Tree Genetics & Genomes (2017) 13:90
DOI 10.1007/s11295-017-1174-4

ORIGINAL ARTICLE

DNA taxonomy in the timber genus *Milicia*: evidence of unidirectional introgression in the West African contact zone

Kasso Daïnou[1,2,3] · **Jean-François Flot**[3] · **Bernd Degen**[4] · **Céline Blanc-Jolivet**[4] · **Jean-Louis Doucet**[5] · **Ludivine Lassois**[5] · **Olivier J. Hardy**[3]

Received: 30 September 2016 / Revised: 7 July 2017 / Accepted: 13 July 2017

Abstract DNA-based techniques are helpful in characterising hybridisation patterns in plant species. To be efficient in disentangling species boundaries and interspecific gene flow, it is recommended to combine various methodologies and types of markers. Here, we used different analytical tools (algorithms implemented in Structure, Tess, NewHybrids and HIest, and the haploweb approach) and three nuclear genetic markers (7 nuclear simple sequence repeat loci (SSRs), 62 single-nucleotide polymorphism loci (SNPs) and a single-copy gene region, *At103*) to revisit hybridisation patterns in the commercially important African tree genus *Milicia*. Samples were collected in the natural ranges of *Milicia regia* and *Milicia excelsa* in West Africa. Using real data sets, simulated purebreds and hybrid genotypes, we found that SNPs yielded results more consistent than SSRs; outputs from the Bayesian and maximum-likelihood analyses differed significantly using the SSRs, whereas they were perfectly congruent using SNPs. A proportion of 12.4% hybrids were detected amongst the SNP genotype samples. A haploweb analysis of *At103* gene sequences confirmed the existence of interspecific hybrids. There was also a clear evidence of advanced generations of hybrids (backcrossed individuals) but only towards *M. regia*. Although more investigation is required for understanding the mechanisms responsible for this asymmetric introgression, we suggest that it may be due to the differences in flowering time between species and between sexes, combined with a maternal inheritance of flowering time.

Keywords Hybridisation · Introgression · Haploweb · Tropical forests · *Milicia* sp.

Communicated by F. Gugerli

Electronic supplementary material The online version of this article (doi:10.1007/s11295-017-1174-4) contains supplementary material, which is available to authorized users.

✉ Kasso Daïnou
k.dainou@natureplus.be; kdainou@ulg.ac.be

[1] Nature+ asbl / TERRA Research Centre, Central African Forests, Gembloux Agro-Bio Tech, University of Liege, Passage des Déportés 2, 5030 Gembloux, Belgium

[2] Université Nationale d'Agriculture, BP 43 Kétou, Benin

[3] Evolutionary Biology and Ecology – CP 160/12, Faculté des Sciences, Université Libre de Bruxelles, Av. F. Roosevelt 50, 1050 Brussels, Belgium

[4] Thünen Institute of Forest Genetics, Sieker Landstrasse 2, 22927 Grosshansdorf, Germany

[5] TERRA Research Centre, Central African Forests, Gembloux Agro-Bio Tech, University of Liege, Passage des Déportés 2, 5030 Gembloux, Belgium

Introduction

Interspecific hybridisation has been known for long as a common phenomenon in sexually reproducing animals and plants, occurring between species that diverged relatively recently and/or have historically maintained substantial gene flow (Mallet 2005; Freeland et al. 2011). Although interspecific hybridisation does not necessarily lead to gene flow between species, when it does, it can have important consequences for the evolution of species and more generally for speciation and diversification processes. Whereas morphological traits were previously used to identify interspecific hybrids, DNA-based approaches have highly improved our ability to detect such individuals. Depending on the degree of species divergence, markers and analysis methods differ in their ability to reveal hybrids. In particular, though plastid and mitochondrial DNA (uniparentally inherited in most cases) can be useful in

Published online: 31 July 2017

phylogeny and species delineation purposes, they are intrinsically unable to identify hybrids without additional information from the nuclear genome (e.g. Fehrer et al. 2007; Quintela et al. 2010).

Nuclear microsatellites (simple sequence repeats (SSRs)) in particular have been commonly used to characterise hybridisation (e.g. Craft et al. 2002; Barilani et al. 2007; Edwards et al. 2008; Lepais et al. 2009; Pollegioni et al. 2012). However, due to allelic homoplasy, microsatellites can sometimes lack power to disentangle species (Sites and Marshall 2004). In addition, microsatellites might have limited resolution in detecting hybrids beyond the first hybrid generation (von Holdt et al. 2013). Although single-nucleotide polymorphism (SNP) loci also have some limitations (low genetic diversity and high sensitivity to ascertainment bias; DeFaveri et al. 2013; Granevitze et al. 2014), it has been argued that they are valuable in investigating hybridisation (Mallet 2005; Twyford and Ennos 2012; Lamer et al. 2015). SSR and SNP data are often analysed using the Bayesian approaches to estimate either the ancestry coefficient of individuals assigned to genetic clusters (assuming the Hardy-Weinberg equilibrium within the clusters; e.g. Pritchard et al. 2000; Chen et al. 2007) or the probability that a given individual belongs to a hybrid class (Anderson and Thompson 2002). Recently, Fitzpatrick (2012) suggested a maximum-likelihood method that provides a joint estimation of both ancestry index and interclass heterozygosity (the proportion of loci that combines alleles from the two parental species), and plotting these two parameters allows to visualise the type of hybrids in a ternary diagram. The most important advantage of the method is that it does not assume that only the six classical categories of individuals (each parent, F1, F2 and backcross towards each parent) exist in the sample; advanced generations of hybrids (F3, F4…) and various crosses between the parents and types of hybrid may occur, providing a continuum of hybrid index (Gompert and Buerkle 2016). Another interesting analytical tool for hybridisation issues was developed by Flot et al. (2010). Following Doyle (1995), Flot et al. (2010) proposed an allele sharing-based approach to analyse diploid nuclear sequences. Species units and putative hybrids are identified on the basis of heterozygous individuals: haplotypes found co-occurring in heterozygous individuals form an allele pool, and the corresponding group of individuals is called the 'field for recombination' (FFR). Although this 'haploweb' approach primarily aims to detect reproductively isolated species, it can also be used to detect infrequent interspecific hybridisation (Papakostas et al. 2016). One drawback of haplowebs, however, is that they require a dense sampling of individuals.

