H0M9C9	2	0.7894
	H0M7C8	1	0.8662
	H7M3C3	2	0.9723
	H7M9C9	1	0.0473*
	H0M6C5	4	0.9970
	H5M5C9	1	0.8854
	H8M9C7	2	0.9675
	H0M7C8'	5	0.9723
	H7M0C4*	3	0.9996
	H9M7C7*	2	0.9547
	H0M0C5*	1	0.7492
	H9M8C9*	3	0.7557
	H3M7C6*	3	0.9156
	H8M7C7*	2	0.8662
Juvenile	H7M0C4*J	5	1.0000
	H9M7C7*J	7	0.9996
	H0M0C5*J	4	0.9823
	H9M8C9*J	6	0.9999
	H3M7C6*J	4	0.9938
	H8M7C7*J	4	0.9995

J at the end of population names differentiates juvenile populations.



Supplementary Material 7. Spatial autocorrelation correlograms of kinship coefficient of Loiselle per distance class for pairwise individuals for 20 individual (a-t) and combined (u) adult populations. 95% confidence interval for the kinship coefficient is shown by black dots (·····) and region of acceptance of the null hypothesis to test for absence of SGS is shown by grey dashes (---). Distance classes are out of scale in axis for presentation purposes.



Supplementary material 8. Expected heterozygosity (He) interaction between land use (Ce = cerrado, Pa = pasture) and size class (Ad = adult, Ju = juvenile) for six populations in three geographical pairwise units.

III. Do demographic persistence strategies of two neotropical savanna trees include clonal reproduction?

1. Introduction

Most perennial plants can combine sexual reproduction with asexual or vegetative reproduction, which, not considering somatic mutations, results in genetically identical individuals (Eckert 2001; Arnaud-Haond *et al.* 2007). The relative significance of sexual and asexual reproduction can vary among populations due to environmental and ecological features, considering that clonal reproduction allows for demographic persistence in habitats where sexual reproduction is challenged (Eckert 2001). Vegetative regeneration is favored in cases of environmental disturbance, being able to muffle the genetic effects of population impacts through a delay in responses among generations, buffering the effects of lack of recruitment for long periods and, thus, preserving genetic diversity (Bond & Midgley 2001, 2003; Del Tredici 2001; Aguilar *et al.* 2008). Progenies of vegetative regeneration are not subjected to distresses involved in regeneration by seeds, such as pollination, seed dispersion errors, seed mortality and competition among saplings, being an important strategy for the demographic persistence in environments under disturbance regimes (Bond & Midgley 2001, 2003).

This strategy can, however, interfere significantly in spatial patterns of pollination and reproduction among individuals, resulting in expressive changes in genetic diversity and spatial structure, being able to alter the evolutionary dynamic locally and regionally (Charpentier 2001; Eckert 2001; Hoebee *et al.* 2006). Species that reproduce vegetatively can become dominant due to their high establishment rates, which can limit species diversity and alter community composition and structure (Hoffmann 1998; Del Tredici 2001; Hoffmann & Moreira 2002; Salazar & Goldstein 2014). Regeneration is the most sensible stage to genetic-structural changes in tree species populations (Finkeldey & Ziehe 2004; Ratnam *et al.* 2014); therefore, quantifying the occurrence and detecting the factors leading to vegetative propagation in natural populations is of great importance to properly plan for the management and conservation of these species.

The Cerrado, one of the richest and most threatened areas of the world (Ratter *et al.* 1997), is the largest neotropical savanna, being located in the center of Brazil (Ratter *et al.* 1997; Furley 1999). This savanna is characterized by harsh

environmental conditions and intense anthropogenic disturbances. It presents strong climate seasonality, with marked dry and rainy seasons (Silva *et al.* 2008). Most of its soils are old, weathered, acid, poor in nutrients and with high concentrations of aluminum (Furley & Ratter 1988; Haridasan 2000). It presents fire regime as an important feature (Miranda *et al.* 2002). Agriculture and cattle ranching have led to a massive devastation of the Cerrado, remaining only 50% of its original area (MMA 2011).

The occurrence of fires largely increase vegetative reproduction via root suckers as a coping mechanism for demographic persistence in some Cerrado species (Hoffmann 1998; Hoffmann & Moreira 2002; Salazar & Goldstein 2014). Disturbances such as drought and seasonal climate, as well as changes in nutrient availability, can lead to formation of bud bank and resprouting (Klimešová & Klimeš 2007). Injuries, both on underground and aboveground areas of the plant, such as those caused by grazing pressure or vegetation thinning, can lead to the development of root suckers, which is explained, in the case of aboveground wounding, by breakage of apical dominance in trees (Koop 1987; Del Tredici 2001; Fraser *et al.* 2004; Klimešová & Klimeš 2007). Therefore, many of the typical features of the Cerrado seem to be capable of inducing clonal reproduction in plants. In fact, asexual reproduction can actually have a greater importance than sexual reproduction in the Cerrado and several species of this biome seem to reproduce vegetatively (Hoffmann 1998; Hoffmann & Moreira 2002; Medeiros *et al.* 2008).

However, the capacity of asexual reproduction is still not actually confirmed for most species. Definite verification of asexual reproduction via root suckers is uncommon, especially through genetic confirmation of existing clones. In plants, vegetative reproduction is narrowly defined as root suckering that lead to the formation of new stems at some distance from the original individual (Hoffmann 1998; Del Tredici 2001; Hoffmann & Moreira 2002). Thus, considering that the distance between new stems and the original individual can be quite large and that the physical connection to it may not exist any longer, definite identification of clones through genotyping is critical. In addition, the causes of increases in vegetative reproduction in species that present this ability are still mostly unknown (Hoffmann & Moreira, 2002), despite its importance for management and conservation strategies.

Dipteryx alata Vog. (Fabaceae) and *Caryocar brasiliense* Camb. (Caryocaraceae) are tree species that are common and well distributed in the

Cerrado. In a study of floristic composition of 376 Cerrado areas by Ratter et al. (2003), these species were present in 27% and 61% of all areas, respectively. They are economically important mostly because of its nut (*D. alata*) and fruit (*C. brasiliense*), which are nutritionally rich and popular foods used in diverse forms by local people (Araújo 1995; Almeida *et al.* 1998; Sano *et al.* 2004). Both species have a significant social role as a source of income for regional harvesting communities (Araújo 1995; Sano *et al.* 2004). Considering their ecological, social and economic importance, *D. alata* and *C. brasiliense* can be considered strategic species for studies in the Cerrado.

D. alata is diploid, with hermaphrodite flowers and is preferentially allogamous (Tarazi *et al.* 2010); it is pollinated mainly by bees (Oliveira & Sigrist 2008); its fruits and seeds are dispersed through gravity and wild mammals (Sano *et al.* 2004). *C. brasiliense*, in turn, is diploid, with hermaphrodite flowers (Araújo 1995) and it is preferentially allogamous (Collevatti et al. 2001a, 2010); it is pollinated by small nectarivorous bats and its fruits and seeds are dispersed through gravity and the action of animals (Gribel 1986; Gribel & Hay 1993). According to Araújo (1995) and Almeida et al. (1998), *C. brasiliense* can reproduce asexually in natural conditions, which was corroborated by Medeiros et al. (2008), who observed aerial, basal and underground sprouting in this species, without, however, confirming it through genetic identification.

In this study, we aimed to better understand the demographic persistence mechanisms for *D. alata* and *C. brasiliense* by trying to answer the following questions: 1) Do these species present the ability to reproduce vegetatively via root suckers in a significant manner, detectable through genetically identical individuals?; 2) Can land use and management affect the quantity of root suckers? To answer these questions, we quantified the occurrence of genetically identical individuals in populations of *D. alata* and *C. brasiliense* in areas presenting different land use and management conditions with microsatellite markers. We expected that: 1) both species present the capacity of root sprouting in a significant manner, since clonal reproduction is a common strategy for demographic persistence of native species in the Cerrado, an environment with historical disturbance regime; 2) land use and management lead to increased quantity of clones, as injuries, such as those caused by cattle grazing and vegetation thinning, can promote root suckering.

2. Methodology

Study areas and sampling

The study areas are located in the Cerrado, which presents tropical climate with dry winter and rainy summer seasons (Aw - Köppen). Six populations per species were selected to form three geographical units composed of paired populations, consisting of one population in native Cerrado vegetation and one in pastureland (Figure 1, Table 1). This pairwise population design was used to control for history and geography on genotypic diversity in each geographical unit. *D. alata* and *C. brasiliense* trees are maintained in pasturelands due to the economic value of its nuts and fruits. The populations analyzed here were previously studied on the demographic consequences of land use and management (Giroldo & Scariot 2015; Ferreira 2016).

