

Development of nuclear SNP markers for the timber tracking of the African tree species Sapelli, *Entandrophragma cylindricum*

Céline Blanc-Jolivet¹ · Birgit Kersten¹ · Nils Bourland^{2,3,4} · Erwan Guichoux⁵ · Adline Delcamp² · Jean-Louis Doucet⁶ · Bernd Degen¹

Received: 31 August 2017 / Accepted: 24 September 2017
© Springer Science+Business Media B.V. 2017

Abstract We describe the development of new nuclear SNP markers for the genetic timber tracking of the geographical origin of Sapelli, *Entandrophragma cylindricum* (Meliaceae). Restriction associated DNA sequencing (RADseq) of two reference individuals yielded 1131 putative SNPs. Among those, 131 were selected to design four MassARRAY multiplexes and screened at 178 individuals. Seventy-two loci were selected for further use in genetic tracking.

Keywords Single nucleotide polymorphism · *Entandrophragma* · RAD sequencing · MassARRAY · Timber tracking

Electronic supplementary material The online version of this article (doi:10.1007/s12686-017-0872-4) contains supplementary material, which is available to authorized users.

✉ Céline Blanc-Jolivet
celine.blanc-jolivet@thuenen.de

¹ Thünen Institute of Forest Genetics, Sieker Landstrasse 2, 22927 Grosshansdorf, Germany

² Service of Wood Biology, Royal Museum for Central Africa, 3080 Tervuren, Belgium

³ Center for International Forestry Research, Bogor 16115, Indonesia

⁴ Resources & Synergies Development Pte Ltd, Hong Leong Building, Singapore 048581, Singapore

⁵ BIOGECO, INRA, University of Bordeaux, 33610 Cestas, France

⁶ Nature +, Winstar Park, 62 rue Provinciale, 1301 Wavre, Belgium

Introduction

Sapelli, *Entandrophragma cylindricum* (Meliaceae), is a heavily exploited hardwood timber species occurring in tropical forests from Africa (Lourmas et al. 2007). This species is classified as Vulnerable in the IUCN Red List of Threatened Species and occurs at low densities (Lourmas et al. 2007). In order to trace the origin of *E. cylindricum* timber and thereby assisting the enforcement of laws on timber trade, we developed molecular markers for genetic timber tracking. Until now, only nuclear microsatellites (nSSRs) have been developed to study the mating system and gene flow (Garcia et al. 2004; Lourmas et al. 2007) and have been used to control the forest concession of origin (Jolivet and Degen 2012). New sets of nSNPs have been recently developed for other valuable tropical timber species for traceability purposes (Blanc-Jolivet et al. 2017; Jardine et al. 2016; Pakull et al. 2016) and present several advantages over nSSRs (Blanc-Jolivet and Liesebach 2015). Here we present a new set of nSNPs for *E. cylindricum* optimized for low-cost tracking of geographical origin.

Marker development and screening

One sample from Cameroon and one sample from DRC were selected for SNP marker development with restriction associated DNA sequencing (RADseq; Table 1). DNA was extracted from dried cambium of leaves following (Dumolin et al. 1995) and sent to Floragenex (Portland, Oregon, USA) for sequencing and SNP detection. RADseq procedure and SNP detection protocols were the same as described in (Blanc-Jolivet et al. 2017; Jardine et al. 2016). We screened a total of 178 samples, recorded as *E. cylindricum* by our field teams, and distributed over the species distribution range, at the newly developed SNP loci (Table 1).

Table 1 *Entandrophragma cylindricum* samples used for the screening of 131 SNP loci

Country	Region	Latitude	Longitude	Number of samples	Timber
Ghana	Edubiase	6.04	− 1.39	6	1
Ghana	Nkrabia	6.01	− 1.55	9	1
Ghana	Bosombepo	6.70	− 2.63	9	
Cameroon	UFA 004	4.62	10.42	10	
Cameroon	Kamba	3.22	11.35	5	1
Cameroon	Egona	4.84	11.31	5	
Cameroon	Djoum	2.26	13.05	10	
Cameroon	Pallisco	3.22	14.26	3 ^a	
Cameroon	CFC	3.30	15.30	5	
Cameroon	FBR Ngoli	2.09	15.36	8	
Congo Brazaville	UFA Pokola	1.33	16.39	13	
Congo Brazaville	UFA Loundougou	2.38	17.06	13	
DRC	Mabali	0.89	18.13	14	
DRC	Gemena	3.17	19.81	9	
DRC	Lisala	2.36	21.30	10	
DRC	Lomami	0.34	23.94	5	
DRC	Yangambi	0.78	24.49	10	
DRC	Kisangani	0.62	25.26	15	
DRC	Yoko 400 ha	0.29	25.30	1 ^a	
DRC	Kisangani-East	0.49	25.93	15	

^aSamples used for RAD sequencing

We selected contigs with only one putative SNP identified from RADseq data to minimize the risk of null alleles. A further criterion was the quality of the sequence data. Primer design and multiplexing for a MassARRAY iPLEX platform (Agena Bioscience) was conducted in Assay Design Suite (ADS) (Agena Bioscience, San Diego, USA). The Genome Transcriptome Facility of Bordeaux (France) conducted all steps from DNA amplification to fragment separation on the MassARRAY system. Interpretation of the spectra was conducted in Typer Viewer v 4.0.24.71 (Agena Bioscience).