Interspecific hybridisation has been relatively little studied in tropical trees, especially in Africa. The present work aimed at revisiting hybridisation and species delimitation issues in an African plant genus through a variety of markers and methods. *Milicia* is a tree genus highly exploited for timber production in Africa for several decades (Nyong'o et al. 1994; Ofori and Cobbinah 2007). The genus was long considered to comprise two species: *Milicia excelsa* is widespread in Africa, whereas *Milicia regia* is restricted to West Africa westwards of the Dahomey Gap; however, a third cryptic species located in central Africa was recently identified (Daïnou et al. 2016). Whilst nuclear data (SSRs, SNPs and gene sequences) have confirmed the genetic divergence between *M. excelsa* and *M. regia*, plastid sequences suggest that populations of *M. excelsa* in West Africa are genetically more related to those of *M. regia* in the same region than to their conspecific in central Africa (Daïnou et al. 2014). Indeed, plastid haplotypes from West African populations of *M. excelsa* and *M. regia* belong to the same clade, whilst haplotypes from central African samples of *M. excelsa* form another clade. Nevertheless, mutual haplotypic exclusivity and the apparent absence of admixture amongst genetic clusters suggested by nuclear SSR markers support the hypothesis that each of these two morpho-groups corresponds to a distinct species (Daïnou et al. 2014). However, the two species are hardly distinguishable morphologically except for subtle differences mostly in leaf traits (Tondeur 1939; Nichols et al. 1998; Joker 2002), which has caused confusions when delimiting the distribution range of the two species (e.g. Chevalier 1912 vs. Chevalier 1917).

M. regia and *M. excelsa* display parapatric geographic distributions. Although Ofori et al. (2001) suspected the existence of interspecific hybrids on the basis of genetic distances between populations, this was never confirmed and the extent of the putative hybridisation and introgression has never been examined. From a dense sampling in the Ivory Coast and Ghana, the area where both species co-occur, using SSRs, SNPs and gene sequences, we asked the following questions: (1) are the different types of markers and analysis methods congruent in indicating the existence of interspecific hybridisation in *Milicia* species in West Africa? (2) If hybridisation is confirmed, how frequent is it and where does it occur? (3) If *Milicia* species hybridise, does this process lead to interspecific gene flow?

Material and methods

Study taxa, sample collection and DNA-based data sets

The two known species of *Milicia*, *M. excelsa* and *M. regia*, are amongst the most important timber trees in West and central Africa. As *M. regia* is restricted to West African evergreen forests, its ranges overlap only in this region. Both species are suspected to be wind-pollinated (Joker 2002), and their seeds are mainly dispersed by bats and parrots (Daïnou et al. 2012). The two species are morphologically so close (including wood

properties) that they are exploited under the same commercial name, iroko or obeche. The West African taxon *M. regia* is listed in the IUCN red list as a vulnerable species, and various African countries have banned the exploitation of iroko.