Areas present different land use and management. Cattle ranching scenarios in the Cerrado can vary significantly in intensity of used technology, cattle density and frequency and intensity of vegetation thinning - a practice that aims to reduce unwanted vegetation in order to increase establishment of exotic African grasses used for cattle grazing (Klink & Moreira 2002; Giroldo & Scariot 2015). To estimate the disturbance caused by cattle ranching and vegetation thinning, we used a modified interaction matrix (Leopold *et al.* 1971) to construct a 0-9 rank index for these variables, as described in chapter I (page 6). Cattle ranching and vegetation thinning indexes are presented in Figure 1.

In each population, ~ 30 adults with diameter at 30 cm above soil (D_{30cm}) ≥ 10 cm and 70 to 100 young individuals of different sizes (hereinafter referred to as juveniles) were sampled. The adults were sampled randomly with a minimum of 30 m distance from each other. For *D. alata*, 10 juveniles with height ≤ 20 cm were sampled in a 10 m radius around each of 10 randomly chosen previously sampled adult tree, in a total of ~ 100 juveniles per population. For *C. brasiliense*, juveniles of such sizes could not be found around trees; thus, ~ 70 juveniles with D_{30cm} ≤ 5 cm were sampled in clusters, a common distribution presented by the species. This sampling strategy was chosen in order to maximize clonal detection; it can be classified as a stratified random sampling design, which is considered appropriate for heterogeneous populations (Arnaud-Haond *et al.* 2007), such as those here

investigated. Fresh leaves were collected from each of the 1.446 sampled individuals, which were also used for analyses in chapters I and II.

Laboratory analyses

Laboratory analyses were performed as described in chapters I (page 8) and II (page 31).

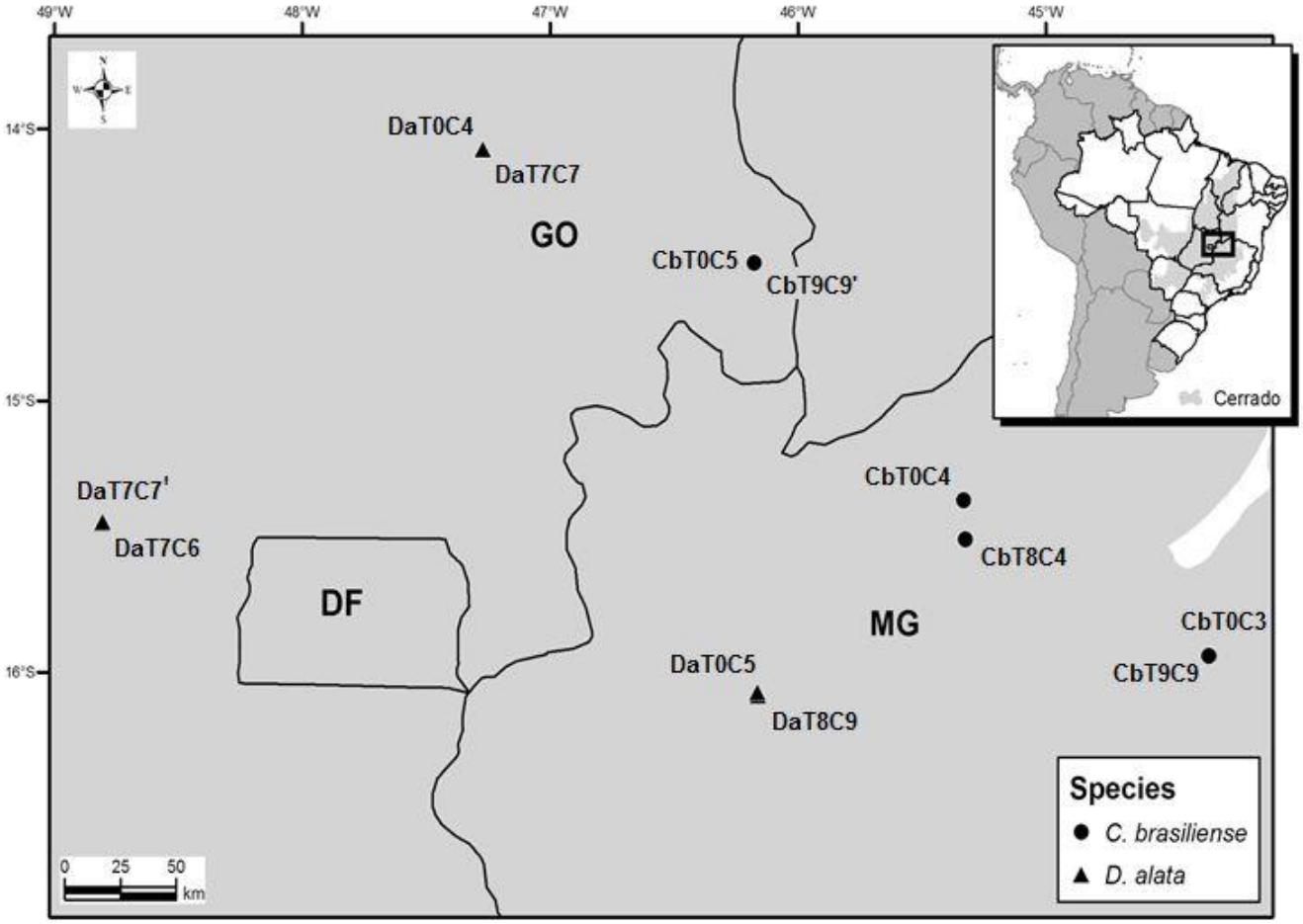


Figure 1 Location of 12 study areas with *D. alata* and *C. brasiliense* populations in the Brazilian Cerrado. Due to the scale used, some points represent more than one area. Areas are named after species (Da for *D. alata* and Cb for *C. brasiliense*), vegetation thinning (T) and cattle ranching (C) indexes, which summarize severity, duration and area of disturbances. ` differentiates populations with equal names.

Table 1. Localization and physiognomy of areas with populations of *C. brasiliense* and *D. alata*.

	Population	Longitude	Latitude	Physiognomy
C. brasilliense	CbT0C4	-45.332	-15.370	Dense Cerrado
	CbT8C4	-45.325	-15.511	Pastureland
	CbT0C5	-46.178	-14.496	Typical Cerrado
	CbT9C9'	-46.176	-14.494	Pastureland
	CbT9C9	-44.338	-15.941	Pastureland
	CbT0C3	-44.343	-15.942	Dense Cerrado
D. alata	DaT0C4	-47.273	-14.076	Dense Cerrado
	DaT7C7	-47.271	-14.073	Pastureland
	DaT0C5	-46.161	-16.085	Typical Cerrado
	DaT8C9	-46.162	-16.075	Pastureland
	DaT7C6	-48.809	-15.450	Woodland
	DaT7C7'	-48.805	-15.447	Pastureland

Data analyses

To distinguish individuals in a reliable manner and to estimate the necessary number of locus to do so, we rounded off alleles, tested for null alleles and for linkage equilibrium, as described in chapter 1 (page 8).

Test for clonality through detection of identical multilocus genotypes (MLG) was executed with the RClone package (Bailleul *et al.* 2016) in R software (R Core Team 2016), which is a R version of the GenClone program (Arnaud-Haond & Belkhir 2007). To test for reliability of loci set used for optimal MLG recognition and to select the numbers of loci to be used in a cost-efficiently manner (Arnaud-Haond *et al.* 2005), genotype accumulation curves were created for each species based on 1,000 resamplings in a Monte Carlo procedure. In addition to the number of distinct MLGs, the clonal diversity index, as proposed by Dorken & Eckert (2001), was estimated per population, which varies from zero (all individuals have the same MLG) to one (all individuals have different MLGs) and corrects for different sampling sizes, as clarified by Arnaud-Haond et al. (2007).

We also performed an identity analysis to confirm the existence of matching genotypes, using for this a minimum number of matching loci of n-1 (n being the number of loci used for each species), and allowing for up to 1 mismatch through fuzzy matching. The Polymorphic Information Content (PIC) was estimated for each locus and for both loci set used. The usefulness of both loci set for identity analysis was examined through the estimation of combined non-exclusion probability. These

analyses were executed on Cervus software (Marshall *et al.* 1998; Kalinowski *et al.* 2007).

To understand if land use and management present associations with the genetic diversity of populations of *D. alata* and *C. brasiliense* in the absence of vegetative reproduction, we used a generalized linear model approach (Lindsey 1997). For this, we estimated basic genetic diversity parameters, as described in chapter 1 (page 8). Two genetic diversity parameters were used as response variables in distinct analyses: allelic richness and expected heterozygosity. We used cattle ranching and vegetation thinning indexes as explanatory variables, which summarize land use and management. We used the gaussian distribution as the variance function and identity as the link function, which resulted in good fit. We tested all possible combinations of explanatory variables as models, which were tested against null models through Chi-square test. This was performed through stats and boot (Davison & Hinkley 1997) packages in R (R Core Team 2016).

