# Genetic Structure and Diversity of *Attalea vitrivir* Populations in Brazilian Cerrado

Bárbara Rayane Ramos Muniz Nassau[1], Lorena Ramos da Mata[2], Márcio de Carvalho Moretzsohn[2],
Vânia Cristina Rennó Azevedo[2], Marcelo Mattos Cavallari[3], Murilo Malveira Brandão[4], Henrique Maia Valério[4],
Marcílio Fagundes[4], Maurício Lopes de Faria[4], Walter Santos de Araújo[4] & Marcio Antonio Silva Pimenta[4]

[1] Programa de Pós-Graduação em Biodiversidade e Uso dos Recursos Naturais, Universidade Estadual de Montes Claros, Montes Claros, MG, Brazil

[2] Embrapa Recursos Genéticos e Biotecnologia, Parque Estação Biológica, Brasília, DF, Brazil

[3] Embrapa Pecuária Sudeste, São Carlos, SP, Brazil

[4] Department of General Biology, Universidade Estadual de Montes Claros, Montes Claros, MG, Brazil

Correspondence: Marcio Antonio Silva Pimenta, Laboratório de Genética da Conservação, Universidade Estadual de Montes Claros, Av. Dr. Ruy Braga s/n, Campus Universitário Professor Darcy Ribeiro, 39401-089, Prédio 6, Centro de Ciências Biológicas e da Saúde, Montes Claros, MG, Brazil. Tel: 55-(38)-3229-8191. Fax: 55-(38)-3229-8033. E-mail: marcio.pimenta@unimontes.br

## Abstract

The Babassu (*Attalea* sp.) is a native palm tree, belonging to the family Arecaceae, with a wide distribution in Brazil. The species *Attalea vitrivir* Zona occurs in the Northwest of Minas Gerais State and Southwest of Bahia State, where the Cerrado is the predominant biome. Babassu is an important native oil resource and one of the main extractive products in Brazil, besides contributing significantly to the economy of some Brazilian States. The results of an ongoing study of the conservation status and genetic diversity of populations of the palm *A.vitrivir* are presented. The distribution of *A. vitrivir* has been reduced by habitat fragmentation, resulting in two main regions of occurrence (northern and southern), with a large gap between them, being only one of them protected by a conservation unit. A total of 84 microsatellite primer pairs designed for *Attalea speciosa* were screened for polymorphism and transferability in *A. vitrivir* individuals and 21 transferable primer pairs were used to better understand the distribution of genetic variability, gene flow and conservation status of *A. vitrivir* populations.The genetic diversity estimates using these primers indicated high levels of genetic variation and showed that a higher variation was located within than among populations, probably due to reproductive isolation. The five populations evaluated were grouped into two distinct groups, coinciding with northern and southern regions. High genetic diversity within populations and high differentiation among both regions indicate limited gene flow. The northern unprotected populations presented high genetic variability and should also be considered for protection for the long term conservation of the species.

**Keywords:** Babassu, Cerrado fragmentation, population genetics, microsatellites

## 1.Introduction

The Cerrado is the largest, richest and, possibly, the most threatened tropical savanna in the world (Silva & Bates, 2002). Myers et al. (2000) ranked the Cerrado among the 25 most important biodiversity hotspots globally, housing more than 12,000 plant species (Mendonça et al., 2008). However, only about 1.2% of the Cerrado is preserved in protected areas (Mittermeier et al., 1999).

The Cerrado is a floristically diverse savanna that covers more than two million square kilometers of Central Brazil and parts of Bolivia and Paraguay (Simon et al., 2009). Savannas are the dominant vegetation, covering about 70% of the region, interspersed by patches of dry forests and complex belts composed of both dry forests and savannas, giving the region a mosaic-like aspect (Silva & Bates, 2002). Boundaries of Cerrado with other biomes (Amazonia, Atlantic Forest, Chaco and Caatinga) appear to have been porous to the ingress and recruitment of lineages from a range of vegetation types (Simon et al., 2009). For these reasons, Cerrado species are not expected to be uniformly distributed throughout the biome.

Babassu is the common name given to several palm trees from *Attalea* (synonym: *Orbignya*). The genus *Attalea* (Arecaceae) is a poorly studied group of Neotropical palms (Pintaud, 2008). Its main center of diversity is the Brazilian Shield (Anderson & Balick, 1988), an ancient pediplane geologically associated to the Cerrado (Simon & Proença, 2000). Of the approximately 33 species of *Attalea* in Brazil, 17 occur in the Cerrado region (Leitman et al., 2015).

*Attalea vitrivir* Zona (synonym: *Orbignya oleifera* Burret.) is a palm tree restricted to the Eastern portion of the Cerrado, between Northwest of Minas Gerais (MG) State and Southwest of Bahia (BA) State (Lorenzi et al., 2010), where the Cerrado meets two other biomes, the Caatinga and the Atlantic Forest (Lopes et al., 2010). According to Glassman (1999), *A. vitrivir* appears between the valleys of the Urucuia (to the south) and São Francisco (to the east) rivers and the borders of Bahia (to the north) and Goiás (to the west) States. Within this region, *A. vitrivir* is protected only in the *River Pandeiros Environmental Protection Area* (EPA Rio Pandeiros), a conservation unit that is considered of extremely high biological importance for the conservation of Brazilian diversity (MMA, 2007).

Apart from this protected area, the range of *A. vitrivir* is highly fragmented and mostly converted into cattle pastures. Although species of *Attalea* can be favored by clearing (Barot et al., 2005) and can form massive stands in open pastures (Anderson et al., 1991), suppression of the Cerrado and population fragmentation can have negative consequences for its long-term conservation. Fragmentation results in the erosion of genetic variation and increases interpopulation genetic divergence due to increased random genetic drift, elevated inbreeding and reduced gene flow (Young et al., 1996). The knowledge of the genetic variation of *A. vitrivir* populations is therefore essential for their management and to establish conservation strategies.

