

Development and characterization of microsatellite markers in the African timber tree species *Cylicodiscus gabunensis* (Fabaceae)

Oriana Bili Bhasin (✉ oriana.bhasin@ulb.be)

Université Libre de Bruxelles - Campus du Solbosch: Université Libre de Bruxelles

<https://orcid.org/0000-0001-6663-1487>

Saskia Sargeant

Université Libre de Bruxelles: Université Libre de Bruxelles

Olivier J. Hardy

Université Libre de Bruxelles: Université Libre de Bruxelles

Short Report

Keywords: *Cylicodiscus gabunensis*, Fabaceae, Microsatellites, Sustainable management, Gene flow, Natural regeneration

Posted Date: June 9th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1718484/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Development and characterization of microsatellite markers in the African timber tree species *Cylicodiscus gabunensis* (Fabaceae)

Oriana B. Bhasin ¹, Saskia Sargeant ¹, Olivier J. Hardy ¹

Abstract

I. **Background:** *Cylicodiscus gabunensis* (Fabaceae) or ‘Okan’ is a Central African multipurpose timber species that is used for heavy construction and traditional medicine. Despite being currently heavily exploited, the species shows a low population density and a natural regeneration deficit in dense forest.

II. **Methods and Results:** We aimed to characterize polymorphic microsatellite markers that can be used to study patterns of genetic structure and gene flow (mating system, pollen and seed dispersal and inbreeding depression) and ultimately, help to develop sustainable forest management practices. We developed 24 polymorphic markers that can be amplified in three PCR multiplexes that were tested in 647 individuals of *C. gabunensis* from three populations. The number of alleles ranged from 4 to 27 and the average observed and expected heterozygosity across loci and populations were $H_o = 0.585$ (0.081-0.936) and $H_e = 0.510$ (0.076-0.914), respectively.

III. **Conclusions:** This set of markers is a useful tool for exploring intra-specific diversity, genetic structure and gene flow of *C. gabunensis*.

Key words

Cylicodiscus gabunensis · Fabaceae · Microsatellites · Sustainable management · Gene flow · Natural regeneration

Acknowledgments

We thank E. Kaymak for her help in the laboratory. We also thank J.-F. Gillet, D. Vanderhoeven, A. Boubady, and S.B. Ango for their help with data collection.

Oriana B. Bhasin

oriana.bhasin@ulb.be

<https://orcid.org/0000-0001-6663-1487>

¹ Evolutionary Biology and Ecology, Université libre de Bruxelles, Brussels, Belgium

Introduction

Cylicodiscus gabunensis HARMS (Fabaceae), locally known as ‘Okan’, is a very large tree (up to 60 m height) endemic to the Congo Basin and the only species of the genus *Cylicodiscus*. This non-gregarious, light-demanding species is distributed from Ivory Coast to the Republic of the Congo where it grows in semi-deciduous and dense humid evergreen forests [1]. It bears white-greenish spike-like inflorescences that are hermaphroditic and most-likely pollinated by insects [2]. Fruits are large dehiscent pods containing numerous winged seeds that are wind-dispersed [1, 3]. Like several other light-demanding species, population structure and population density of *C. gabunensis* are variable according to the forest type and are scarce in dense humid evergreen forests [2, 4]. The species is of high socio-economical value. Its dense and hard wood is harvested for heavy constructions including bridges, gates, marine constructions and heavy floor [1]. The tree is also sacred for some communities and its bark and leaves are used for multiple purposes including traditional medicine [2, 3].

C. gabunensis was still unknown in the timber market a few years ago but the interest for the species has increased rapidly in the last two decades and it is now the sixth most exploited species in the Congo Basin [5, 6]. Despite the high socio-economical value of *C. gabunensis*, very little is known about its ecology and genetic structure [2]. Both seed and pollen dispersal affects gene flow and spatial genetic structure (SGS), i.e the non-random spatial distribution of genotypes [7]. By decreasing population density, selective logging is expected to increase SGS because of higher local genetic drift [8]. Intra-specific genetic diversity and gene flow also influence the reproductive success and the capacity of tree populations to adapt to habitat fragmentation. Understanding the processes involved in the reproduction of *C. gabunensis* is therefore crucial to develop sustainable forest management practices, including the definition of a minimal diameter of exploitation and a minimal post-exploitation density allowing to maintain a sufficient potential for regeneration and species adaptation [4, 6, 9, 10].