3. Results

No loci combination showed linkage disequilibrium in more than 50% of *D. alata* populations, indicating that such associations are most likely related to specific populations and not to loci combination. For *C. brasiliense*, only one loci pair showed linkage disequilibrium in a significant proportion of populations (83%), which we opted to maintain in the analysis. No locus presented evidence for null alleles in more than 50% of the populations analyzed for both species, indicating that excess of homozygosity in specific populations might have enabled overestimation of null allele frequencies.

Test for clonality through RClone did not detect any clone in any of the populations, for both *D. alata* and *C. brasiliense* (Table 2). Consequently, clonal diversity index values were equal to one for all populations (Table 2), meaning that all individuals in the populations have different MLGs. No matching genotypes were found through identity analyses for neither species (Table 3).

The estimated PIC showed that the loci set used for *C. brasiliense* is considerably polymorphic, while the set used for *D. alata* is only moderately polymorphic and, thus, not so informative (Table 4). In spite of this, the MLG accumulation curves indicate that both loci set used were exceedingly reliable, since

just 8 out of 15 loci for *D. alata* and 4 out of 9 loci for *C. brasiliense* were able to differentiate all MLGs efficiently (Figure 2). Combined non-exclusion probabilities for identity analysis values were 3.25E-12 and 8.88E-13 for *D. alata* and *C. brasiliense*, respectively, also indicating high consistency.

Cattle ranching and vegetation thinning did not present significant associations with allelic richness or expected heterozygosity for *D. alata* or *C. brasiliense* adult and juvenile populations; no model was statistically different from the null model (results not shown).

Table 2. Number of sampled units (n), number of unique multilocus genotypes (MLG) and clonal diversity index (R) of *D. alata* and *C. brasiliense* populations.

	Population	n	MLG	R
D. alata	DaT0C4	137	137	1
	DaT7C7	157	157	1
	DaT0C5	138	138	1
	DaT8C9	146	146	1
	DaT7C6	125	125	1
	DaT7C7'	132	132	1
	Average	139.2	139.2	-
C. brasiliense	CbT0C4	100	100	1
	CbT8C4	101	101	1
	CbT0C5	101	101	1
	CbT9C9'	100	100	1
	CbT9C9	107	107	1
	CbT0C3	102	102	1
	Average	101.8	101.8	-

Table 3. Number of individuals compared, pairwise comparison and matching genotypes found through identity analysis for all individuals of *D. alata* and *C. brasiliense.*

	D. alata	*C. brasiliense*
Number of individuals compared	835	611
Number of pairwise comparison	348195	186355
Number of matching genotypes found	0	0

Table 4. Polymorphic Information Content (PIC) for each locus and loci set used for *D. alata* and *C. brasiliense.*

D. alata		*C. brasiliense*	
Locus	PIC	Locus	PIC
DaE06	0.297	Cb03	0.798
DaE12	0.448	Cb05	0.853
DaE20	0.241	Cb06	0.894
DaE34	0.792	Cb09	0.784
DaE41	0.528	Cb11	0.837
DaE46	0.529	Cb12	0.815
DaE63	0.473	Cb13	0.535
DaE67	0.617	Cb20	0.840
Do06	0.582	Cb23	0.823
Do08	0.578		
Do17	0.779		
Do20	0.474		
Do24	0.756		
Do25	0.742		
Do35	0.744		
Average	0.572		0.798

4. Discussion

Marker sets used for both species were very reliable, and only approximately half of the loci used in each set was necessary to distinguish all MLGs efficiently. Although the set used for *D. alata* is not exceptionally polymorphic, this is simply a natural consequence of the low genetic diversity pattern for the species (Collevatti *et al.* 2013), and not a limitation of the markers used. Considering that working with molecular markers is expensive and that the sets used should be efficient in distinguishing individuals and not necessarily high in number of markers, the choice of markers should be optimized in a cost-efficient manner (Arnaud-Haond *et al.* 2005). Thus, based on the genotype accumulation curves and the PIC for individual markers, we suggest the following sets of markers for future studies that aim to differentiate individuals cost-efficiently: DaE34, Do17, Do24, Do35, Do25, DaE67, Do06 and Do08 for *D. alata*; Cb06, Cb05, Cb20 and Cb11 for *C. brasiliense*.

Our results showed that none of the populations of neither species presented genetically identical individuals, though these are Cerrado native species and in spite of the existence of disturbances that are proven to promote root suckering, such as cattle grazing and vegetation thinning. Our results also imply that land use and management did not inflict root sprouting in the studied populations of *D. alata* and *C. brasiliense*. Sampling effort was not an issue that could have restricted clone

detection. Although optimal sampling effort can be very complex to determine, sample sizes in clonal organisms studies do not normally go beyond 30 - 50 individuals per locality (Arnaud-Haond *et al.* 2007), while our sample sizes per area varied from 100 to 157. In addition, our results demonstrate that the number of loci used was more than sufficient to differentiate individuals.

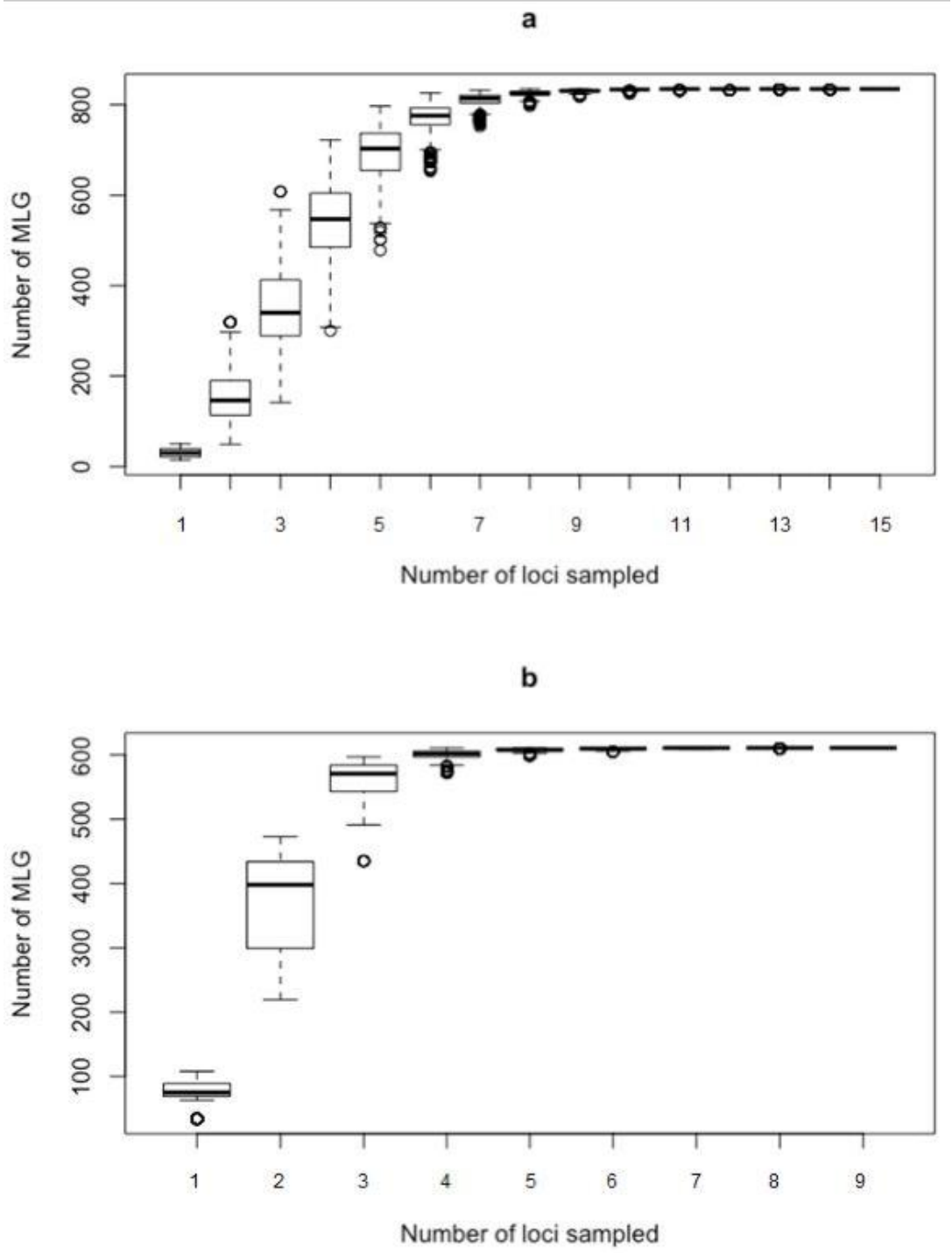


Figure 2 Number of MLGs accumulated with number of loci for *D. alata* (a) and *C. brasiliense* (b) populations.