Here we present the results of a study aimed at a better insight into the distribution of genetic variation of *A. vitrivir* populations in this mosaic landscape assessed by 21 microsatellite or SSR (simple sequence repeat) markers, transferred from *Attalea speciosa* (synonym: *Orbignya phalerata*). SSR markers are useful for a variety of applications in plant genetics because of their reproducibility, multi-allelic nature, codominant inheritance, relative abundance and good genome coverage (Varshney et al., 2005). Because of these advantages, SSR markers have been increasingly used to evaluate the genetic diversity and population structure of several palm trees (Oliveira et al., 2014; Lanes et al., 2015; Ibanes et al., 2015; Ithnin et al., 2017; Soares et al., 2019).

## 2.Material and Methods

### 2.1 Sampling

Two main regions of *A. vitrivir* occurrence have been identified: one in the northern part of the species distribution in southern Bahia State and in the extreme northern region of Minas Gerais State, and the other one further south in central-northern of Minas Gerais State. These two regions are separated by approximately 90 km where no populations were observed, suggesting that northern (N) and southern (S) populations are geographically isolated.

Sampling was undertaken in these two isolated regions (Figure 1): three sampling sites in the northern region, with distances of 17.1 km between N1 and N2 and 16.5 km between N2 and N3; and two sampling sites in the southern region, with 18.6 km distance between S1 and S2. Twenty-five individuals were sampled in population N1, 26 in N2, S1 and S3, and 38 individuals in population N3. A minimum distance of 50 m between sampled individuals was observed, in order to cover the whole area within each population.

The southern sampling sites are located within the River Pandeiros Environmental Protection Area (EPA), in north of MG, Brazil (15°26′10″ S and 44°40′44″ W). The EPA-Pandeiros is located in a transition zone between the Caatinga and Cerrado biomes. The predominant vegetation are Cerrado, gallery forests, and seasonally deciduous forests (Silva et al., 2009).
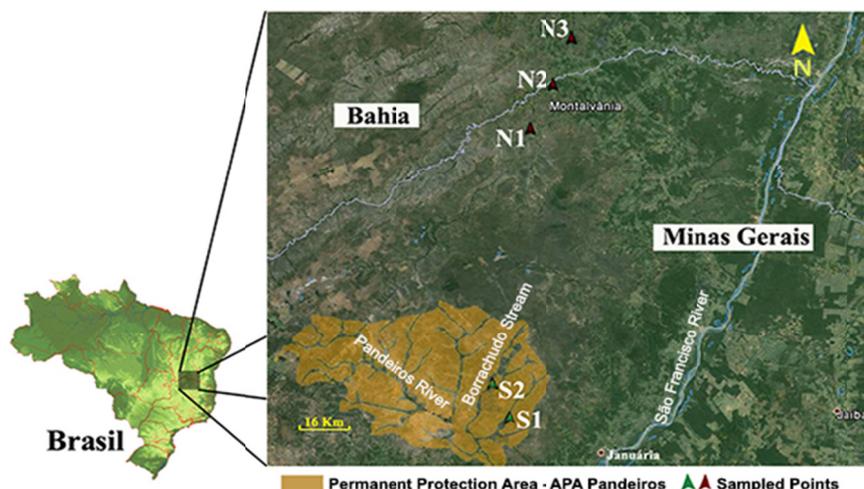
Figure 1. Sampling locations of N1, N2, N3, S1, and S2 populations. The yellow region corresponds to the Rio Pandeiros Environmental Protection Area (EPA-Pandeiros). Source: Google Earth (2019)

### 2.2 DNA Extraction and Amplification

Total genomic DNA was extracted from young leaflets as described by J. J. Doyle and J. L. Doyle (1987). The quality and quantity of the DNA were evaluated by 1% agarose gel electrophoresis.

A total of 84 microsatellite primer pairs developed for *Attalea speciosa* (unpublished data) were screened for polymorphism and transferability against a subsample of 12 *A. vitrivir* individuals, being 2-3 from each of the five populations analyzed, randomly selected. This screening was performed on 2% agarose gels stained with ethidium bromide. Of the 84 primer pairs screened, 21 (Table 1) were selected and used to genotype the 141 individuals sampled. The forward primers were labeled with a fluorescent dye (HEX or 6-FAM, Applied Biosystems, Foster City, CA, USA). The primers were multiplexed according to the fluorescence, annealing temperature and size of the amplified alleles. PCR reactions were performed with 1X PCR buffer (600 mM Tris-$SO_4$, 180 mM $(NH_4)_2SO_4$, 2.0 mM $MgSO_4$, pH 8.9) (1.3 µl), 0.2 mM of each dNTP (1.3 µl), 0.2 mg/ml BSA (1.3 µl), 0.1 µM forward primer (0.15 µl), 0.1 µM reverse primer (0.15 µl), 1 U HF Platinum *Taq* DNA polymerase (0.2 µl), 3.0 ng DNA (3 µl), and 5.6 µl of ultra-pure water for a final volume of 13 µl.

Amplifications were carried out on ABI 9700 thermocyclers (Applied Biosystems), with the following conditions: 95 °C for 15 min (1 cycle); 95 °C for 30 s, 50-60 °C for 1 min (annealing temperature depending on primer pair), 72 °C for 1 min (30 cycles); and a final extension for 30 min at 72 °C (1 cycle).

The PCR products were denatured and size fractioned using capillary electrophoresis in an ABI 3700 automatic DNA analyzer (Applied Biosystems). Loading samples contained 1 µl of the PCR product diluted 1:10 in ultra-pure water, 8.5 µl of Hi-Di formamide (Applied Biosystems) and 0.5 µl of ROX-labeled size-standards. Allele sizing of the electrophoretic data thus obtained was done using Gene Mapper 4.1 (Applied Biosystems).