Nuclear simple sequence repeat markers (SSR), or microsatellites, have gained considerable importance in plant genetics because of their high level of transferability to related species, co-dominant inheritance, important genome coverage, high polymorphism and amenability to high-throughput genotyping [11]. They are used to study the mating system, gene flow, inbreeding depression and the determinants of reproductive success through parentage analyses, for example in African trees [12–14]. They are also used to study the genetic structure of populations, phylogeographic patterns and other studies of genetic variation [8, 14]. The natural regeneration deficit and low observed densities of *C. gabunensis* in dense forest suggest that special precautions should be taken to achieve sustainable exploitation of this species [2, 5]. Therefore, we aimed to develop polymorphic microsatellites primers for *C. gabunensis* using high-throughput sequencing. These nuclear markers will be characterized in three populations to study patterns of genetic structure and gene flow (mating system, pollen and seed dispersal and inbreeding depression).

Material & Methods

Microsatellite development

DNA was isolated from 15–25 mg of silica dried leaf samples of two *Cylicodiscus gabunensis* individuals collected in the Republic of Congo (OH5903, 3.500°S, 12.305°E) and in Cameroon (JFG1811, 3.855°N, 14.768°E), using the Nucleospin 96 Plant II kit (Macherey-Nagel, Duren, Germany). We sheared these DNA extracts using a Bioruptor® Pico (Diagenode SA.) in volume 100 µL of TE 1X pH 8 (Life Technologies AM9858), 6 cycles of 12 s ON and 90 s OFF to obtain fragments of c. 500 bp. We further used the AMPure XP protocol (Agencourt) to remove too small (<400bp) and too large (>800 bp) DNA fragments. We then prepared non-enriched genomic libraries using the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs E7645) and the protocol of Mariac et al. [15]. These libraries were sequenced on an Illumina MiSeq platform at GIGA (Liège, Belgium). The sequencing run generated 1.8 to 2.0 million paired-end 250 bp reads, which were merged with FLASH [16] resulting in 1.5 to 1.6 million sequences with a modal length of 360 to 390 bp (up to 492bp). We used the QDD pipeline [17] to identify microsatellite markers, for which suitable primers were proposed using the parameters described in Malausa et al. [18]. More than 30,000 microsatellite loci were detected. After selecting pure microsatellites with at least 12 repetitions of di-, tri- or tetra-nucleotide repeats, with primers situated at least 16 bp from the microsatellite sequence and a primer design ‘A’ (i.e. best quality according to QDD output), we ordered primers for 45 loci. On each forward primer we added one of four universal sequences (Q1-

Q4, Table 1) associated to fluorescent dyes (6-FAM, NED, VIC and PET, respectively), allowing low-cost M13-like fluorescent labelling [19].

Multiplex reactions and polymorphism tests

Amplification of these 45 primers was tested directly in 24 multiplexed reactions (12 mixes of 4 loci, 8 mixes of 6 loci and 4 mixes of 2 loci) on three individuals (AB501, JFG1654 and NB0518) from three distant populations (Cameroon, Gabon and Ghana) for evaluating successful primers which were polymorphic and showed readable and distinct peak patterns. Amplifications reactions were carried out using Type-it Microsatellite PCR kit (QIAGEN) in 15 μ L volume of 7.5 μ L of Type It, 3 μ L of Q solution, 0.1 μ L of the unlabeled forward primer (0.07 μ M), 0.15 of the Q-tailed reverse primer (0.1 μ M), 0.15 μ L of Q1–Q4 labeled primer (10 μ M each), 1.5 μ L of DNA and H₂O [19]. The following PCR conditions were used: initial premelt phase at 95°C (5 min), 28 cycles of denaturation at 95°C (30 s), annealing at 56°C (90 s), elongation at 72°C (60 s) followed by 12 cycles of denaturation at 94°C (30 s), annealing at 53°C (45 s), elongation at 72°C (45 s) and a final elongation at 60°C (30 min). Genotyping reaction was performed by mixing 12 μ L of highly deionized (Hi-Di) formamide, 0.3 μ L of MapMarker 400 and 1.3 μ L of each PCR product and using an ABI 3730 DNA Analyzer (Applied Biosystems) capillary sequencer. Electropherograms were analyzed on Geneious Prime v. 2022.0.1 [20] to assess peaks readability and level of polymorphism. Out of the initially 45 selected pairs of primers, 37 successfully amplified on the three *C. gabunensis* samples. However, only 24 pairs of primer were finally selected after excluding unreadable or monomorphic loci.