Disturbance frequency and severity (measured by lost biomass) and site productivity (soil moisture and fertility) are important determinants of sprouting frequency, because these factors greatly affect the size of storage organs (Iwasa & Kubo 1997; Bellingham & Sparrow 2000). This is consistent with the reality of the Cerrado, since many plants in this region present underground storage organs (Miranda *et al.* 2002; Medeiros *et al.* 2008; Simon *et al.* 2009). The frequency and severity of disturbances were high enough at least in the populations sampled in pastures, comprising half of the populations analyzed, in a way that allowed for the development of storage organs and, eventually, root-sprouting. Cattle grazing and vegetation thinning are continuous activities in these populations and vegetation thinning usually removes most of the aboveground biomass of individuals of both species, characterizing considerable frequency and severity of disturbance. Also, most Cerrado soils are acid and poor in nutrients and this region undergoes a dry season for roughly half of the time; thus, site productivity should not be a constraint to the development of storage organs.

Plants can present different adaptive traits to persist demographically in recurrently disturbed environments: some cope by resprouting, usually presenting poor regeneration from seeds; other species, rely on a large production of seeds and do not form below-ground reserves (Bellingham & Sparrow 2000; Bond & Midgley 2001; Hoffmann & Moreira 2002; Klimešová & Klimeš 2007). *D. alata* presents a considerably high production of fruits, which fluctuates seasonally and depends of site, varying from no fruits up to 5,000 fruits per tree, with rough average estimates of 1,500 fruits per plant (Sano *et al.* 2004; Sano & Simon 2008). *C. brasiliense* also presents variable production of fruits, with averages that vary from 11 to 132, reaching values of up to 2,160 fruits per plant (Santana & Naves 2003; Zardo & Henriques 2011; Ferreira *et al.* 2015). These facts corroborate with our results, indicating that both *D. alata* and *C. brasiliense* seem to rely on production of fruits to persist demographically rather than on vegetative reproduction.

The ability to sprout is common and could be the ancestral state in woody angiosperms (Bond & Midgley 2001). Most clades present component species with varied sprouting abilities, which is a feature that is not conserved along phylogenetic lineages (Vesk & Westoby 2004). This could be because different environmental conditions lead to selective pressure against the trade-offs between persistence and

recruitment, leading to evolutionary switches between sprouting and non-sprouting, which have occurred repeatedly throughout evolution (Bond & Midgley 2001).

This is corroborated by studies in the Cerrado. Adaptations to fire such as root-sprouting occurred in this region when C4 flammable grasses became dominant, leading to the development of a fire regime, and savannas expanded around the world, less than 10 million years ago (Simon *et al.* 2009; Simon & Pennington 2012). Most Cerrado species belong to genera that are not restricted to this region, with species from fire-free biomes that do not present fire-adapted features, which were originated independently and in parallel many times (Simon *et al.* 2009; Simon & Pennington 2012). Along with other evidences, this information led these authors to believe that the evolutionary barrier to entry the Cerrado is weak, due to the ease of evolution of adaptations to fire regimes. This probably describes accurately the cases of *D. alata* and *C. brasiliense*. Both *Dipteryx* and *Caryocar* genera have species in many different biomes, occurring predominantly in the Amazon rainforest (Souza Neto 2012; Pinto *et al.* 2014), which is likely the source biome for *C. brasiliense* and *D. alata*.

However, more recent historical climatic changes in the last glaciation period during the late Pleistocene (last glaciation maximum - 21,000 years ago) shows that the distribution of *D. alata* and *C. brasiliense* went through range shifts (Collevatti *et al.* 2012, 2013). During this period, the climate in the Cerrado region was drier, which led to an expansion of this savanna. For *D. alata*, this led to population extinction or shrinkage in most part of the species distribution in central Brazil, followed later (6,000 years ago) by population expansion, which significantly affected the patterns of genetic diversity of the species (Collevatti *et al.* 2013). For *C. brasiliense*, range shifts also occurred, but in a much less pronounced way; the drier climate led to range retraction with population subdivision, consistent with the “multiple refugia” theory, with a slight westward migration, resulting in a loss of chloroplast lineages (Collevatti *et al.* 2012). These studies indicate that *D. alata* could be considerably less adapted to drier climates in comparison to *C. brasiliense*, which is corroborated by current distribution of both species, since *D. alata* is more restricted to forest physiognomies of the Cerrado, which present moister soils. This, in turn, might indicate that *D. alata* could possibly not have been successful in evolving traits that enable adaptation to drier climates, such as root sprouting, which is considered to be

a feature associated not only to adaptation to fire, but also to drought (Simon & Pennington 2012).

In conclusion, our findings indicate the opposite of our expectations: 1) *D. alata* and *C. brasiliense* do not present the ability to reproduce clonally in a significant manner; 2) land use and management intensity do not inflict detectable root sprouting in these species. The history of *D. alata* in the last glaciation period corroborates with our results, indicating the possibility of this species not having acquired the ability to reproduce clonally at all. As for *C. brasiliense*, although we cannot refute the possibility that it might have evolved the ability to root sprout, our results indicate that this feature has not become an important ecological trait for the demographic persistence of the species.

In situations where resources are limited and reproduction possibilities are restricted, such as in extreme disturbances, absence of clonal reproduction could lead to reductions in genetic diversity in future populations, possibly causing bottlenecks. This could be especially threatening to *D. alata* populations, since the species already has low levels of genetic diversity due to historic events (Collevatti *et al.* 2013). However, both *D. alata* and *C. brasiliense* are trees with: long lifespan, allowing for persistence; high levels of phenotypic plasticity, facilitating adaptation; and large size and high level of pollen and seed production, generating gene flow that will reduce loss of genetic diversity (Vranckx *et al.* 2012). Thus, serious negative consequences of not reproducing vegetatively are unlikely to hinder their persistence. This is corroborated by our generalized linear model analyses results, which show that *D. alata* and *C. brasiliense* adult and juvenile populations were able to withstand the effects of cattle ranching and vegetation thinning without significant losses of genetic diversity.

Supplementary Material

Supplementary Material 1. Genetic diversity parameters for the adult and juvenile *D. alata* and *C. brasiliense* populations.

	Size class	Population	n	A	Ar	He	Ho	f (95% CI)
D. alata	Adult	DaT0C4	16.1	4.8000	5.8667	0.5513	0.4495	0.1886* (0.0276 to 0.3673)
		DaT7C7	23.1	5.1333	5.6475	0.5895	0.5267	0.1054 (-0.0691 to 0.2860)
		DaT0C5	26.2	4.0000	4.4997	0.5175	0.4389	0.1538 (-0.0061 to 0.3162)
		DaT8C9	24.7	4.0000	4.6058	0.5171	0.4297	0.1708 (-0.0185 to 0.3550)
		DaT7C6	26.6	4.7333	4.9626	0.5476	0.4961	0.0945 (-0.1160 to 0.3155)
		DaT7C7'	27.1	4.6000	4.7946	0.5564	0.6211	-0.1188 (-0.2810 to 0.0393)
	Juvenile	DaT0C4	105.	6.6667	5.4279	0.5244	0.4309	0.1789* (0.0279 to 0.3360)
		DaT7C7	102.	7.2667	5.8419	0.6102	0.4960	0.1876* (0.0454 to 0.3235)
		DaT0C5	94.5	5.1333	4.7687	0.4977	0.4295	0.1373 (-0.0435 to 0.3143)
		DaT8C9	101.	7.0000	5.5304	0.5668	0.4377	0.2284* (0.0753 to 0.3882)
		DaT7C6	81.9	5.7333	5.1059	0.5005	0.4704	0.0604 (-0.0966 to 0.2138)
		DaT7C7'	89.8	6.6000	6.0721	0.5443	0.4868	0.1060 (-0.0489 to 0.2452)
		Average	59.9	5.4722	5.2603	0.5436	0.4761	0.1244
C. brasiliense	Adult	CbT0C4	28.4	8.8889	9.7778	0.7927	0.7510	0.0535 (-0.0190 to 0.1316)
		CbT8C4	25.9	9.1111	9.8889	0.7806	0.7598	0.0269 (-0.0856 to 0.1364)
		CbT0C5	27.9	10.1111	10.888	0.8235	0.7915	0.0395 (-0.0253 to 0.1047)
		CbT9C9'	28.4	10.3333	11.111	0.7818	0.7598	0.0285 (-0.0731 to 0.1453)
		CbT9C9	28.1	8.7778	9.5556	0.7925	0.8191	-0.0342 (-0.1532 to 0.0466)
		CbT0C3	36.4	8.2222	8.5811	0.7001	0.7307	-0.0443 (-0.1224 to 0.0482)
	Juvenile	CbT0C4	64.4	11.3333	10.502	0.7919	0.7111	0.1026* (0.0030 to 0.2159)
		CbT8C4	66.4	9.5556	8.6787	0.7489	0.7366	0.0165 (-0.0592 to 0.0939)
		CbT0C5	67.2	12.8889	11.450	0.8283	0.7881	0.0488 (-0.0030 to 0.1025)
		CbT9C9'	66.1	11.4444	10.501	0.7850	0.7748	0.0129 (-0.0830 to 0.1235)
		CbT9C9	61.1	10.7778	9.5148	0.6969	0.6465	0.0728 (-0.0169 to 0.1586)
		CbT0C3	56.4	10.1111	9.7892	0.7658	0.7553	0.0135 (-0.1271 to 0.1385)
		Average	46.4	10.1296	10.020	0.7740	0.7520	0.0281

n = sample size; A = allele number per locus; Ar = allelic richness per locus; He = expected heterozygosity; Ho = observed heterozygosity; f = inbreeding coefficient. * Significantly different from zero based on 95% confidence intervals estimated by 10,000 bootstraps.