Table 1. Primer pair sequences, annealing temperatures (Ta), fragment size range, total number of alleles and expected heterozygosity (He) of the 21 SSR markers used for genetic analysis of *A. vitrivir*. F = forward, R = reverse

| Locus | Sequence (5'→3') | Ta (ºC) | Fragmentsize (pb) | No. alleles | He |
|-------|------------------|---------|-------------------|-------------|-----|
| OPH 03 | F: GATGTGTTTGCAACGAAGAGAG<br>R: TCTCCAACATAAACACCTGACG | 54 °C | 100-138 | 7 | 0.535 |
| OPH 10 | F: TGATCTGGGTTACCTAAGTCCC<br>R: TGTCGAATGGTATTCAAGGATG | 58 °C | 330-366 | 9 | 0.590 |
| OPH 11 | F: AGAGCGAGACGAGAAGAAGAGA<br>R: CGTTAGAGGGAAACCACTTTCA | 60 °C | 91-125 | 11 | 0.724 |
| OPH 12 | F: ACAGAGGATCACTATTTGCCGT<br>R: ATAAATGCTTGCGTCCGTTG | 58 °C | 110-139 | 13 | 0.853 |
| OPH 13 | F: ATAAGGCCATCTCCTATTGCAC<br>R: AGAGAAGAAGACGAGCGAAAGA | 53 °C | 105-133 | 14 | 0.867 |
| OPH 25 | F: ATTTTGATTGATGGAGACCATGT<br>R: ATGTCATATTTTCCGTGCCATT | 56 °C | 122-151 | 7 | 0.620 |
| OPH 27 | F: AAGAGAGAAAGAGGGAGAGCGT<br>R: CCGATTAGAAGATTTCTCCCCT | 60 °C | 213-231 | 8 | 0.628 |
| OPH 30 | F: GATGGAAAAGAAAACATCAGGG<br>R: TTATCTCACGCATGCACTCTCT | 56 °C | 178-215 | 15 | 0.799 |
| OPH 33 | F: GTGGTCAGAGGCCTTGGTTG<br>R: AGTGCCTGCCCGAACTTGT | 60 °C | 224-260 | 15 | 0.744 |
| OPH 35 | F: CAGAGGGAAAGAGAGAGGAAGA<br>R: CTCTCAATGGCTTTCCATAACC | 56 °C | 311-331 | 10 | 0.660 |
| OPH 38 | F: GTCAATTTCAAAAGCCAGGTTC<br>R: TGGAATCTTAGTGGACAGCAAA | 56 °C | 122-170 | 7 | 0.823 |
| OPH 41 | F: AGGGCTCTTCCTCTCCCTC<br>R: CTAGGCCCAAACTGAGATAGCA | 60 °C | 168-202 | 8 | 0.703 |
| OPH 43 | F: GAACTTTTCTGTTGCTGTCCTTC<br>R: ATAAGCCATAGCAACCACACCT | 60 °C | 130-178 | 9 | 0.677 |
| OPH 45 | F: GCCCAAAATTGAAGTGAAGAGT<br>R: CCAATGTCATCAATAAGCCTCA | 56 °C | 116-174 | 9 | 0.575 |
| OPH 46 | F: CTTTAGATCTCGAGCACCGACT<br>R: CGAGATAGGCTAGGCAACAAGT | 60 °C | 131-192 | 12 | 0.729 |
| OPH 49 | F: TACACGTTTAATGATGCAAGCC<br>R: CTTTATGCTAGGGTTCCACCTG | 58 °C | 143-175 | 14 | 0.799 |
| OPH 50 | F: AATTAGTCGAAGATGGTCCCAA<br>R: AGCGTACTTGCCGTTGTCTTAT | 60 °C | 91-129 | 14 | 0.691 |
| OPH 52 | F: AATATCTTCGAGCAAGGTGGAT<br>R: GAGAAAGGCTCATAAGCAGGAT | 60 °C | 212-242 | 13 | 0.584 |
| OPH 53 | F: GAAAGGGAGACGTCAGAGACAC<br>R: ACCTTTTCATTCTCCTTCCCTC | 60 °C | 127-135 | 6 | 0.490 |
| OPH 58 | F: CTTCCCCTTCACTTCTCCTTCT<br>R: AGATGGCCTCACTAGACCTCAA | 60 °C | 122-138 | 7 | 0.636 |
| OPH 59 | F: TACATGAGTGTGTGTTGCTTGC<br>R: CTCGGTGGGATCCTAGTTCA | 60 °C | 122-185 | 10 | 0.702 |

## 2.3 Statistical Analyses

Total and mean number of alleles per locus, the expected heterozygosity (He) according to Hardy-Weinberg equilibrium of Nei (1978) for each locus and population, the observed heterozygosity (Ho) and the intrapopulation fixation index (*f*) were estimated using the software GDA 1.0 (Lewis & Zaykin, 2002). An Analysis of Molecular Variance (AMOVA) was performed, to assess the distribution of genetic variation among and within populations using GenAlEx v.6.5 (Peakall & Smouse, 2012). The genetic structure of populations was also analyzed using the Bayesian grouping model and the software Structure 2.3.4 (Pritchard & Wen, 2012). The

number of clusters (K) ranged from one to five. Using the admixture model, four independent runs for each K value were performed with 100,000 Markov chain replications and Monte Carlo simulations (MCMC) and 50,000 burn-ins. The most likely number of clusters (K) was estimated by ΔK, as described by Evanno et al. (2005) and using Structure Harvester (Earl & Von Holdt, 2012). The program HICKORY (Holsinger et al., 2002) was used to estimate the population differentiation ($G_{ST}$-B), by the Bayesian method. Based on this parameter, the gene flow (Nm) was estimated by Nm = 0.25 (1 − $G_{ST}$-B)/$G_{ST}$-B (McDermott & McDonald, 1993). Genetic distances based on allelic frequencies of the five populations were estimated using the modified Rogers' genetic distance (Goodman and Stuber, 1983) and the software BOOD (Coelho, 2002). The resulting diagonal matrix was then submitted to cluster analysis using UPGMA (unweighted pair-group method analysis) and the software NTSYS 2.21 (Rohlf, 2009). The reliability of the generated dendrogram was tested by bootstrap analysis also using the BOOD program with 10,000 iterations. Principal coordinate analysis (PCoA) was performed among the 141 individuals, after conversion of the pairwise genetic distance matrix to a covariance-standardized matrix using the GenAlEx 6.5 software (Peakall & Smouse, 2012). The program Barrier 2.2 was used to analyze the presence of spatial genetic discontinuities among the populations by means of *Delaunay* triangulation verified by the Monmonier algorithm (Manni et al., 2004) with geographic coordinates acquired using the Global Positioning System (GPS) and the genetic distances. The Mantel test was applied using the program PC-Ord 4.14 (McCune & Mefford, 1997) to assess the correlation between the genetic distance matrix and the geographic distance matrix.