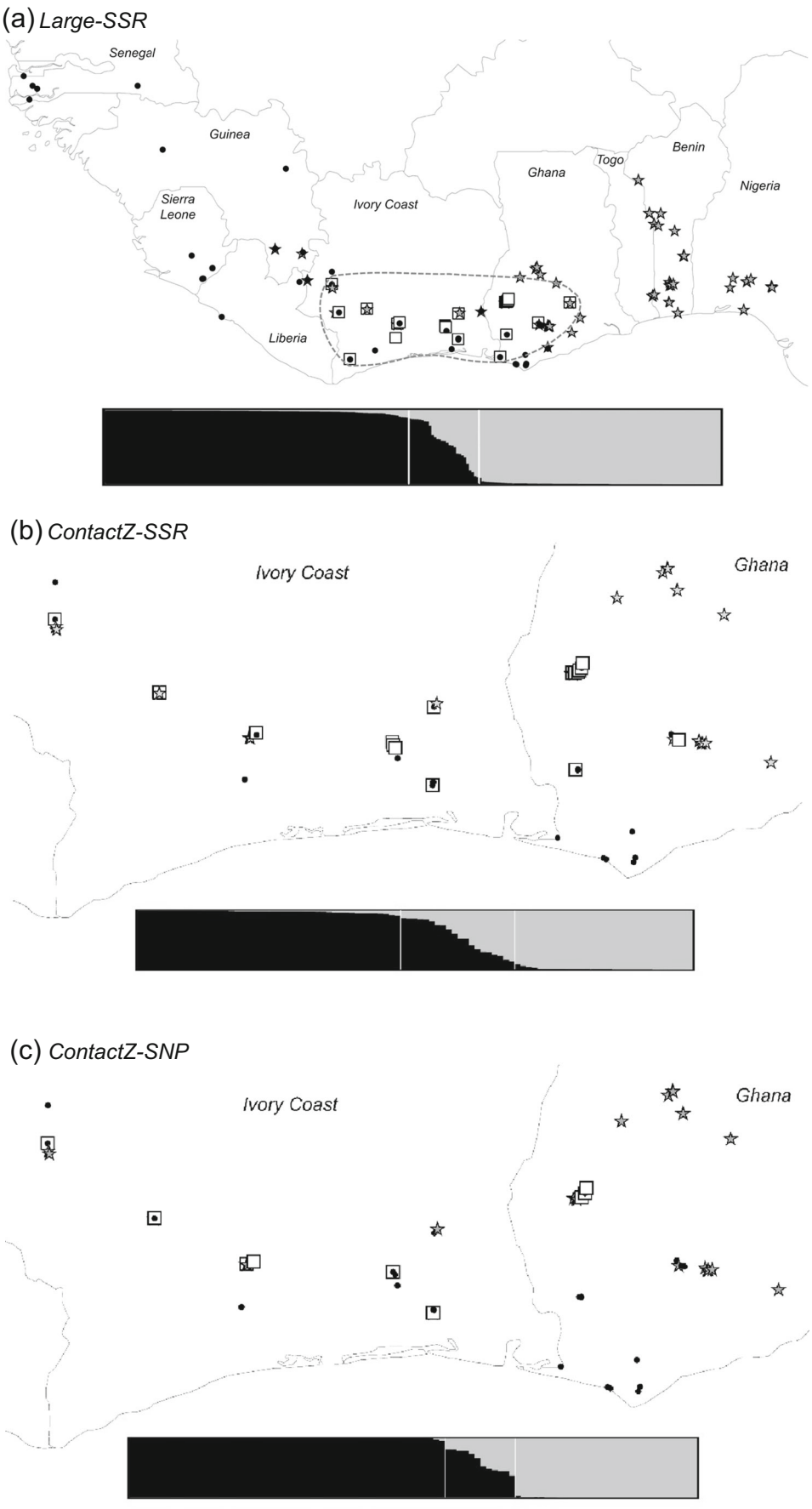
This work focused on populations of *Milicia* in West Africa (Fig. S1 in Online Resource). Two sets of samples were independently collected in this region (dried pieces of leaves or fragments of cambium). The first sampling was made up of 243 georeferenced individuals collected from Senegal to Nigeria (Fig. 1). These samples were genotyped following the method detailed in Daïnou et al. (2010, 2012) using seven nuclear microsatellite markers (SSRs): *Mex51*, *Mex63*, *Mex81*, *Mex95*, *Mex137*, *Mex163a* and *Mex202* (Ouinsavi et al. 2006). This data set will be called Large-SSR hereafter. The second sampling totalling 98 individuals took place exclusively in the Ivory Coast and Ghana in the contact zone of the two species (Fig. 1) and will be referred to as the ContactZ sample. This sample was first genotyped for 62 biallelic SNP markers which were identified from two individuals of *M. excelsa* from Benin (West Africa) and Kenya (East Africa), whereas the selection of the SNPs applied here was carried out based on a subsample of 89 individuals comprising 16% of trees morphologically assigned to *M. regia* (Blanc-Jolivet et al. 2017). The ContactZ sample was also genotyped for the same seven SSR loci employed on Large-SSR. These two data sets will be called ContactZ-SNP and ContactZ-SSR hereafter. Finally, 48 individuals of the ContactZ sample were selected from the total pool of specimens to cover the sampling area range in the Ivory Coast and Ghana, and the single-copy nuclear region *At103* (Li et al. 2008) was sequenced for each of them. *At103* was the only polymorphic region amongst the set of 12 nuclear regions we tested on our samples (*Adh*, *GScp*, *AdhC*, *RnI1*, *TPI 6rn/4rn*, *LF4 Cl1R*, *LF4 Cl4R*, *Agt1*, *Apg1*, *At103*, *PEPC E1/E2* and *PEPC E2/E3*; Li et al. 2008, J. Duminil unpublished). The sequencing method for *At103* is detailed in Daïnou et al. (2014).

Simulations for fixing the ancestry coefficient threshold to identify purebreds and verifying the performance of the various analytical tools

As a preliminary step, we used simulations to test the capacity of the seven SSR loci to disentangle purebred *Milicia* from interspecific hybrids through simulated genotypes. The simulations require first to define pure parental species, ideally from allopatric populations (e.g. Quintela et al. 2010). Because *M. excelsa* is known to be absent from Senegal to Sierra Leone and *M. regia* should not occur eastwards of Togo, we considered the respective areas for harbouring appropriate samples for such simulations. However, hybrids—if they exist—may exceptionally be found in populations far away from Ghana and the Ivory Coast, so that we decided to check first the purity of the candidate pure populations before the simulations. The Bayesian algorithms implemented into Tess (Chen et al. 2007), Structure (Pritchard et al. 2000) and NewHybrids (Anderson and Thompson 2002) were used for this purpose. Using the dataset Large-SSR, we a priori considered the Beninese and Nigerian samples, totalling 53 individuals, as pure representatives of *M. excelsa* and similarly, 45 individuals from Senegal to Sierra Leone as a priori pure individuals of *M. regia*. Tess and Structure were run for $K_{max} = 2$ clusters under conditions of no spatial autocorrelation and independent allele frequencies. NewHybrids was run under the 'Jeffrey priors' settings assuming six genotype frequency categories: purebred *M. excelsa* (*Pe*), purebred *M. regia* (*Pr*), F1 hybrids (*F1*), F2 hybrids (*F2*), backcrossed F1 to purebred *M. excelsa* (*BckExc*) and backcrossed F1 to purebred *M. regia* (*BckReg*). The outputs confirmed our assumptions. Tess assigned each of the 98 candidates to one of the two clusters with a minimum $q = 0.97$. Structure outputs yielded the same trend with a minimum $q = 0.93$, and NewHybrids provided the same outcome with a minimum posterior probability $q = 0.91$ except for one tree of the *M. regia* sample that presented a posterior probability of 0.87 for its expected category. Although the assignment probability of that individual from Tess and Structure were >0.98, we decided to exclude it from the following analyses. From that point, we considered these two groups as pure parental populations, assuming a minimum ancestry coefficient $q = 0.90$ to separate purebreds from hybrids.