Considerações finais

As populações adultas de *C. brasiliense* e *D. alata* apresentaram algumas semelhanças quanto à importância de variáveis para a diversidade genética intrapopulacional: em ambas as espécies, foram detectadas associações importantes entre diversidade genética e as variáveis índice de criação de gado e conteúdo de areia no solo, além do extrativismo de frutos não ter sido importante. A associação negativa entre criação de gado e diversidade genética e a ausência de associação entre extrativismo de frutos e diversidade genética para ambas as espécies foram os resultados mais relevantes em termos práticos deste trabalho, indicando a necessidade de construir políticas públicas que permitam a conservação genética destas espécies em ambientes com criação de gado e estimulem o extrativismo de frutos como estratégia de conservação. As populações avaliadas das duas espécies, contudo, também apresentaram algumas diferenças quanto à importância de variáveis para a diversidade genética intrapopulacional: enquanto as populações de *C. brasiliense* apresentaram um gradiente latitudinal de diversidade genética, aquelas de *D. alata* apresentaram um gradiente altitudinal. É possível que restrições na distribuição de *D. alata* durante a última glaciação e consequente redução de sua diversidade genética tenham atenuado um gradiente latitudinal de diversidade adquirido por múltiplos processos que atuaram durante milhões de anos, enquanto tal situação não ocorreu para *C. brasiliense*. A existência de um gradiente altitudinal apenas para *D. alata*, por sua vez, pode estar relacionada com a maior divergência genética interpopulacional apresentada por esta espécie em comparação com *C. brasiliense*, visto que o isolamento populacional é um dos principais fatores que podem levar à existência de gradientes altitudinais de diversidade genética. Além disso, associações importantes foram detectadas entre diversidade genética e densidade populacional para as populações de *C. brasiliense,* enquanto isso não foi observado para as populações de *D. alata*. Apesar de *D. alata* apresentar certa agregação demográfica, essa característica ocorre de maneira muito marcante em populações de *C. brasiliense*, possivelmente acarretando em consequências genéticas mais acentuadas para esta espécie. Assim, diferenças quanto aos resultados para *D. alata* e *C. brasiliense* estão relacionadas a características ecológicas e evolutivas intrínsecas de cada espécie, o que, por sua vez, evidencia como a conservação e manejo de populações naturais

devem ser planejados considerando particularidades de espécies e populações, e não de maneira generalizada.

Nenhuma das espécies avaliadas apresentou reprodução clonal, um resultado surpreendente considerando as diversas afirmações existentes na literatura a respeito da ampla ocorrência de reprodução vegetativa em espécies do Cerrado. Tal premissa, assim, necessita de revisão, o que deve ser alcançado através do desenvolvimento de estudos com o objetivo de determinar claramente se demais espécies do Cerrado apresentam reprodução vegetativa. Para isto, contudo, é necessário que não se confunda reprodução vegetativa com outros tipos de brotação. Assim, a caracterização de reprodução clonal deve ser feita apenas para plantas com caules distanciados da planta mãe. Além disso, tal caracterização deve ser feita preferencialmente através de identificação genotípica, o que irá reduzir erros como: não identificação de indivíduos clonais por estes não mais apresentarem conexão física com a planta mãe; identificação de plântulas como indivíduos clonais devido à aparente inexistência de sementes ou grande profundidade da raiz.

Referências bibliográficas

Adams RI, Hadly EA (2012) Genetic diversity within vertebrate species is greater at lower latitudes. *Evolutionary Ecology*, **27**, 133–143.

Aguilar R, Quesada M, Ashworth L, Herrerias-Diego Y, Lobo J (2008) Genetic consequences of habitat fragmentation in plant populations: susceptible signals in plant traits and methodological approaches. *Molecular Ecology*, **17**, 5177–5188.

Akaike H (1998) *Information Theory and an Extension of the Maximum Likelihood Principle*. Springer New York.

Allendorf FW (1986) Genetic drift and the loss of alleles versus heterozygosity. *Zoo Biology*, **5**, 181–190.

Allendorf FW, England PR, Luikart G *et al.* (2008) Genetic effects of harvest on wild animal populations. *Trends in ecology & evolution*, **23**, 327–37.

Allendorf F, Luikart G (2007) *Conservation and the Genetics of Populations*. Blackwell Publishing.

Almeida S, Proença C, Sano S, Ribeiro J (1998) *Cerrado: espécies vegetais úteis*. Embrapa Cerrados, Planaltina.

Araujo F (1995) A review of *Caryocar brasiliense* (Caryocaraceae)—an economically valuable species of the central Brazilian cerrados. *Economic Botany*, **49**, 40–48.

Armando M (1994) O impacto do fogo na rebrota de algumas espécies de árvores do Cerrado. Universidade de Brasília.

Arnaud-Haond S, Alberto F, Teixeira S *et al.* (2005) Assessing genetic diversity in clonal organisms: Low diversity or low resolution? Combining power and cost efficiency in selecting markers. *Journal of Heredity*, **96**, 434–440.

Arnaud-Haond S, Belkhir K (2007) GENCLONE: A computer program to analyse genotypic data, test for clonality and describe spatial clonal organization. *Molecular Ecology Notes*, **7**, 15–17.

Arnaud-Haond S, Duarte CM, Alberte F, Serrao EA (2007) Standardizing methods to adress clonality in population studies. *Molecular Ecology*, **16**, 5115–5139.

Avise J (2008) The history, purview, and future of conservation genetics. In: *Conservation biology: evolution in action.* (eds Carrol S, Fox C), pp. 5–15. Oxford University Press, New York.

Bailleul D, Stoeckel S, Arnaud-Haond S (2016) RClone: A package to identify MultiLocus Clonal Lineages and handle clonal data sets in R. *Methods in Ecology and Evolution*, **7**, 966–970.

Banks SC, Cary GJ, Smith AL *et al.* (2013) How does ecological disturbance influence genetic diversity? *Trends in Ecology & Evolution*, **28**, 670–679.

Bellingham PJ, Sparrow AD (2000) Resprouting as a life history strategy in woody plant communities. *Oikos*, **89**, 409–416.

Blackshaw J, Blackshaw A (1994) Heat stress in cattle and the effect of shade on production and behaviour: a review. *Australian Journal of Experimental*

Agriculture, **34**, 285–295.

Bond WJ, Midgley JJ (2001) Ecology of sprouting in woody plants: the persistence niche. *Trends in Ecology & Evolution*, **16**, 45–51.

Bond WJ, Midgley JJ (2003) The evolutionary ecology of sprouting in woody plants. *International Journal of Plant Sciences*, **164**, S103–S114.

Broadhurst LM, Lowe A, Coates DJ *et al.* (2008) Seed supply for broadscale restoration: maximizing evolutionary potential. *Evolutionary applications*, **1**, 587–97.

Burnham K, Anderson D (2002) *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer New York.

Cardoso M, Marcuzzo F, Barros J (2014) Classificação climática de Köppen-Geiger para o estado de Goiás e o Distrito Federal. *Acta Geográfica*, **8**, 40–55.

Carrazza R, Ávila J (2010) *Aproveitamento integral do fruto do Baru (*Dipteryx alata*)*. Instituto Sociedade, População e Natureza, Brasíla.

Charpentier A (2001) Consequences of clonal growth for plant mating. *Evolutionary Ecology*, **15**, 521–530.

Collevatti R, Brondani R, Grattapaglia D (1999) Development and characterization of microsatellite markers for genetic analysis of a Brazilian endangered tree species *Caryocar brasiliense*. *Heredity*, **83**, 748–756.