## 3. Results

A total of 84 microsatellite primer pairs developed for *Attalea speciosa* (unpublished data) were screened for polymorphism and transferability against a subsample of 12 *A. vitrivir* individuals. By using the same annealing temperaure optimized for *A. speciosa*, 69 primer pairs (82.1%) were fully transferable for *A. vitrivir*. Of these, 21 were selected based on the polymorphism detected and the quality of the amplified products (Table 1). These 21 primer pairs were fluorescently labeled with HEX or 6-FAM and used for the genetic analyses of the 141 individuals included in the present study. The 21 microsatellite loci displayed considerable levels of polymorphism. The number of alleles per locus ranged from 6 (OPH 53) to 15 (OPH 30 and OPH 33), with an average of 10.4, while He values ranged from 0.490 (OPH 53) to 0.867 (OPH 13), with an average of 0.687 (Table 1).

### 3.1 Genetic Diversity

The results of the genetic diversity analyses obtained for the five populations are shown in Table 2. The average number of alleles per locus was homogenous among the three northern populations, but different from S1 and S2, with 8.5 and 5.5 alleles per locus, respectively. The five populations showed high levels of expected heterozygosity (He), varying from 0.633 (population S2) to 0.750 (population S1). The mean value of observed heterozygosity (Ho) was 0.518, and ranged from 0.480 (population S1) to 0.540 (populations N2). The fixation index (*f*) ranged from 0.154 (population S2) to 0.367 (population S1) indicating high and variable levels of inbreeding in these populations (Table 2).

Table 2. Number of individuals analyzed per population (N), mean number of alleles per locus (A), expected (He) and observed heterozygosity (Ho), and fixation index (*f*) obtained for the five populations of *A. vitrivir* estimated using 21 microsatellite loci and the software GDA

| Population | N | A | He | Ho | *f* |
|---|---|---|---|---|---|
| N1 | 25 | 7.3 | 0.735 | 0.505 | 0.318 |
| N2 | 26 | 7.2 | 0.714 | 0.540 | 0.247 |
| N3 | 38 | 7.9 | 0.684 | 0.527 | 0.231 |
| S1 | 26 | 8.5 | 0.750 | 0.480 | 0.367 |
| S2 | 26 | 5.5 | 0.633 | 0.537 | 0.154 |
| Mean | | 7.3 | 0.703 | 0.518 | 0.267 |

### 3.2 Genetic Structure of the Populations

The results of the AMOVA partitioning of genetic variation among *A. vitrivir* populations in Minas Gerais and Bahia States are shown in Table 3. There were significant differences among and within populations (at 1% probability of error), but most of the total genetic variation was found among individuals within populations (93.3%), and only 6.7% was found among populations (Table 3). Using a Bayesian approach with HICKORY, the genetic diversity due to allele frequency differences among sub-populations ($G_{ST}$-B), similar to Nei's $G_{ST}$,

was 0.079. The gene flow (Nm) was estimated at 2.914 for the five populations and at 1.384 among the two regions sampled.

Table 3. Analysis of Molecular Variation (AMOVA) of 141 *A. vitrivir* individuals collected in five populations obtained with 21 microsatellite loci and the software GenAlEx

| Source of Variation | d.f. | SQ | VC | %VT |
| --- | --- | --- | --- | --- |
| Among populations | 4 | 8.167 | 0.029 | 6.7** |
| Within populations | 277 | 112.500 | 0.406 | 93.3** |
| Total | 281 | 120.667 | 43.536 | 100 |

*Note*. Degrees of freedom (d.f.), sums of squares (SQ), variance components (VC), percentage of total variation (% VT), **significant at 1% probability of error.

The Structure analysis revealed that the most probable number of real populations was two (K = 2), showing that the five sampled populations form two distinct groups (Figure 2). The two groups coincided with the northern and southern regions.
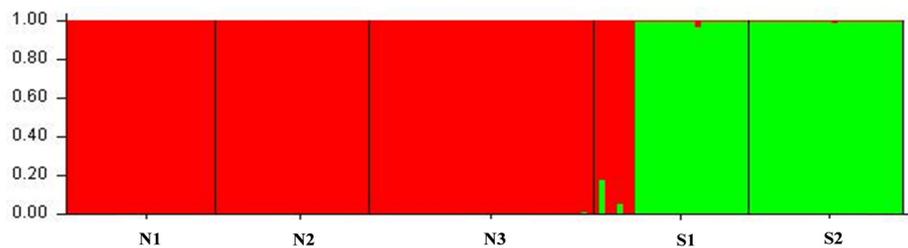


Figure 2. Set of sampled populations grouped by the Bayesian method performed by Structure with K = 2. The colored vertical lines symbolize each individual. Individuals of similar color belong to the same group

The analysis using $F_{ST}$ distances with *Delaunay* triangulation and verified by the Monmonier algorithm using Barrier 2.2 software confirmed the results of the Structure analysis by showing the existence of two principal geographic populations, northern and southern (Figure 3). This analysis revealed similarity between the three northern populations and between the two southern populations. In other words, *Delaunay* triangulation indicated the presence of a barrier that separates the northern populations (N1, N2 and N3) from the southern populations (S1 and S2). The Mantel test indicated that there was no correlation between the genetic and geographical distance matrices (r = 0.212; P = 0.357).
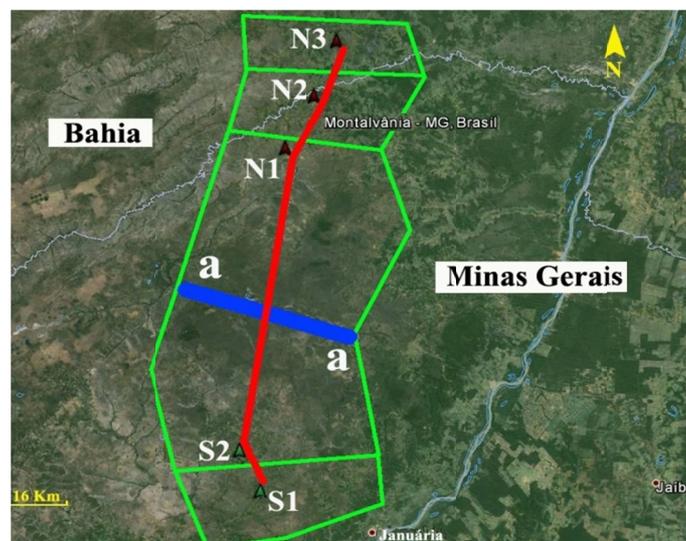