The next step was to verify whether the programs Tess, Structure and NewHybrids were able to identify hybrids of different generations in a mixed sample when running on these combinations of markers and samples. To do so, we used the pure species-specific samples to generate five simulated data sets (*Sim1-SSR*, *Sim2-SSR*, etc.) each comprising 100 samples for each of the pure populations (*Pe* and *Pr*) and 10 samples for each of the four hybrid categories (*F1*, *F2*, *BckExc* and *BckReg*). Simulated genotypes were obtained using Hybridlab (Nielsen et al. 2006). Each simulated dataset totalled 240 individuals including 17% hybrids.

Because the SNP data (ContactZ-SNP) may be more accurate in detecting hybrids (see the "Results" section) and although it was difficult to identify a priori purebreds in the contact zone, we also conducted a simulation to verify (i) whether these markers were able to detect different categories of hybrids and (ii) if they might underestimate or overestimate hybrid proportions. Based on the outputs from Tess and NewHybrids on ContactZ-SNP (see the "Results" section), we considered individuals assigned to a cluster with $q > 0.90$ as purebreds, and as done with Large-SSR, we created a simulated data set of 240 individuals comprising 100 *Pe*, 100 *Pr* and 10 individuals of each hybrid category (*F1*, *F2*, *BckReg* and *BckExc*). Tess and NewHybrids were run on that simulated data set, and we compared the percentages of correct assignment. The software Structure was not used for


(a) Large-SSR
Senegal
Guinea
Sierra Leone
Liberia
Ivory Coast
Ghana
Togo
Benin
Nigeria
(b) ContactZ-SSR
Ivory Coast
Ghana
(c) ContactZ-SNP
Ivory Coast
Ghana

◀ **Fig. 1** Results of clustering of *Milicia* samples using Tess (Chen et al. 2007). Datasets Large-SSR (**a**), ContactZ-SSR (**b**) and ContactZ-SNP (**c**). The *dashed area* on Large-SSR represents the contact area where putative hybrids were detected. For each data set, the *black circles* stand for pure *M. regia* individuals, *stars* represent pure *M. excelsa* and *squares* are putative hybrids. Individual membership coefficients are shown as *bar plots* below each map (*black shade* for *M. regia* and *grey shade* for *M. excelsa*); the *vertical white lines in the bars* delimit the putative interspecific hybrids ($q < 0.90$)

this because previous results showed that it performs very similarly to Tess.

The method proposed by Fitzpatrick (2012) to identify hybrids requires that parental allele frequencies be known for the loci used and the approach perform better with diagnostic markers (one allele fixed in each parental population). The author also showed that at least 50 loci are necessary for the reliable detection of a hybrid class and uncertainty in parental allele frequencies may affect the final output. We have no diagnostic locus amongst the SSR loci, and the SNPs exhibited only few diagnostic markers in the studied populations (Table S1 in Online Resource). These limitations led us to test first the method using some of the simulated data sets from SSRs (only one of the five simulated datasets was randomly selected) and SNPs as described above. The method is implemented in the R package HIest (Hybrid index estimation; available at http://cran.r-project.org/web/packages/HIest/index.html), and the interclass heterozygosity index H was plotted against the ancestry coefficient S (similar to q in Tess and Structure).

Detection of putative hybrids in the real data sets

Using nuclear SNPs and SSRs From the simulated data sets, Tess and Structure resulted in similar performances in detecting pure individuals, whereas NewHybrids was slightly better in identifying the hybrids (see the "Results" section). We retained Tess and NewHybrids for the analyses on the real data sets. We first ran Tess on each dataset (Large-SSR, ContactZ-SSR and ContactZ-SNP; $K_{max} = 2$ for 20 runs of 100,000 sweeps each; admixture model; no consideration of individuals' location) to assign the samples to one of the two species. The 50% best runs (with the lowest values of DIC) were retained, and the membership coefficients of each individual in regard to the two genetic clusters were averaged using Clumpp (Jakobsson and Rosenberg 2007). As mentioned in the "Simulations for fixing the ancestry coefficient threshold to identify purebreds and verifying the performance of the various analytical tools" section, we considered all individuals with a coefficient membership $q \geq 0.90$ as pure individuals based on the minimum threshold derived from the previous verification. A graph of assignment probabilities was constructed using Distruct (Rosenberg 2004) to illustrate the final output of each data set. Second, NewHybrids was run on ContactZ-SSR and ContactZ-SNP with the same settings as on the simulations. Third, the CRAN package HIest was used to plot the ancestry coefficient S against the interclass heterozygosity H for ContactZ-SSR and ContactZ-SNP. We used SPAGeDi 1.5 (Hardy and Vekemans 2002) to calculate allele frequencies in parental populations ($q \geq 0.90$) required to run HIest.

Using the nuclear gene region *At103* Following the protocol described in Daïnou et al. (2014), chromatograms were imported into CodonCode Aligner (version 6.0.2; CodonCode Corporation, Dedham, MA, USA) to align the forward and reverse chromatograms of each individual. This alignment step was successful for 39 of the 48 sequenced individuals. For eight other individuals, each forward and reverse chromatograms contained multiple double peaks, as expected in the case of length-variant heterozygotes (Flot et al. 2006). We manually corrected the double peaks and used Champuru (Flot 2007) for retrieving the haplotypes of these individuals, which were added to those from the first 39 individuals. The algorithms PHASE (Stephens and Donnelly 2003) and SeqPHASE (Flot 2010) were used to infer the corresponding haplotypes. For further analyses, only the haplotypes with a posterior probability ≥0.90 were retained; this corresponded to a subsample of 38 individuals. A median-joining network (Bandelt et al. 1999) was constructed using the program Network (Fluxus Technologies) and turned into a haploweb by adding curves connecting haplotypes co-occurring in heterozygous individuals.