Collevatti RG, Estolano R, Garcia SF, Hay JD (2010) Short-distance pollen dispersal and high self-pollination in a bat-pollinated neotropical tree. *Tree Genetics & Genomes*, **6**, 555–564.

Collevatti R, Grattapaglia D, Hay J (2001a) High resolution microsatellite based analysis of the mating system allows the detection of significant biparental inbreeding in *Caryocar brasiliense*, an endangered tropical tree species. *Heredity*, **86**, 60–67.

Collevatti R, Grattapaglia D, Hay J (2001b) Population genetic structure of the endangered tropical tree species *Caryocar brasiliense*, based on variability at microsatellite loci. *Molecular Ecology*, **10**, 349–356.

Collevatti R, Grattapaglia D, Hay J (2002) Evidences for multiple maternal lineages of *Caryocar brasiliense* populations in the Brazilian Cerrado based on the analysis of chloroplast DNA sequences and. *Molecular Ecology*, **12**, 105–115.

Collevatti RG, Hay JD (2011) Kin structure and genotype-dependent mortality: A study using the neotropical tree *Caryocar brasiliense*. *Journal of Ecology*, **99**, 757–763.

Collevatti RG, Lima-Ribeira MS, Souza-Neto AC *et al.* (2012) Recovering the demographical history of a Brazilian Cerrado tree species *Caryocar brasiliense*: coupling ecological niche modeling and coalescent analyses. *Natureza & Conservação*, **10**, 169–176.

Collevatti RG, Nabout JC, Diniz-Filho JAF (2011) Range shift and loss of genetic diversity under climate change in *Caryocar brasiliense*, a Neotropical tree species. *Tree Genetics & Genomes*, **7**, 1237–1247.

Collevatti RG, Telles MPC, Nabout JC, Chaves LJ, Soares TN (2013) Demographic

history and the low genetic diversity in *Dipteryx alata* (Fabaceae) from Brazilian Neotropical savannas. *Heredity*, **111**, 97–105.

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

Davies ID, Cary GJ, Landguth EL, Lindenmayer DB, Banks SC (2016) Implications of recurrent disturbance for genetic diversity. *Ecology and Evolution*, **6**, 1181–1196.

Davison A, Hinkley D (1997) *Bootstrap Methods and Their Application*. Cambridge University Press, Cambridge.

DeSalle R, Amato G (2004) The expansion of conservation genetics. *Nature Reviews Genetics*, **5**, 702–712.

Dick CW, Hardy OJ, Jones FA, Petit RJ (2008) Spatial scales of pollen and seed-mediated gene flow in tropical rain forest trees. *Tropical Plant Biology*, **1**, 20–33.

Diniz-Filho JAF, Nabout JC, Bini LM *et al.* (2009) Niche modelling and landscape genetics of *Caryocar brasiliense* ("Pequi" tree: Caryocaraceae) in Brazilian Cerrado: an integrative approach for evaluating central–peripheral population patterns. *Tree Genetics & Genomes*, **5**, 617–627.

Dorken ME, Eckert CG (2001) Severely reduced sexual reproduction in northern populations of a clonal plant, *Decodon verticillatus* (Lythraceae). *Journal of Ecology*, **89**, 339–350.

Doyle J, Doyle J (1987) Isolation of plant DNA from fresh tissue. *Focus*, **12**, 13–15.

Eckert CG (2001) The loss of sex in clonal plants. *Evolutionary Ecology*, **15**, 501–520.

Ehinger M, Fontanillas P, Petit E, Perrin N (2002) Mitochondrial DNA variation along an altitudinal gradient in the greater white-toothed shrew, *Crocidura russula*. *Molecular Ecology*, **11**, 939–945.

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.

Ferreira J (2016) Impactos causados pelo extrativismo, uso da terra e manejo na persistência de populações de *Dipteryx alata* Vog. (Baru) no Cerrado. Universidade de Brasília.

Ferreira G, Naves R, Chaves L, Veloso. VR, Souza E (2015) Production of fruits from natural populations of Pequi trees in the state of Goiás, Brazil. *Revista Brasileira de Fruticultura*, **37**, 121–129.

Finkeldey R, Ziehe M (2004) Genetic implications of silvicultural regimes. *Forest Ecology and Management*, **197**, 231–244.

Fischer M, Bossdorf O, Gockel S *et al.* (2010) Implementing large-scale and long-term functional biodiversity research: The Biodiversity Exploratories. *Basic and Applied Ecology*, **11**, 473–485.

Forzza R, Baumgratz J, Bicudo C *et al.* (2010) *Catálogo de Plantas e Fungos do Brasil*. Andrea Jakobsson Estúdio, Instituto de Pesquisas Jardim Botânico do

Rio de Janeiro, Rio de Janeiro.

Françoso RD, Brandão R, Nogueira CC *et al.* (2015) Habitat loss and the effectiveness of protected areas in the Cerrado Biodiversity Hotspot. *Natureza & Conservação*, **13**, 35–40.

Frankham R (2010) Challenges and opportunities of genetic approaches to biological conservation. *Biological Conservation*, **143**, 1919–1927.

Fraser EC, Lieffers VJ, Landhäusser SM (2004) Wounding of aspen roots promotes suckering. *Canadian Journal of Botany*, **82**, 310–315.

Furley PA (1999) The nature and diversity of neotropical savanna vegetation with particular reference to the Brazilian cerrados. *Global Ecology and Biogeography*, **8**, 223–241.

Furley PA, Ratter JA (1988) Soil resources and plant communities of the central Brazilian Cerrado and their development. *Journal of Biogeography*, **15**, 97–108.

Gaoue OG, Lemes MR, Ticktin T, Sinsin B, Eyog-Matig O (2014) Non-timber forest product harvest does not affect the genetic diversity of a tropical tree despite negative effects on population fitness. *Biotropica*, **46**, 756–762.

Gibbs JP (2001) Demography versus habitat fragmentation as determinants of genetic variation in wild populations. *Biological Conservation*, **100**, 15–20.

Giroldo AB, Scariot A (2015) Land use and management affects the demography and conservation of an intensively harvested Cerrado fruit tree species. *Biological Conservation*, **191**, 150–158.

Goodell K, Elam D, Nason J, Ellstrand N (1997) Gene flow among small populations of a self-incompatible plant: an interaction between demography and genetics. *American journal of botany*, **84**, 1362.

Goudet J (2005) HIERFSTAT, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes*, **5**, 184–186.

Gribel R (1986) Ecologia da polinização e da dispersão de *Caryocar brasiliense* Camb. (Caryocaraceae) na região do Distrito Federal. Universidade de Brasília.

Gribel R, Hay J (1993) Pollination ecology of *Caryocar brasiliense* (Caryocaraceae) in Central Brazil Cerrado vegetation. *Journal of Tropical Ecology*, **9**, 199–211.

Gurevitch J, Scheiner S, Fox G (2006) *The Ecology of Plants*. Sinauer Associates, Sunderland.

Hahn T, Kettle CJ, Ghazoul J *et al.* (2012) Patterns of genetic variation across altitude in three plant species of semi-dry grasslands. *PloS one*, **7**, e41608.

Hardy O, Vekemans X (2002) SPAGeDI: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, **2**, 618–620.

Haridasan M (2000) Nutrição mineral de plantas nativas do cerrado. *Revista Brasileira de Fisiologia*, **12**, 54–64.

Hartl D, Clark A (1997) *Principles of Population Genetics*. Sinauer Associates, Sunderland.

Hedrick PW (2011) *Genetics of populations*. Jones and Barlett Publishers, Sudbury.

Hoebee SE, Menn C, Rotach P, Finkeldey R, Holderegger R (2006) Spatial genetic structure of *Sorbus torminalis*: The extent of clonal reproduction in natural stands of a rare tree species with a scattered distribution. *Forest Ecology and Management*, **226**, 1–8.

Hoffmann WA (1998) Post-burn reproduction of woody plants in a neotropical savanna: the relative importance of sexual and vegetative reproduction. *Journal of Applied Ecology*, **35**, 422–433.

Hoffmann W, Moreira A (2002) The role of fire in population dynamics of woody plants. In: *The Cerrados of Brazil: Ecology and Natural History of a Neotropical Savanna* (eds Oliveira P, Marquis R), pp. 159–177. Columbia University Press, New York.

Hughes AR, Inouye BD, Johnson MTJ, Underwood N, Vellend M (2008) Ecological consequences of genetic diversity. *Ecology Letters*, **11**, 609–623.

Iwasa Y, Kubo T (1997) Optimal size of storage for recovery after unpredictable disturbances. *Evolutionary Ecology*, **11**, 41–65.

Jugran AK, Bhatt ID, Rawal RS, Nandi SK, Pande V (2013) Patterns of morphological and genetic diversity of *Valeriana jatamansi* Jones in different habitats and altitudinal range of West Himalaya, India. *Flora - Morphology, Distribution, Functional Ecology of Plants*, **208**, 13–21.