Figure 3. Barrier drawn by BARRIER software separating the two large groups

### 3.3 Genetic Relationships

Genetic distances based on allelic frequencies of the five populations were estimated by the modified Rogers' genetic distance using 21 microsatellite loci. A dendrogram based on UPGMA was constructed for the five populations (Figure 4), and evidenced the separation of the five populations into two main groups, being one composed of the populations located in the northern region and another composed of the southern populations. The cophenetic correlation of the UPGMA clustering of this matrix was very high (0.946, significant at 1% probability by the Mantel test). To assess the variability within populations, a principal coordinate analysis was performed (Figure 5). The separation of the southern and the northern populations were again evident, but this analysis also showed that the southern populations were clearly separated into two subgroups, while the northern populations showed to be more homogenous, with N1, N2, and N3 individuals scattered throughout the northern group.
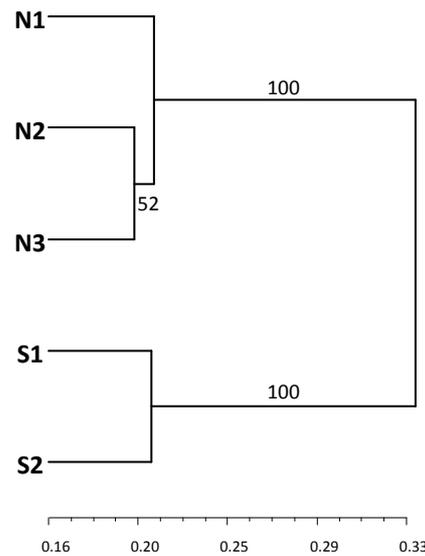


Figure 4. Dendrogram based on genetic distances of the five *A. vitrivir* populations estimated by the modified Rogers' coefficient (Goodman & Stuber, 1983) generated by UPGMA. Bootstrap support values are the result of 10,000 permutations
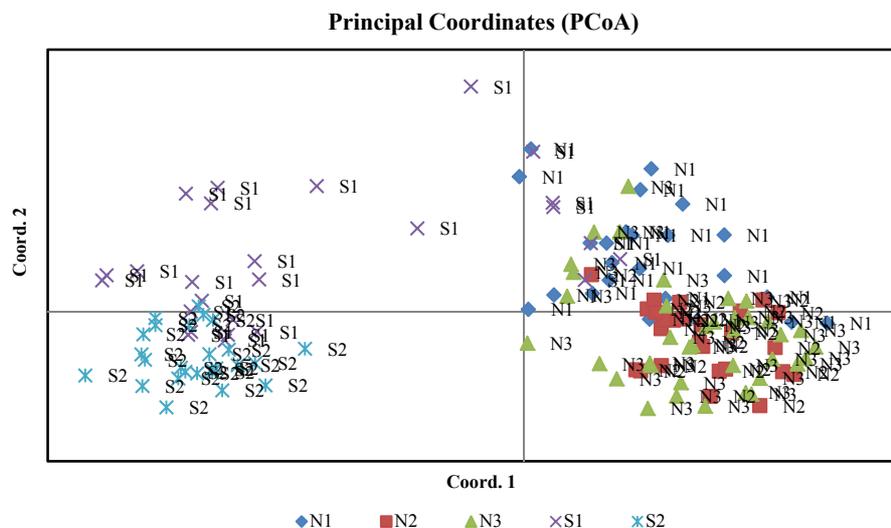


Figure 5. Principal coordinate analysis based on 21 microsatellite loci of the 141 *A. vitrivir* individuals, collected in five populations (S1, S2, N1, N2, and N3)

## 4. Discussion

Recently the genetic structure of basically the same populations analyzed here were investigated by using nine ISSR (Inter Simple Sequence Repeat) primers (Santos et al., 2015). Now we presented the results of an analysis using a considerably higher number of markers in addition of having a codominant inheritance, which allowed a more precise estimates of some genetic parameters.

The possibility of using microsatellite markers developed for one species in genetic studies of related species greatly reduces the cost of the analysis, since the development of microsatellite markers is still expensive and time consuming. On the present study, 82.1% of the primers were fully transferable from *A. speciosa* to *A. vitrivir*. The few studies on palms have also shown, in general, high cross species/genera transferability of microsatellite loci (Zaki et al., 2012; Oliveira et al., 2014; Xiao et al., 2016; Bazzo et al., 2018). Indeed, the efficient transferability of 21 primer pairs to amplify microsatellite loci in *A. vitrivir* corroborates the applicability of heterologous primers to genetic studies in palm trees. These primers detected high levels of genetic variation, with an average He of 0.687 and amplifying an average of 10.4 alleles each (Table 1); thus, confirming their high utility for population genetics studies of *A. vitrivir*.

The values obtained for expected and observed heterozygosity (Table 2) were high and comparable to those reported for other palm species, such as *Syagrus coronata* (Simplicio et al., 2017), *Attalea speciosa* (Ibanes et al., 2015), *Astrocaryum jauari* (Oliveira et al., 2014), *Oenocarpus bataua* (Ottewell et al., 2012), *Bactris gasipaes* (Billotte et al., 2004); and considerably greater than found in *Pseudophoenix* sp. (Rodriguez-Peña et al., 2014). The observed heterozygosity was lower than the expected heterozygosity in all populations, indicating an excess of homozygotes. This was also evidenced by the fixation indices, which were all positive and ranged from 0.154 (population S2) to 0.367 (population S1). Although the excess of homozygotes can be an evidence of inbreeding, it should be considered that our estimates of heterozygosity and fixation indices might have been overestimated by the occurrence of null alleles, as suggested by Chybicki and Burczyk (2009).

The AMOVA, performed with two hierarchical levels, showed that the highest percentage of variation (93.3%) was found within populations, while among population component contained 6.7% of the total variation (Table 3). Comparable results have been found for perennial palms with an outcrossing breeding system, such as *Cocos nucifera* (Lebrun et al., 2005; Rajesh et al., 2014), *Geonomas hottiana* (Silva et al., 2011), *Astrocaryum* (Oliveira et al., 2014), *Elaeis oleifera* (Ithnin et al., 2017). These results are also in accordance with those obtained by Büttow et al. (2010), who observed higher variation (83.68%) within populations of *Butia capitata*. These authors reported that efficient mechanisms for the dispersion of pollen and seeds increase intrapopulation variation, as believed to be the case for *A. vitrivir*. The knowledge of genetic diversity distribution is important to determine strategies for future collection expeditions and for the conservation and use of *A. vitrivir* germplasm. The high genetic divergence found between the populations (northern and southern) may be associated with their reproductive isolation.