Results

Outputs from the simulated SSR and SNP genotypes

Based on the minimum assignment probability observed with pure individuals, we assumed that purebreds should present an ancestry coefficient $q \geq 0.90$ for SSR genotypes. For *M. excelsa*, Tess yielded a perfect score (100% of correct assignment) that was close to the ones of Structure and NewHybrids (99.4 and 98.0%, respectively; Fig. 2). The same trend was obtained regarding the recognition of pure individuals of *M. regia*. All the tested programs resulted in similar average assignment percentages in the range of 97.0–99.6% (Fig. 2 and Table S2 in Online Resource). Performances slightly differed when detecting hybrids from SSR-simulated genotypes (Fig. 2). Globally, there was a decreasing capacity in identifying hybrids beyond F1; the mean correct identification can be ranked as follows: *F1* > *F2* > *BckExc* > *BckReg* (Fig. 2 and Table S2 in Online Resource). When all the simulated categories of hybrids were pooled together (*F1*, *F2*, *BckReg* and *BckExc*), NewHybrids correctly recognised

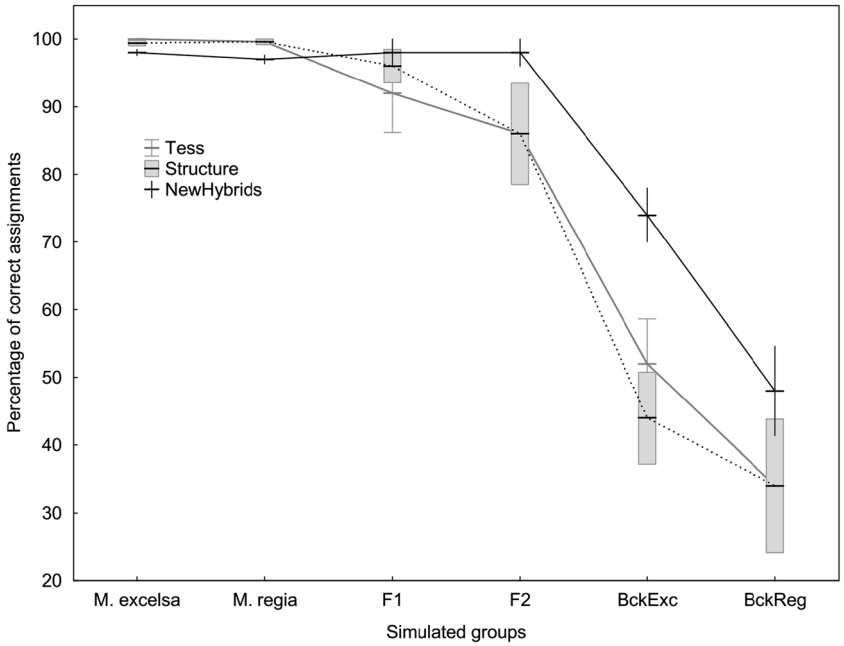


Fig. 2 Correct assignment percentages (mean and standard error based on five data sets) of Tess (Chen et al. 2007), Structure (Pritchard et al. 2000) and NewHybrids (Anderson and Thompson 2002) from data sets that comprised simulated purebreds and interspecific hybrid individuals of *Milicia excelsa* and *M. regia*

79.5% of the hybrids, whereas Tess and Structure yielded 66.0 and 65.0% success rates, respectively (Table S2 in Online Resource). NewHybrids better identified *BckExc* individuals than those of *BckReg*. Most of the simulated samples of the latter group were considered as purebreds of *M. regia* (Table S2 in Online Resource).

Hence, given this microsatellite dataset of *Milicia*, Tess and Structure seemed more appropriate than NewHybrids in recognising pure individuals, but overall, the detected pure samples should be overestimated because a substantial proportion of hybrids from backcrosses may be included in the pure species groups. On the other hand, NewHybrids seemed more reliable in detecting hybrids. Taking into account the limits of each program, we retained Tess (the results of which were very close to those of Structure in our data set) and NewHybrids to analyse the real data sets.

Regarding the simulated SNP genotypes, Tess perfectly identified all the simulated purebreds (minimum assignment coefficient $q_{min} = 0.98$) and the hybrids. Individuals of F1 and F2 were easily recognised (with q ranging from 0.41 to 0.61), whereas individuals of *BckExc* and *BckReg* exhibited assignment coefficients closer to their respective parents (with q ranging from 0.67 to 0.88). No misclassification was observed in NewHybrids outputs. Each purebred and hybrid individual was assigned to its category with a minimum assignment coefficient of 0.98. This congruence may indicate that the outputs from ContactZ-SNP were more reliable than ContactZ-SSR.

From both SSR- and SNP-simulated genotypes, HIest confirmed that the limited number of diagnostic markers in our data sets may affect the performance of its analyses. However, the SNP loci tended to produce an output more congruent with the expectations. The different simulated categories (*Pe*, *Pr*, *F1*, *F2*, *BckExc* and *BckReg*) were globally placed at the expected positions in the ternary plot surface, whereas SSRs yielded unclear barriers amongst these groups (Fig. 3a, b).

Clustering analysis on real data sets using Tess

Large-SSR confirmed that *M. regia* occurs in the western part of West Africa, whereas *M. excelsa* is exclusively found eastwards. The two species co-occur in Ghana and the Ivory Coast. A fraction of 11.5% of the sample (28 trees) located in the two aforementioned countries could not be assigned to any cluster at a threshold of $q = 0.90$ (Fig. 1a). The other data sets, ContactZ-SSR and ContactZ-SNP (Ivory Coast to Ghana), yielded results that were mostly congruent. (i) The majority of individuals with $q < 0.90$ were found in the Ivory Coast. (ii) In Ghana where we had a better north-south distribution of the samples, pure populations of *M. excelsa* did not occur or may be very scarce in the coastal zone, whereas *M. regia* may not occur in the northern part of the sampling zone (Fig. 1b, c). Globally, SNP loci detected less putative hybrids than SSRs, with a proportion of 12.4% for ContactZ-SNP compared to 20.6% for ContactZ-SSR (Table 1). The two types of markers agreed well in identifying pure *M. excelsa* individuals with 87.9% of congruence when considering all individuals assigned to *M. excelsa* by any of the programs. The corresponding percentage was 72.4 for *M. regia*.