Microsatellite marker data analysis

The selected 24 markers were combined in three multiplexed reactions to carry out population genetic analyses on 649 individuals from two populations from Cameroon (Alpicam with N=242 individuals 3°46'N and 14°43'E, and Dja with N=52 individuals at 3°25'N, 12°78'E) and one from Gabon (CEB with N=355 individuals between 0°30'–1°00'S and 12°30'–14°00'E). The number of alleles (N_a), observed heterozygosity (H_o), expected heterozygosity (H_e), Weir and Cockerham's fixation index (F_{IS}) and deviations from Hardy–Weinberg equilibrium (HWE) were estimated with SPAGeDi 1.5 [21]. CERVUS 3.0.7 was used to measure the polymorphism information content (PIC) [22]. We also jointly estimated the population inbreeding and the frequencies of null alleles for each population using INEst v2.2 [23] with the Bayesian approach.

Results & Discussion

Microsatellite sequences were deposited on GenBank (Table 1). One or two alleles per locus were observed, indicating that *C. gabunensis* is diploid. The 24 markers genotyped in the three populations showed 4 to 27 alleles with an average of 9.96 alleles per locus. PIC values ranged from 0.108 and 0.934 with an average of 0.560 (Table 2). The average expected and observed heterozygosity were respectively 0.585 (0.081–0.936) and 0.510 (0.076–0.914). Of the 24 loci studied, nine loci significantly deviated from HWE in at least one of the populations, five of them usually showing heterozygosity deficiency while four of them showed slight heterozygote excess in population Alpicam (Table 2). Heterozygosity deficiency was associated with higher frequencies of null alleles suggesting that it is probably not a consequence of inbreeding, as supported by the absence of inbreeding in each population according to INEst. However, the frequencies of null alleles were always below 0.20 except for three markers in at least one population (Cyl07, Cyl14 and Cyl36; Table 2).

In conclusion, we developed 24 microsatellite markers that can be used to study intra-specific diversity, genetic structure and gene flow of *C. gabunensis*. They are also expected to be polymorphic enough to conduct paternity or parentage analyses. The biological knowledge that can be acquired with these markers will help determine which practices allow a sustainable exploitation of this timber species. The number of microsatellite loci developed is expected to be sufficient to obtain accurate measures of genetic variation and structure.

Statements & Declarations

Funding

We thank the “Fonds pour la Formation à la Recherche dans l'Industrie et l'Agriculture” (FRIA, Belgium) for providing a PhD grant to OB (FRIA/FC 27793). Molecular analyses were funded by the F.R.S-FNRS WISD grant X.3040.17 (project AFRITIMB).