$F_{ST}$ values (similar to $G_{ST}$-B) are related to the level of genetic differentiation, with 0.0-0.05 indicating a low differentiation among populations; 0.05-0.25 moderate; and values > 0.25 indicating a high differentiation, according to Yeh (2000). Therefore, the mean $G_{ST}$-B value of 0.079 obtained here indicate that the populations evaluated possess a moderate level of differentiation and, consequently, that diversity exists within populations.

The dendrogram constructed for the five populations (Figure 4) evidenced the separation of the five populations into two main groups, being one composed of the populations located in the northern region and another composed of the southern populations. These results were corroborated by the Structure analysis (Figure 2), which also grouped the individuals into two clusters (K = 2), corresponding to the northern and southern populations. The separation of the southern and the northern populations were also evident by the principal coordinate analysis (Figure 5). This analysis also showed that the southern populations were clearly separated into two subgroups, while the northern populations showed to be more homogenous, with N1, N2, and N3 individuals scattered throughout the northern group. The differentiation of S1 and S2 was not expected, due to their geographical proximity. This can probably be explained by S1 population being composed of well-spaced plants, far apart from each other, which results in a limited gene flow and higher level of inbreeding.

The North and South populations are separated by approx. 90 km, a relatively short distance considering the landscape in which they are inserted. Along the road from South to North populations, no other population was observed. Despite the fragmentation, one would expect to find *A. vitrivir* individuals in this geographic region, and their absence may indicate a limited dispersion of the species. For comparison, *Acrocomia aculeata* is widespread in Brazil and is present even in very fragmented landscapes. The Delaunay triangulation indicated the presence of a barrier that separates the northern populations from the southern populations (Figure 3), despite

the absence of conspicuous geographic barriers in this region. The isolation of north/south population was also evidenced by the gene flow, which was estimated as 2.914 migrant individuals per generation for the five populations, but was as low as 1.384 among the southern and northern populations. These observations indicate a fragility of the species considering the accelerated fragmentation of the landscape: with limited dispersal and gene flow, populations will tend to be more significantly impacted by genetic drift and are more prompted to allele loss.

The genus *Attalea* is represented in Brazil by 32 species (Leitman et al., 2015). While some species such as *Attalea speciosa*, *A. phalerata* and *A. maripa* are widely distributed across different biomes, others occur in small geographic areas, despite de vastness of the apparently homogeneous landscape. That is the case of *A. vitrivir*, wich occurs in a limited range within the Cerrado. Although the landscape in which *A. vitrivir* occurs is highly fragmented and modified by livestock farming, habitat loss may not be the main responsible for its limited distribution. It may be explained by reproductive biology, dispersal syndromes and geographic/edaphoclimatic barriers.

## 5. Conclusions

The transferability of microsatellite markers developed for *Attalea speciosa* to *A. vitrivir* was successful. The information obtained in this study shows that these transferred markers are a robust tool for the genetic analysis of *A. vitrivir* populations. The genetic diversity estimates using these primers indicated high levels of genetic variation and showed that a higher variation was located within than among populations, probably due to reproductive isolation. Some levels of inbreeding probably occur in this species, since the proportion of heterozygotes were smaller than expected for the five populations analyzed. The five populations evaluated were grouped into two distinct groups, coinciding with northern and southern regions. The high genetic diversity within populations not included in the protection area (N1, N2 and N3) and their differentiation from the southern populations evidenced the need for the expansion of conservation areas. Only the southern populations are under protection of a conservation unit, the EPA Rio Pandeiros. Therefore, the northern region (N1, N2 and N3 populations) needs to be included in future species management and conservation plans because it also has a high genetic diversity, is genetically differentiated from the southern populations index and requires more attention.

### Support for Conservation

The high genetic diversity within populations not included in the protection area (N1, N2 and N3), evidences the need for new proposals for the expansion of conservation areas, since their occurrence is restricted and populations S1 and S2 did not exhibit sufficient genetic diversity for the conservation of the species. In addition, we recommend that in order to preserve the genetic variability of the species, germplasm banks should be created and maintained. We also recommend that environmental education programs be established in these areas (northern and southern) and studies be undertaken to describe the reproductive biology of *A. vitrivir* for a better understanding of the dynamics of this species.

## References

Aguiar, R., Quesada, M., Ashowrth, L., Herrerias-Diego, Y., & Lobo, J. (2008). Genetic consequences of habitat fragmentation in plant populations: Susceptible signals in plant traits and methodological approaches. *Molecular Ecolology, 17*(24), 5177-5188. https://doi.org/10.1111/j.1365-294X.2008.03971.x

Anderson, A. B., & Balick, M. J. (1988). Taxonomy of the Babassu complex (*Orbignya* spp.: Palmae). *Systematic Botany, 13*(1), 32-50. https://doi.org/10.2307/2419239

Anderson, A. B., May, P. H., &Balick, M. J. (1991). *The subsidy from nature: Palm forests, peasantry, and development on an Amazon frontier* (p. 233). Columbia University Press. New York, USA.

Barot, S., Mitja, D., Miranda, I., Meija, G. D., & Grimaldi, M. (2005). Reproductive plasticity in an Amazonian palm. *Evolutionary Ecology Research, 7*(7), 1051-1065.