When we considered only the ContactZ-SNP data set, the cluster of putative hybrids was dominated by individuals genetically closer to F1 and *M. regia* than to *M. excelsa* (Fig. 4a). The discontinuity in the range of assignment probabilities (Fig. 4a) also suggested that interspecific hybridisation should

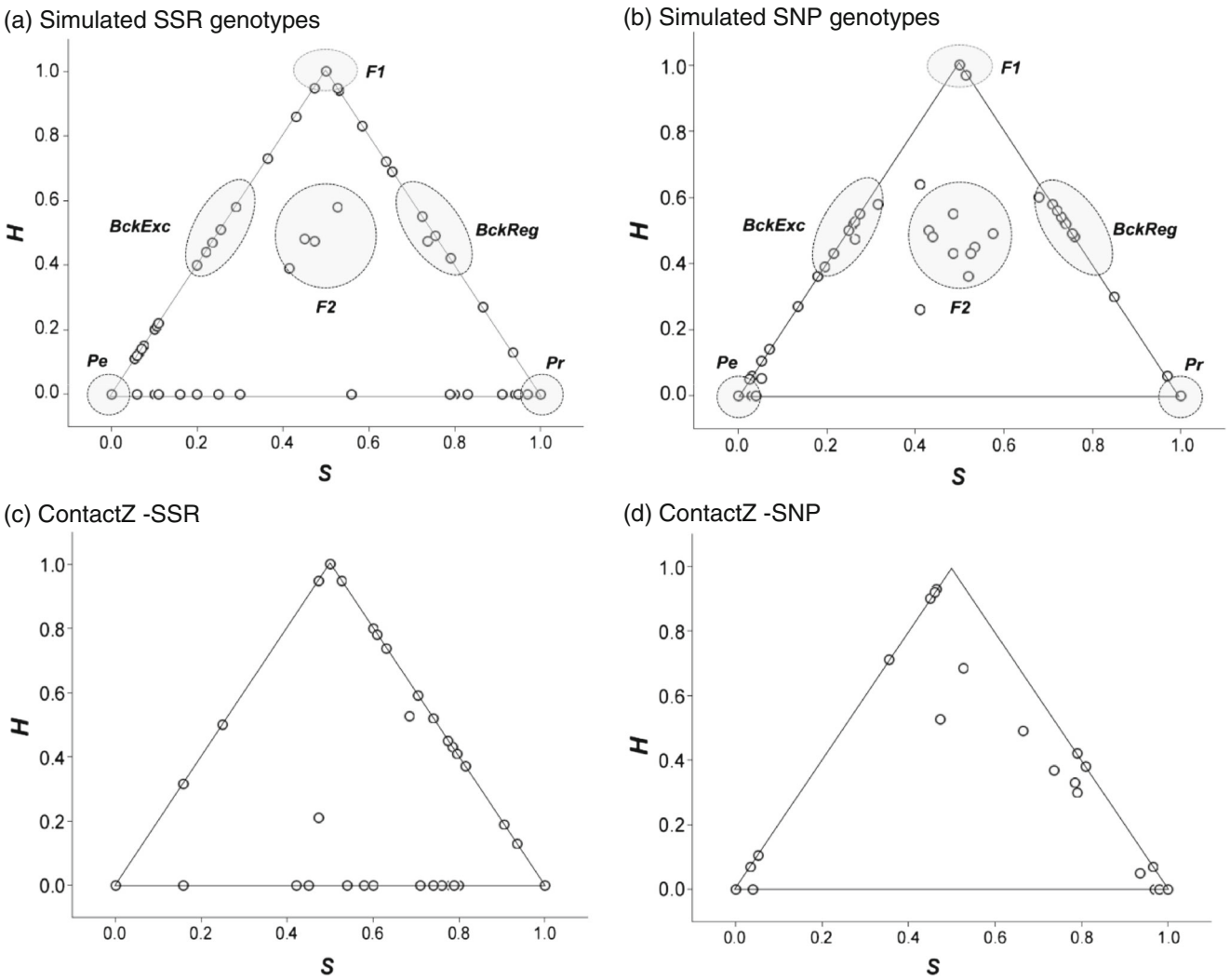


Fig. 3 Ternary plots of ancestry coefficient *S* against interclass heterozygosity index *H* (fraction of loci at which individuals have alleles from both pure populations) based on simulated SSR genotypes (**a**), simulated SNP genotypes (**b**), ContactZ-SSR (**c**) and ContactZ-SNP (**d**). The *shaded areas* in **a**, **b** show the expected positions of the simulated categories *Pe*, *Pr*, *F1*, *F2*, *BckReg* and *BckExc*

not be particularly common in this contact zone, in contradiction with the outcomes from ContactZ-SSR that indicated a higher occurrence of interspecific hybridisation (Fig. 4b). The diversity parameters of the different groups (parental populations and group of non-assigned individuals) following Tess are presented in Table S3.