Kalinowski S, Taper M, Marshall T (2007) Revising how the computer program cervus accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*, **16**, 1099–1106.

Karunaratne W, Edirisinghe J (2008) Diversity of bees at different altitudes in the Knuckles Forest Reserve. *Ceylon Journal of Science (Biological Sciences)*, **37**, 61–72.

Klimešová J, Klimeš L (2007) Bud banks and their role in vegetative regeneration – A literature review and proposal for simple classification and assessment. *Perspectives in Plant Ecology, Evolution and Systematics*, **8**, 115–129.

Klink C, Moreira A (2002) Past and current human occupation, and land use. In: *The Cerrados of Brazil: Ecology and Natural History of a Neotropical Savanna* (eds Oliveira P, Masquis R), pp. 69–88. Columbia University Press, New York.

Koop H (1987) Vegetative reproduction of trees in some European natural forests. *Vegetatio*, **72**, 103–110.

Kramer A, Ison J, Ashley M (2008) The paradox of forest fragmentation genetics. *Conservation Biology*, **22**, 878–885.

Laikre L (2010) Genetic diversity is overlooked in international conservation policy implementation. *Conservation Genetics*, **11**, 349–354.

Lambin E, Geist H, Lepers E (2003) Dynamics of land-use and land-cover change in tropical regions. *Annual Review of Environment and Resources*, **28**, 205–241.

Leberg PL (2008) Estimating allelic richness: Effects of sample size and bottlenecks. *Molecular Ecology*, **11**, 2445–2449.

Leite GLD, Nascimento AF, Alves SM *et al.* (2012) The mortality of *Caryocar brasiliense* in northern Minas Gerais State, Brazil. *Acta Scientiarum. Agronomy*, **34**, 131–137.

Lemke TO (1984) Foraging ecology of the long-nosed bat, *Glossophaga Soricina*, with respect to resource availability. *Ecology*, **65**, 538–548.

Leopold LB, Clarke F, Hanshaw B, Balsley J (1971) *A procedure for evaluating environmental impact*. US Dept. of the Interior, Washington.

Li J, Xu H, Song Y *et al.* (2016) Geography plays a more important role than soil composition on structuring genetic variation of pseudometallophyte *Commelina communis*. *Frontiers in plant science*, **7**, 1085.

Lindsey J (1997) *Applying Generalized Linear Models*. Springer, New York.

Loiselle B, Sork V, Nason J, Graham C (1995) Spatial genetic structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *American Journal of Botany*, **82**, 1420.

Loveless MD, Hamrick JL (1984) Ecological determinants of genetic structure in plant populations. *Annual Review of Ecology and Systematics*, **15**, 65–95.

Lowe AJ, Boshier D, Ward M, Bacles CFE, Navarro C (2005) Genetic resource impacts of habitat loss and degradation; reconciling empirical evidence and predicted theory for neotropical trees. *Heredity*, **95**, 255–273.

Mander Ü, Helming K, Wiggering H (2007) Multifunctional land use: meeting future demands for landscape goods and services. In: *Multifunctional Land Use* (eds Mander Ü, Wiggering H, Helming K), pp. 1–13. Springer-Verlag, Berlin, Heidelberg.

Marshall T, Slate J, Kruuk L, Pemberton J (1998) Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology*, **7**, 639–655.

Martin PR, McKay JK (2004) Latitudinal variation in genetic divergence of populations and the potential for future speciation. *Evolution*, **58**, 938–945.

Medeiros MB, Miranda HS, Bond WJ *et al.* (2008) Post-fire resprouting and mortality in Cerrado woody plant species over a three-year period. *Edinburgh Journal of Botany*, **65**, 53–68.

Melo DB, Diniz-Filho JAF, Oliveira G *et al.* (2011) Optimizing sampling efforts for *ex situ* conservation of genetic variability of *Dipteryx alata* Vogel. *BMC Proceedings*, **5**, P18.

Mengli Z, Willms WD, Bing H, Laroche A (2005) Effects of heavy grazing pressure on the random amplified polymorphic DNA marker diversity of mountain rough fescue (*Festuca campestris* Rydb.) in south western Alberta. *Canadian Journal of Plant Science*, **85**, 623–629.

Miranda H, Bustamante M, Miranda A (2002) The fire factor. In: *The Cerrados of Brazil: Ecology and Natural History of a Neotropical Savanna* (eds Marquis P, RJ O), pp. 51–68. Columbia University Press, New York.

MMA (2011) *Monitoramento do bioma Cerrado 2009 - 2010. Monitoramento do desmatamento nos biomas brasileiros por satélite: acordo de cooperação técnica MMA/IBAMA.* Ministério do Meio Ambiente, Brasília.

Moilanen A, Franco A, Early R *et al.* (2005) Prioritizing multiple-use landscapes for conservation: methods for large multi-species planning problems. *Proceedings of the Royal Society of London B: Biological Sciences*, **272**, 1885–1891.

Nabout J, Oliveira G, Magalhães M, Terribile L, Almeida F (2011) Global climate change and the production of “pequi” fruits (*Caryocar brasiliense*) in the Brazilian Cerrado. *Natureza & Conservação*, **9**, 55–60.

Neville H, Dunham J, Rosenberger A, Umek J, Nelson B (2009) Influences of wildfire, habitat size, and connectivity on trout in headwater streams revealed by patterns of genetic diversity. *Transactions of the American Fisheries Society*, **138**, 1314–1327.

Ohsawa T, Ide Y (2008) Global patterns of genetic variation in plant species along vertical and horizontal gradients on mountains. *Global Ecology and Biogeography*, **17**, 152–163.

Oksanen J, Kindt R, Legendre P *et al.* (2008) *The vegan package*.

Oliveira W de (2009) Ecologia populacional e extrativismo de frutos de *Caryocar brasiliense* Camb. no Cerrado no norte de Minas Gerais. Universidade de Brasília.

Oliveira W, Scariot A (2010) *Boas práticas de manejo para o extrativismo sustentável do pequi*. Embrapa Recursos Genéticos e Biotecnologia.

Oliveira MIB, Sigrist MR (2008) Fenologia reprodutiva, polinização e reprodução de *Dipteryx alata* Vogel (Leguminosae-Papilionoideae) em Mato Grosso do Sul, Brasil. *Revista Brasileira de Botânica*, **31**, 195–207.

Oosterhout C Van, Hutchinson W, Wills D, Shipley P (2004) MICRO- CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535–538.

Petit RJ, El Mousadik A, Pons O (1998) Identifying Populations for conservation on the basis of genetic markers. *Conservation Biology*, **12**, 844–855.

Pinto RB, Francisco VM da CR, Mansano V de F (2014) Morphological study of fruits, seeds and embryo in the tropical tribe Dipterygeae (Leguminosae-Papilionoideae). *Rodriguésia*, **65**, 89–97.

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

Pivello VR, Norton GA (1996) FIRETOOL: An expert system for the use of prescribed fires in Brazilian Savannas. *The Journal of Applied Ecology*, **33**, 348–356.

Prasanth V, Chandra S, Hoisington D, Jayashree B (2006) *AlleloBin - a program for allele binning of microsatellite markers based on the alogirithm of Idury and Cardon, 1997.* International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Delhi.

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

Pritchard J, Wen X, Falush D (2010) *Documentation for Structure software: Version 2.3*.

R Core Team (2016) *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna.

Rajora OP, Pluhar SA (2003) Genetic diversity impacts of forest fires, forest harvesting, and alternative reforestation practices in black spruce (*Picea mariana*). *Theoretical and Applied Genetics*, **106**, 1203–1212.

Ramos-Neto MB, Pivello VR (2000) Lightning fires in a Brazilian Savanna National Park: Rethinking management strategies. *Environmental Management*, **26**, 675–684.

Rao R, Hodgkin T (2002) Genetic diversity and conservation and utilization of plant genetic resources. *Plant Cell, Tissue and Organ Culture*, **68**, 1–19.

Ratnam W, Rajora OP, Finkeldey R *et al.* (2014) Genetic effects of forest management practices: Global synthesis and perspectives. *Forest Ecology and Management*, **333**, 52–65.

Ratter J, Bridgewater S, Ribeiro J (2003) Analysis of the floristic composition of the Brazilian Cerrado vegetation III: comparison of the woody vegetation of 376 areas. *Edinburgh Journal of Botany*, **60**, 57–109.

Ratter J, Ribeiro J, Bridgewater S (1997) The Brazilian Cerrado vegetation and threats to its biodiversity. *Annals of Botany*, **80**, 223–230.