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Author Contributions

OB and OJH designed the study. OB contributed to collection of plant material. OB and SS performed the laboratory analyses. OB conducted the data analyses. OB and OJH wrote the first draft and all authors contributed to the final version of the manuscript.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Louppe D, Oteng-Amoako AA, Brink M (2008) Plant Resources of Tropical Africa 7(1). Timbers 1. Montpellier
2. Makemba RN, Tosso F, Moupela C, et al (2019) *Cylicodiscus gabunensis* Harms : une espèce prisée dans le commerce international (synthèse bibliographique). *Biotechnol Agron Soc Environ* 23:. <https://doi.org/10.25518/1780-4507.18006>
3. Hédin L (1929) Note sur le bois d'Adoum (*Cylicodiscus gabunensis* Harms). *Rev Bot appliquée d'agriculture Colon* 446–448
4. Karsenty A, Gourlet-Fleury S (2006) Assessing sustainability of logging practices in the Congo Basin's managed forests: The issue of commercial species recovery. *Ecol Soc* 11:. <https://doi.org/10.5751/ES-01668-110126>
5. Seka MJ, Bégin J, Avana TML, et al (2018) Régénération naturelle de l'Okan (*Cylicodiscus gabunensis* Harms) et facteurs influençant la probabilité d'observation des semis. 11:26–38
6. Karsenty A, Ferron C (2017) Recent evolutions of forest concessions status and dynamics in Central Africa. *Int For Rev* 19:10–26. <https://doi.org/10.1505/146554817822295957>
7. Vekemans X, Hardy OJ (2004) New insights from fine-scale spatial genetic structure analyses in plant populations. *Mol Ecol* 13:921–935. <https://doi.org/10.1046/j.1365-294X.2004.02076.x>
8. Bizoux JP, Daïnou K, Bourland N, et al (2009) Spatial genetic structure in *Milicia excelsa* (Moraceae) indicates extensive gene dispersal in a low-density wind-pollinated tropical tree. *Mol Ecol* 18:4398–4408. <https://doi.org/10.1111/j.1365-294X.2009.04365.x>
9. Schwartz G, Falkowski V, Peña-Claros M (2017) Natural regeneration of tree species in the Eastern Amazon: Short-term responses after reduced-impact logging. *For Ecol Manage* 385:97–103. <https://doi.org/10.1016/j.foreco.2016.11.036>
10. Aleksić JM, Mataruga M, Daničić V, et al (2022) High pollen immigration but no gene flow via-seed into a Genetic Conservation Unit of the endangered *Picea omorika* after disturbance. *For Ecol Manage* 510:. <https://doi.org/10.1016/j.foreco.2022.120115>
11. Kalia RK, Rai MK, Kalia S, et al (2011) Microsatellite markers: An overview of the recent progress in plants. *Euphytica* 177:309–334. <https://doi.org/10.1007/s10681-010-0286-9>
12. Monthe FK, Hardy OJ, Doucet JL, et al (2017) Extensive seed and pollen dispersal and assortative mating in the rain forest tree *Entandrophragma cylindricum* (Meliaceae) inferred from indirect and direct analyses. *Mol Ecol* 26:5279–5291. <https://doi.org/10.1111/mec.14241>
13. Hardy OJ, Delaide B, Hainaut H, et al (2019) Seed and pollen dispersal distances in two African legume timber trees and their reproductive potential under selective logging. *Mol Ecol* 28:3119–3134
14. Angbonda DMA, Monthe FK, Bourland N, et al (2021) Seed and pollen dispersal and fine-scale spatial genetic structure of a threatened tree species: *Pericopsis elata* (HARMS) Meeuwen (Fabaceae). *Tree Genet Genomes* 17:. <https://doi.org/10.1007/s11295-021-01509-8>
15. Mariac C, Scarcelli N, Pouzadou J, et al (2014) Cost-effective enrichment hybridization capture of chloroplast genomes at deep multiplexing levels for population genetics and phylogeography studies. *Mol Ecol Resour* 14:1103–1113. <https://doi.org/10.1111/1755-0998.12258>
16. Magoč T, Salzberg SL (2011) FLASH: Fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 27:2957–2963. <https://doi.org/10.1093/bioinformatics/btr507>
17. Megléc E, Costedoat C, Dubut V, et al (2009) QDD: A user-friendly program to select microsatellite markers and design primers from large sequencing projects. *Bioinformatics* 26:403–404. <https://doi.org/10.1093/bioinformatics/btp670>
18. Malausa T, Gilles A, Megléc E, et al (2011) High-throughput microsatellite isolation through 454 GS-FLX Titanium pyrosequencing of enriched DNA libraries. *Mol Ecol Resour* 11:638–644.

- <https://doi.org/10.1111/j.1755-0998.2011.02992.x>
19. Micheneau C, Dauby G, Bourland N, et al (2011) Development and characterization of microsatellite loci in *Pericopsis elata* (Fabaceae) using a cost-efficient approach. *Am J Bot* 98:268–270. <https://doi.org/10.3732/ajb.1100070>
 20. Kearse M, Moir R, Wilson A, et al (2012) Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28:1647–1649. <https://doi.org/10.1093/bioinformatics/bts199>
 21. Hardy OJ, Vekemans X (2002) spagedi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol Ecol Notes* 618–620. <https://doi.org/10.1046/j.1471-8278>
 22. Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol Ecol* 16:1099–1106. <https://doi.org/10.1111/j.1365-294X.2007.03089.x>
 23. Chybicki IJ, Burczyk J (2009) Simultaneous estimation of null alleles and inbreeding coefficients. *J Hered* 100:106–113. <https://doi.org/10.1093/jhered/esn088>
 24. Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. *Nat Biotechnol* 18:233–234
 25. Culley TM, Weller SG, Sakai AK, Putnam KA (2008) Characterization of microsatellite loci in the Hawaiian endemic shrub *Schiedea adamantis* (Caryophyllaceae) and amplification in related species and genera. *Mol Ecol Resour* 8:1081–1084

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Table1SSROBMolBiolRep.pdf](#)
- [Table2SSROBMolBiolRep.pdf](#)