Hybridisation pattern derived from NewHybrids outputs

Using NewHybrids, ContactZ-SSR exhibited more putative hybrids than ContactZ-SNP: 36 vs. 12 (Table 1). When these results were compared to the Tess outputs, the patterns from SNP markers were more consistent as both programs provided exactly the same outcomes. The whole sample of 97 individuals may comprise 12 putative hybrids (12.4% of the sample; Table 1), 54 pure individuals of *M. regia* and 31 pure trees of *M. excelsa*. NewHybrids considered only one of the 12

Table 1 Number of individuals of *Milicia* identified as putative interspecific hybrids in a sample of 97 individuals from a contact area genotyped at SNPs and SSRs

	Samples from the contact zone (ContactZ; $n = 97$)		
	SSR	SNP	SSR and SNP
Tess	20	12	8
NewHybrids	36	12	12
Tess and NewHybrids	20	12	8

Detection was derived from Tess (Chen et al. 2007) and NewHybrids (Anderson and Thompson 2002) (assignment threshold was fixed at $q = 0.90$ for purebreds)

SNPs single-nucleotide polymorphism loci, *SSRs* simple sequence repeat loci (or microsatellite)

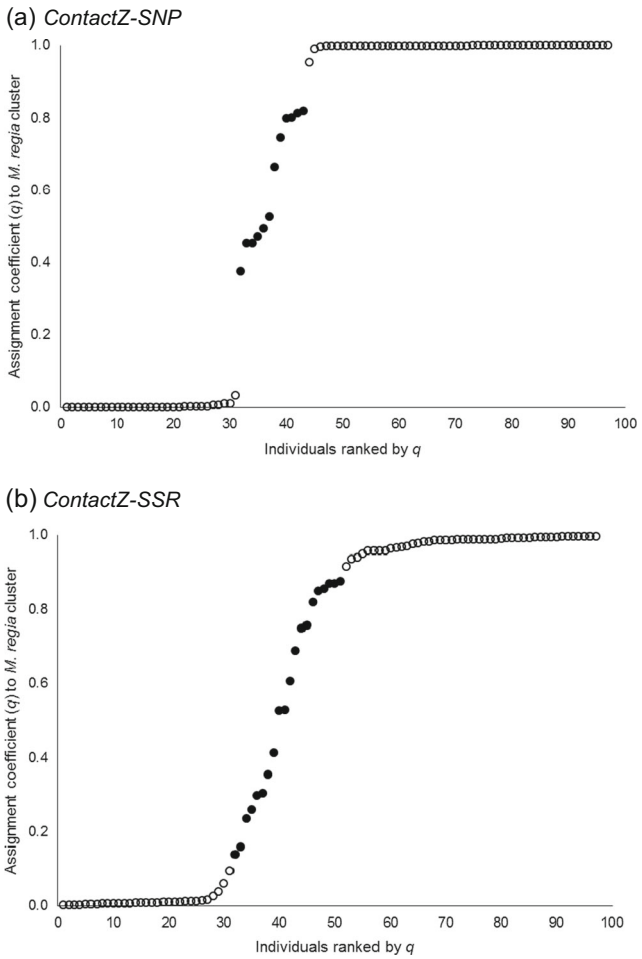


Fig. 4 Assignment probability q for 97 *Milicia* trees genotyped with SSRs and SNPs. The results are based on the Bayesian clustering algorithm implemented in Tess (Chen et al. 2007). In each figure, individuals were ranked from the lowest to the highest values of q. Individuals represented as a *black circle* are the putative hybrids whenever $0.10 < q < 0.90$

putative hybrids as genetically close to *M. excelsa* ($q = 0.82$ for *BckExc*; Table S2 in Online Resource), whereas half of that sample (six trees) were identified as backcrossed to *M. regia*, with the q ranging from 0.90 to 1.00. Three individuals were assigned to *F1* with a high posterior probability ($q = 1$). The remaining two individuals may belong to the F2 group ($q = 0.94$–1.00).

Pattern of hybridisation according to HIest

Although the power of the microsatellites appeared rather small for use in HIest, the results from the real data sets ContactZ-SSR and ContactZ-SNP were quite congruent (Fig. 3c, d): (i) F1 hybrids and backcross to *M. regia* may represent the majority of the hybrids; (ii) although HIest on ContactZ-SSR suggested possible backcrosses to *M. excelsa*, such events seemed much less frequent than the scenario provided by Tess. More interestingly, the 12 individuals previously identified as hybrids by the SNPs using Tess and NewHybrids were exactly the same with intermediate values of S (0.36–0.81) and significant values of H (0.33–0.93; $H = 0$ is expected for pure individuals, whereas $H = 1$ is expected for *F1* individuals).

At103 haploweb: confirmation of interspecific hybridisation

The alignment length of *At103* sequences in West African samples of *Milicia* was 406 bp. A total of 15 variable sites (14 polymorphic sequence characters and 1 microsatellite variant) defined 10 haplotypes. The sequenced samples recognised as purebreds of *M. regia* according to SNPs harboured five haplotypes numbered 1–5, whereas all those assigned to the *M. excelsa* cluster harboured five other haplotypes numbered 6–10 (Fig. 5). Hence, the eight individuals that were classified as hybrids according to the SSR analysis but as pure breeds according to SNPs did not show interspecific heterozygosity at the *At103* locus (Tables S4 and S5 in Online Resource).

Amongst the 12 individuals that were identified as hybrids by Tess and NewHybrids based on SNP markers, eight were successfully sequenced for the nuclear gene region *At103* (Table S4 in Online Resource). Four of them, which were assigned to *BckReg* or *F2* according to NewHybrids and SNPs, harboured the *At103* characteristic of *M. regia*. The four remaining individuals, assigned in NewHybrids to the categories *F1* (two samples), *F2* (one sample) and *BckReg* (one sample), were heterozygous at *At103* with a combination of haplotypes from the two species (Fig. 5). This confirmed that SNPs may be more accurate than SSRs in separating the two species and identifying hybrids in West Africa.