Reatto A, Correia J, Spera S, Martins E (2008) Solos do bioma Cerrado: Aspectos pedológicos. In: *Cerrado: Ecologia e Flora, v.1.* (eds Sano S, Almeida S, Ribeiro J), pp. 107–149. Embrapa Informação Tecnológica.

Salazar A, Goldstein G (2014) Effects of fire on seedling diversity and plant reproduction (sexual vs. vegetative) in neotropical savannas differing in tree density. *Biotropica*, **46**, 139–147.

Sano S, Ribeiro J, Brito M (2004) *Baru: biologia e uso.* Embrapa Cerrados, Planaltina.

Sano S, Simon M (2008) Produtividade de baru (*Dipteryx alata* Vog.) em ambientes modificados, durante 10 anos. In: *Anais do II Simpósio Internacional sobre Savanas Tropicais*, pp. 1–7. Embrapa Cerrados.

Santana J, Naves R (2003) Caracterização de ambientes de Cerrado com alta densidade de Pequizeiros (*Caryocar brasiliense* Camb.) na região sudeste do estado de Goiás. *Pesquisa Agropecuária Tropical*, **33**, 1–10.

Schrey AW, Ashton KG, Heath S, McCoy ED, Mushinsky HR (2011a) Fire alters patterns of genetic diversity among 3 lizard species in Florida Scrub habitat. *The Journal of Heredity*, **102**, 399–408.

Schrey AW, Fox AM, Mushinsky HR, Mccoy ED (2010) Fire increases variance in genetic characteristics of Florida Sand Skink (*Plestiodon reynoldsi*) local populations. *Molecular Ecology*, **20**, 56–66.

Schrey AW, Grispo M, Awad M *et al.* (2011b) Broad-scale latitudinal patterns of genetic diversity among native European and introduced house sparrow (*Passer domesticus*) populations. *Molecular Ecology*, **20**, 1133–1143.

Shaanker R, Ganeshaiah K, Rao M, Aravind N (2004) Ecological consequences of forest use: From genes to ecosystem - A case study in the Biligiri Rangaswamy Temple Wildlife Sanctuary, South India. *Conservation and Society*, **2**, 347–363.

Shan D, Zhao M, Han B, Han G (2006) Examining the genetic diversity of *Stipa grandis* under various grazing pressures. *Acta Ecologica Sinica*, **26**, 3175–3182.

Shen D, Bo W, Xu F, Wu R (2014) Genetic diversity and population structure of the Tibetan poplar (*Populus szechuanica var. tibetica*) along an altitude gradient. *BMC genetics*, **15 Suppl 1**, S11.

Silva F, Assad E, Evangelista B (2008) Caracterização climática do bioma Cerrado. In: *Cerrado: Ecologia e Flora, v.1.* (eds Sano S, Almeida S, Ribeiro J), pp. 69–88. Embrapa Informação Tecnológica, Brasília.

Simon MF, Grether R, de Queiroz LP *et al.* (2009) Recent assembly of the Cerrado, a neotropical plant diversity hotspot, by *in situ* evolution of adaptations to fire. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 20359–20364.

Simon MF, Pennington T (2012) Evidence for adaptation to fire regimes in the tropical savannas of the Brazilian Cerrado. *International Journal of Plant Sciences*, **173**, 711–723.

Smith AL, Bull CM, Gardner MG, Driscoll DA (2014) Life history influences how fire affects genetic diversity in two lizard species. *Molecular Ecology*, **23**, 2428–2441.

Soares TN, Chaves LJ, de Campos Telles MP, Diniz-Filho JAF, Resende LV (2008a) Landscape conservation genetics of *Dipteryx alata* ("baru" tree: Fabaceae) from Cerrado region of central Brazil. *Genetica*, **132**, 9–19.

Soares TN, Chaves LJ, Telles MP de C, Diniz-Filho JAF, Resende LV (2008b) Spatial distribution of intrapopulational genetic variability in *Dipteryx alata*. *Pesquisa Agropecuária Brasileira*, **43**, 1151–1158.

Soares T, Melo D, Resende L (2012) Development of microsatellite markers for the neotropical tree species *Dipteryx alata* (Fabaceae). *American Journal of Botany*, **99**, e72–e73.

Souza Neto A (2012) Filogegrafia e fluxo gênico em espécies do gênero *Caryocar*. Universidade Federal de Goiás.

Stanley D, Voeks R, Short L (2012) Is non-timber forest product harvest sustainable in the less developed world? A systematic review of the recent economic and ecological literature. *Ethnobiology and Conservation*, **1**, 1–39.

Stein RJ, Höreth S, de Melo JRF *et al.* (2016) Relationships between soil and leaf mineral composition are element-specific, environment-dependent and geographically structured in the emerging model *Arabidopsis halleri*. *New Phytologist*, 1–13.

Suárez NM, Betancor E, Fregel R, Rodríguez F, Pestano J (2012) Genetic signature of a severe forest fire on the endangered Gran Canaria blue chaffinch (*Fringilla teydea polatzeki*). *Conservation Genetics*, **13**, 499–507.

Tarazi R, Moreno M, Gandara F *et al.* (2010) High levels of genetic differentiation and selfing in the Brazilian cerrado fruit tree *Dipteryx alata* Vog.(Fabaceae). *Genetics and Molecular Biology*, **33**, 78–85.

Thiel-Egenter C, Gugerli F, Alvarez N *et al.* (2009) Effects of species traits on the genetic diversity of high-mountain plants: a multi-species study across the Alps and the Carpathians. *Global Ecology and Biogeography*, **18**, 78–87.

Ticktin T (2004) The ecological implications of harvesting non-timber forest products.

Journal of Applied Ecology, **41**, 11–21.

Del Tredici P (2001) Sprouting in temperate trees: A morphological and ecological review. *The Botanical Review*, **67**, 121–140.

Uchiyama K, Goto S, Tsuda Y, Takahashi Y, Ide Y (2006) Genetic diversity and genetic structure of adult and buried seed populations of *Betula maximowicziana* in mixed and post-fire stands. *Forest Ecology and Management*, **237**, 119–126.

Vellend M, Geber MA (2005) Connections between species diversity and genetic diversity. *Ecology Letters*, **8**, 767–781.

Venables W, Ripley B (2002) *Modern Applied Statistics with S*. Springer.

Vesk PA, Westoby M (2004) Sprouting ability across diverse disturbances and vegetation types worldwide. *Journal of Ecology*, **92**, 310–320.

Vinson CC, Ribeiro DO, Harris SA, Sampaio I, Ciampi AY (2009) Isolation of polymorphic microsatellite markers for the tetraploid *Dipteryx odorata*, an intensely exploited Amazonian tree species. *Molecular Ecology Resources*, **9**, 1542–1544.

Vranckx G, Jacquemyn H, Muys B, Honnay O (2012) Meta-analysis of susceptibility of woody plants to loss of genetic diversity through habitat fragmentation. *Conservation Biology*, **26**, 228–237.

Wang X, Huang Y, Long C (2013) Assessing the genetic consequences of flower-harvesting in *Rhododendron decorum* Franchet (Ericaceae) using microsatellite markers. *Biochemical Systematics and Ecology*, **50**, 296–303.

Ward M, Dick CW, Gribel R, Lowe AJ (2005) To self, or not to self… A review of outcrossing and pollen-mediated gene flow in neotropical trees. *Heredity*, **95**, 246–254.

Weir B (1996a) *Genetic data analysis II: Methods for discrete population genetic data.* Sinauer Associates, Sunderland.

Weir B (1996b) *Genetic Data Analysis*. Sinauer Associates, Sunderland.

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

Wernsdörfer H, Caron H, Gerber S *et al.* (2011) Relationships between demography and gene flow and their importance for the conservation of tree populations in tropical forests under selective felling regimes. *Conservation Genetics*, **12**, 15–29.

Xu W, Bai W, Wen G, Huai H, Liu A (2013) Influences of harvesting on genetic diversity and population structure of *Anemone altaica* (Ranunculaceae), a traditional Chinese medicinal herb. *Biochemical Systematics and Ecology*, **47**, 121–125.

Yan X-B, Guo Y-X, Zhao C, Liu F-Y, Lu B-R (2009) Intra-population genetic diversity of two wheatgrass species along altitude gradients on the Qinghai-Tibetan Plateau: its implication for conservation and utilization. *Conservation Genetics*, **10**, 359–367.

Young A, Boyle T, Brown T (1996) The population genetic consequences of habitat

fragmentation for plants. *Trends in Ecology & Evolution*, **10**, 413–418.

Zardo R (2008) Efeito do impacto da extração de frutos na demografia do pequi (*Caryocar brasiliense*) no Cerrado do Brasil central. Universidade de Brasília.

Zardo RN, Henriques RPB (2011) Growth and fruit production of the tree *Caryocar brasiliense* in the Cerrado of central Brazil. *Agroforestry Systems*, **82**, 15–23.