Data analysis and marker selection

Data analysis was conducted using genotypes grouped by country or by region (Table 1) to detect differentiation at several geographical scales. The most informative loci were selected according to their high amplification success, high correlation among geographic and genetic distances, and high average genetic differentiation δ among groups (Gregorius 1987) and designed in four MassARRAY multiplexes. The performance of the selected set of loci was tested with a self-assignment test using both data grouping, and a comparison with the results provided by the complete set of loci. The likeliest population was estimated with the criteria of Rannala and Mountain (1997) in GDA-NT (Degen, unpublished) and with a nearest-neighbor approach using only samples with more than 50% available data and a

percentile of 0.02 for the number of nearest-neighbor (Degen et al. 2017).

Results

RADseq data of the two *E. cylindricum* individuals from Cameroon and DRC yielded a total of 1131 putative SNPs under the “stringent” SNP discovery criteria, which provides loci with high base quality and sequencing depth (Floragenex). 131 loci (130 nuclear and one chloroplast, P0002101) were designed in four MassARRAY multiplexes for the screening of 178 individuals (Table S1). Considering all other loci except P0000831 (no allele calling possible), 14–100% amplified with an average of 83% per sample. Indeed, amplification rate per locus ranged from 43 to 100% with an average of 83%, suggesting the presence of null alleles in the samples from DRC. Amplification success was very high in the three timber samples and was over 98%. Estimation of the correlations among genetic and geographic distances yielded values up to 0.97 when samples were grouped per country (Table S2), indicating a strong differentiation among West African (Ghana) and Central African countries. Based on high differentiation and high correlation among genetic and geographic distances, a final set of 72 nuclear loci was designed for MassARRAY genotyping (Table S2 and Table S3). Self-assignment success with

the Bayesian approach from Rannala and Mountain (1997) slightly decreased when the final set of loci was used and data grouped per country (from 70 to 69%), but increased from 35 to 37% when data was grouped per region. Using the nearest-neighbour approach significantly improved the self-assignment success per region (50 and 46% with all and the selected set of loci respectively), while the results per country were not modified (69%).

The set of nuclear SNP loci described in this study will be useful for the genetic timber tracking at a large geographical range. However, low amplification success in many samples from DRC raises questions of species identification in the field and cross-amplification of the markers on other *Entandrophragma* species. Genotyping of a larger set of samples over the distribution range of Sapelli together with interspecies tests of the markers will allow a better understanding of the present results on amplification success.

Acknowledgements We would like to thank Dr. E. Foli (Ghana), Prof. B. Sonké (Cameroon), Prof. J. Loumeto (Congo Brazzaville), C. Botefa (DRC), A. Monkengo no Mpenge (DRC), Prof. C. Amani (DRC) and G. Assimbo (DRC) for their support during the sampling. The work was financially supported by the International Tropical Timber Organization (project PD 620/11 Rev.1). Genotyping was performed at the Genomic and Sequencing Facility of Bordeaux (grants from the Conseil Regional d'Aquitaine n°20030304002FA and 20040305003FA, the European Union, FEDER n°2003227 and Investissements d'avenir, N°ANR-10-EQPX-16-01).

References

- Blanc-Jolivet C, Liesebach M (2015) Tracing the origin and species identity of *Quercus robur* and *Quercus petraea* in Europe: a review. *Silvae Genetica* 64(4):182–193
- Blanc-Jolivet C, Kersten B, Daïnou K, Hardy O, Guichoux E, Delcamp A, Degen B (2017) Development of nuclear SNP markers for genetic tracking of Iroko, *Milicia excelsa* and *Milicia regia*. *Conserv Genetic Resour*. doi: 10.1007/s12686-017-0716-2
- Degen B, Blanc-Jolivet C, Stierand K, Gillet E (2017) A nearest neighbour approach by genetic distance to the assignment of individual trees to geographic origin. *Forensic Sci Int Genet* 27:132–141
- Dumolin S, Demesure B, Petit RJ (1995) Inheritance of chloroplast and mitochondrial genomes in pedunculate oak investigated with an efficient PCR method. *Theor Appl Genet* 91(8):1253–1256
- Garcia F, Noyer JL, Risterucci AM, Chevallier MH (2004) Genotyping of mature trees of *Entandrophragma cylindricum* with microsatellites. *J Hered* 95(5):454–457
- Gregorius H-R (1987) The relationship between the concepts of genetic diversity and differentiation. *Theor Appl Genet* 74(3):397–401
- Jardine DI, Blanc-Jolivet C, Dixon RRM, Dormontt EE, Dunker B, Gerlach J et al (2016) Development of SNP markers for Ayous (*Triplochiton scleroxylon* K. Schum) an economically important tree species from tropical West and Central Africa. *Conserv Genetic Resour* 8(2):129–139
- Jolivet C, Degen B (2012) Use of DNA fingerprints to control the origin of sapelli timber (*Entandrophragma cylindricum*) at the forest concession level in Cameroon. *Forensic Sci Int Genet* 6(4):487–493
- Lourmas M, Kjellberg F, Dessard H, Joly HI, Chevallier MH (2007) Reduced density due to logging and its consequences on mating system and pollen flow in the African mahogany *Entandrophragma cylindricum*. *Heredity* 99(2):151–160
- Pakull B, Mader M, Kersten B, Ekue MRM, Dipelet UGB, Paulini M et al (2016) Development of nuclear, chloroplast and mitochondrial SNP markers for Khaya sp. *Conserv Genetic Resour* 8(3):283–297
- Rannala B, Mountain JL (1997) Detecting immigration by using multilocus genotypes. *Proc Natl Acad Sci USA* 94(17):9197–9201