Bazzo, B. R., Carvalho, L. M., Carazzolle, M. F., Pereira, G. A. G., & Colombo, C. A. (2018). Development of novel EST-SSR markers: In the macaúba palm (*Acrocomia aculeata*) using transcriptome sequencing and cross-species transferability in Arecaceae species. *BMC Plant Biology, 18*, 276. https://doi.org/10.1186/s12870-018-1509-9

Billotte, N., Couvreur, T., Marseillac, N., Brottier, P., Perthuis, B., Vallejo, M., & Pintaud, J. C. (2004). A new set of microsatellite markers for the peach palm (*Bactris gasipaes* Kunth); characterization and across-taxa

utility within the tribe Cocoeae. *Molecular Ecology Notes, 4*, 580-582. https://doi.org/10.1111/j.1471-8286.2004.00741.x

Büttow, M. V., Castro, C. M., Schwartz, E., Tonietto, A., & Barbieri, R. L. (2010). Caracterização molecular de populações de *Butia capitata* (Arecaceae) do sul do Brasil através de marcadores AFLP. *Revista Brasileira Fruticultura, 32*(1), 230-239. https://doi.org/10.1590/S0100-29452010005000001

Chybicki, I. J., & Burczyk, J. (2009). Simultaneous estimation of null alleles and inbreeding coefficients. *Journal of Heredity, 100*, 106-113. https://doi.org/10.1093/jhered/esn088

Coelho, A. S. G. (2002). *BOOD: Assessment of errors associated with estimates of distances/genetic similarities across the bootstrap procedure with variable number of markers*. UFG, Goiânia, GO, Brazil.

Doyle, J. J., & Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletim, 19*, 11-15.

Earl, D. A., & Von Holdt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources, 4*(2), 359-361. https://doi.org/10.1007/s12686-011-9548-7

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: A simulation study. *Molecular Ecol*ogy*, 14*, 2611-2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Glassman, S. F. (1999). *A taxonomic treatment of the palm subtribe Attaleinae (tribe Cocoeae)* (Illinois Biological Monographs, Vol. 59). University of Illinois Press, Urbana. https://doi.org/10.5962/bhl.title.49920

Goodman, M. M., & Stuber, C. W. (1983). Races of maize VI. Isozyme variation among races of maize in Bolívia. *Maydica, 28*,169-187.

Google Earth. (2019). *Google Earth Pro.* (Version 7.3). Retrieved from https://www.google.com.br/earth/download/gep/agree.html

Holsinger, K. E., Lewis, P. O., & Dey, D. K. (2002). Bayesian approach to inferring population structure from dominant markers. *Molecular Ecology, 11*, 1157-1164. https://doi.org/10.1046/j.1365-294X.2002.01512.x

Ibanes, B., Sebbenn, A. M., Azevedo, V. C. R., Moreno, M. A., Gandara, F. B., Tambarussi, E. V., … Carvalhaes, M. A. (2015). Genetic diversity and spatial genetic structure in populations of *Orbignya phalerata* Mart. under different exploitation intensities in the Brazilian savanna. *SilvaeGenetica, 64*, 201-211. https://doi.org/10.1515/sg-2015-0019

Ithnin, M., Teh, C. K., & Ratnam, W. (2017). Genetic diversity of *Elaeis oleifera* (HBK) Cortes populations using cross species SSRs: Implication's for germplasm utilization and conservation. *BMC Genetics, 18*(1), 37. https://doi.org/10.1186/s12863-017-0505-7

Lanes, E. C. M., Motoike, S. Y., Kuki, K. N., Nick, C., & Freitas, R. D. (2015). Molecular Characterization and Population Structure of the Macaw Palm, *Acrocomia aculeata* (Arecaceae), Ex Situ Germplasm Collection Using Microsatellites Markers. *Journal of Heredity, 106*(1), 102-112. https://doi.org/10.1093/jhered/esu073

Lebrun, P., Berguer, A., Hodgkin, T., & Baudouin, L. (2005). Biochemical and molecular methods for characterizing coconut diversity. In P. Batugal, V. R. Rao, & J. Oliver (Eds.), *Coconut Genetic Resources* (pp. 225-247). International Plant Genetic Resources Institute, Regional Office for Asia, the Pacific and Oceania (IPGRIAPO), Serdang, Selangor DE, Malaysia.

Leitman, P., Soares, K., Henderson, A., Noblick, L., & Martins, R. C. (2015). *Arecaceae in Lista de Espécies da Flora do Brasil*. Jardim Botânico do Rio de Janeiro. Retrieved from http://floradobrasil.jbrj.gov.br/jabot/floradobrasil/FB15686

Lewis, P., & Zaykin, D. (2002). *Genetic data analysis: Computer programe for the analyses of allelic data* [Version 1.0 (d2)]. Alburquerque: Universityof New Mexico.

Lopes, L. E., Neto, S. D. A., Leite, L. O., Moraes, L. L., & Capurucho, J. M. (2010). Birds from Rio Pandeiros, southeastern Brazil: A wetland in an arid ecotone. *Revista Brasileira de Ornitologia, 18*(4), 267-282.

Lorenzi, H., Noblick, L. R., Kahn, F., & Ferreira, E. (2010). *Flora Brasileira—Arecaceae (Palmeiras)*. Instituto Plantarum, Nova Odessa: São Paulo.

Manni, F., Guérard, E., &, Heyer, E. (2004). Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by using Monmonier's algorithm. *Human Biology, 76*(2), 173-190. https://doi.org/10.1353/hub.2004.0034

McCune, B., & Mefford, M. J. (1997). *Multivariate analysis of ecological data* (Version 3.0). Gleneden Beach: MjM Software, Software.

McDermott, J. M., & McDonald, B. A. (1993). Gene flow in plant pathosystems. *Annual Reviews of Phytopathology, 31*, 353-373. https://doi.org/10.1146/annurev.py.31.090193.002033

Mendonça, R. D., Felfili, J. M., Walter, B. M. T., Silva Júnior, M. D., Rezende, A. V., Filgueiras, T. S., ... Fagg, C. W. (2008). Flora Vascular do Bioma Cerrado—Checklist com 12.356 espécies. In S. M. Sano, S. P. Almeida, & J. F. Ribeiro (Eds.), *Cerrado: Ecologia e flora* (pp. 421-1279). Brasília, Embrapa Cerrados.

Mittermeier, R. A., Myers, N., Mittermeier, C. G., & Robles Gil, P. (1999). *Hotspots: Earth's biologically richest and most endangered terrestrial ecoregions* (p. 431). Mexico City, Mexico: CEMEX, S.A.; Agrupación Sierra Madre, SC.

MMA (Ministério do Meio Ambiente). (2007). *Áreas prioritárias para conservação, uso sustentável e repartição dos benefícios da biodiversidade brasileira* (MAPA Versão 2.2). Brasília, Brazil.

Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature, 403*(6772), 853-858. https://doi.org/10.1038/35002501

Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics, 89*, 586-590.

Oliveira, L. D. S., Ramos, S. L. F., Lopes, M. T. G., Dequigiovanni, G., Veasey, E. A., Macedo, J. L. V., … Lopes R. (2014). Genetic diversity and structure of *Astrocaryum jauari* (Mart.) palm in two Amazon river basins. *Crop Breeding and Applied Biotechnology, 14*, 166-173. https://doi.org/10.1590/1984-70332014v14n3a25

Ottewell, K., Grey, E., Castillo, F., & Karubian, J. (2012). The pollen dispersal kernel and mating system of an insect-pollinated tropical palm, *Oenocarpus bataua*. *Heredity, 109*, 332-339. https://doi.org/10.1038/hdy.2012.40

Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics, 28*, 2537-2539. https://doi.org/10.1093/bioinformatics/bts460

Pintaud, J. C. (2008). An overview of the taxonomy of *Attalea* (Arecaceae). *Revista Peruana de Biologia, 15*(1), 55-63.

Pritchard, J. K., & Wen, W. (2012). *Structure software version 2.3.4*. Retrieved from http://pritch.bsb.uchicago.edu

Rajesh, M. K., Samsudeen, K., Jerard, B. A., Rejusha, P., & Karum, A. (2014). Genetic and phylogenetic relationships of coconut populations from Amini and Kadmat Islands, Lakshadweep (India). *Emirates Journal of Food and Agriculture, 26*, 898-906. https://doi.org/10.9755/ejfa.v26i10.18055

Rodriguez-Peña, R. A., Jestrow, B., Meerow, A. W., Clase, T., Jiménez-Rodríguez, F., Griffith, M. P., … Francisco-Ortega, J. (2014). Genetic diversity and conservation of *Pseudophoenix* (Arecaceae) in Hispaniola. *Botanic Journal of the Linnean Society, 176*(4), 469-485. https://doi.org/10.1111/boj.12223

Rohlf, F. (2009). *NTSYSpc: Numerical taxonomy system* (Ver. 2.21c). Exeter Software: Setauket, New York.

Santos, R. R. M., Cavallari, M. M., Pimenta, M. A. S., Abreu, A. G., Costa, M. R., & Guedes, M. L. (2015). Population genetic structure of *Attalea vitrivir* Zona (Arecaceae) in fragmented areas of southeast Brazil. *Geneticsand Molecular Research, 14*(2), 6472-6481. https://doi.org/10.4238/2015.June.11.23

Silva, J. M. C., & Bates, J. M. (2002). Biogeographic Patterns and Conservation in the South American Cerrado: A Tropical Savanna Hotspot The Cerrado, which includes both forest and savanna habitats, is the second largest South American biome, and among the most threatened on the continent. *BioScience, 52*(3), 225-234. https://doi.org/10.1641/0006-3568(2002)052[0225:BPACIT]2.0.CO;2

Silva, J. O., Jesus, M. F., Fagundes, M., & Fernandes, G. W. (2009). Esclerofilia, taninos e insetos herbívoros associados a *Copaifera lagsdorffii* Desf. (Fabaceae: Caesalpinioideae) em área de transição Cerrado-Caatinga no Brasil. *Ecologia Austral, 19*, 197-206.

Silva, M. S., Vieira, F. A., & Carvalho, D. (2011). Diversity and genetic structure in natural populations of *Geonoma schottiana* Mart (Arecaceae): Implications for Conservation. *Cerne, 17*(2), 195-201. https://doi.org/10.1590/S0104-77602011000200006

Simon, M. F., & Proença, C. (2000). Phytogeographic patterns of Mimosa (Mimosoideae, Leguminosae) in the Cerrado biome of Brazil: An indicator genus of high-altitude centers of endemism? *Biological Conservation, 96*(3), 279-296. https://doi.org/10.1016/S0006-3207(00)00085-9

Simon, M. F., Grether, R., de Queiroz, L. P., Skema, C., Pennington, R. T., & Hughes, C. E. (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*, 106*(48), 20359-20364. https://doi.org/10.1073/pnas.0903410106

Simplicio, R. R., Pereira, D. G., & Waldschmidt, A. M. (2017). Transferability of microsatellite markers in S*yagrus coronate* (Mart.) Becc. (Arecaceae), an iconic palm tree from the Brazilian semiarid region. *Genetics and Molecular Research, 16*(2), 1-6. https://doi.org/10.4238/gmr16029699

Soares, L. A. S. S., Cazetta, E., Santos, L. R., França, D. S., & Gaiotto, F. A. (2019). Anthropogenic Disturbances Eroding the Genetic Diversity of a Threatened Palm Tree: A Multiscale Approach. *Frontiers in Genetics, 10*, 1090. https://doi.org/10.3389/fgene.2019.01090

Varshney, R. K., Graner, A., & Sorrells, M. E. (2005). Genic microsatellite markers in plants: features and applications. *TRENDS in Biotechnology, 23*(1), 48-55. https://doi.org/10.1016/j.tibtech.2004.11.005

Xiao, Y., Xia, W., Ma, J., Mason, A. S., Fan, H., Shi, P., … Peng, M. (2016). Genome-Wide Identification and Transferability of Microsatellite Markers between Palmae Species. *Frontiers in Plant Sciences, 7*, 1578. https://doi.org/10.3389/fpls.2016.01578

Yeh, F. C. (2000). Population genetics. In A. Young, D. Boshier, & T. Boyle (Eds.), *Forest conservation genetics: principles and practice* (pp. 21-37). Collingwood: CSIRO Publishing. https://doi.org/10.1079/9780851995045.0021

Young, A., Boyle, T., & Brown, T. (1996). The population genetic consequences of habitat fragmentation for plants. *Trends in Ecology & Evolution, 11*(10), 413-418. https://doi.org/10.1016/0169-5347(96)10045-8

Zaki, N. M., Singh, R., Rosli, R., & Ismail, I. (2012). *Elaeisoleifera* genomic-SSR markers: Exploitation in oil palm germplasm diversity and cross-amplification in Arecaceae. *International Journal of Molecular Sciences, 13*, 4069-4088. https://doi.org/10.3390/ijms13044069

**Copyrights**