Discussion

Interspecific hybridisation in plants, which tends to be the rule rather than the exception (Hegarty and Hiscock 2005), is very poorly documented in tropical biomes (Rieseberg and Wendel 1993; Ley and Hardy 2017). It has been suspected in *Milicia* in West Africa, mainly on the basis of leaf traits (Hawthorne 1995). Genetic signs of possible hybridisation between *M. excelsa* and *M. regia* have been reported by Daïnou et al. (2014) using microsatellites. Here, we used three types of nuclear markers (SSRs, SNPs and the single-copy gene *At103*) to test whether interspecific hybridisation occurs in *Milicia*. The hybrids tend to be limited to the contact zone in the Ivory Coast and Ghana, although recurrent backcrosses may obscure any chance of detecting later generations of hybrids far away from that zone of co-occurrence. The Bayesian algorithm implemented in NewHybrids pointed out advanced generations of hybrids as also indicated by the Tess clustering outputs on SNP genotypes. This

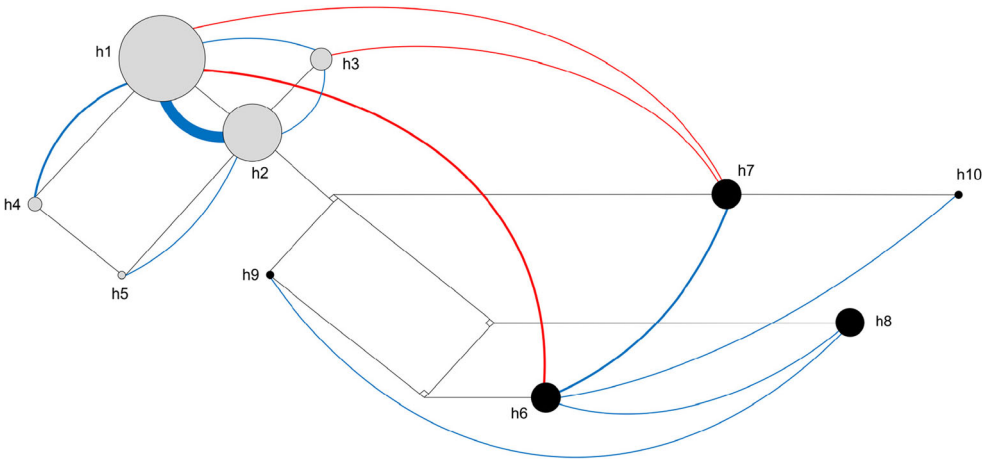


Fig. 5 Haploweb of *Milicia* sequences based on the nuclear gene *At103*. Haplotypes h1–h5 are found in *M. regia* (*grey circles*); h6–h10 represent *M. excelsa* (*black circles*). The *coloured curves* stand for heterozygous links, with the *blue curves* for intraspecific heterozygotes and the *red curves* for interspecific heterozygotes. The *sizes of the circles* are proportional to the number of haplotypes, and the *widths of the curves* are proportional to the number of heterozygous individuals

can be interpreted as a sign of hybrid fertility. Microsatellites suggested more hybrids than did SNPs and *At103*, although we expected the reverse pattern according to the results based on simulated genotypes. Sequences at *At103* were more in accordance with SNPs than with SSRs. Hence, our results highlighted differences between the types of markers in their ability to document hybridisation, but combining various approaches provided us a clear pattern. In addition, introgression was found to be unidirectional in *Milicia*: No backcrossing towards *M. excelsa* was detected.

Recurrent but asymmetric introgression between *M. regia* and *M. excelsa*

Based on genetic distances computed from RAPD loci in five populations of *Milicia*, Ofori et al. (2001) suggested that Ghanaian populations of *Milicia* located in the moist semi-deciduous forest zone are either composed of both species in similar abundances or a mixture of the two species and their putative hybrids. The second hypothesis is in agreement with our results. Indeed, 11 of the 12 individuals with a hybrid status according to SNP data were located in a latitudinal belt ca. 110 km wide at 6–7°N in the Ivory Coast and Ghana (Table S4 in Online Resource). Those latitudes are at the heart of the moist semi-deciduous forests of West Africa (Poorter et al. 2004; Oduro et al. 2012).

Recurrent hybridisation appears to occur as we detected different generations of hybrids. As 51 individuals were collected within the 6–7°N latitudinal belt, the proportion of hybrids in the contact zone can be estimated at 21.5%. This is likely an underestimate since backcrossed individuals appear common (they totalled 7 out of the 12 hybrids according to the SNP data), and obviously, some of these hybrids will not be detected as they will be genetically very close to one of the parental species (Table S4 in Online Resource). Taking into account this high proportion of hybrids in the contact zone, the hybrid fertility pattern suggested by advanced generations of hybrids and the fact that the hybrids reach the adult stage (diameter of 25–70 cm at breast height; Table S4 in Online Resource), there seem to be weak barriers to mating between the two species. However, the unidirectional hybridisation pattern we found raises questions regarding the ease of this hybridisation.

Asymmetric introgression may occur through two mechanisms: either because one of the hybridising species largely dominates in terms of population density (Levin et al. 1996; Rhymer and Simberloff 1996) or due to particular biological barriers, either prezygotic or postzygotic, that limit hybrid formation (Rieseberg and Carney 1998). In terms of population density, Nichols et al. (1998) and Hawthorne (1995) reported that the population density of *Milicia* clearly increased from the wet evergreen forests to the semi-deciduous ones. Unfortunately, this information was based on inventory data that did not distinguish the two species due to their high morphological similarity. Based on the distribution maps in Poorter et al. (2004), both species should be well represented in the moist semi-deciduous forests. Other works dealing with forest composition in this type of forest in Ghana mentioned only *M. excelsa* (e.g. Addo-Fordjour et al. 2009; Pappoe et al. 2010), probably because *M. regia* is expected more southwards in wetter forests. Our random sampling suggested that both species are well represented, 31 individuals of *M. excelsa* vs. 54 of *M. regia*, knowing that some backcrossed hybrids may lead to an overestimation of *M. regia* occurrence. Therefore, unbalanced population densities should not explain the absence or scarcity of backcrosses involving *M. excelsa*.

Prezygotic barriers may include temporal flowering gap, and gametic competition or incompatibility (Rieseberg and Carney 1998) should therefore be examined. *Milicia* species are deciduous and wind pollinated (Joker 2002; Njokuocha 2006). To our knowledge, only one study examined the flowering phenology of the two species, precisely in a population of the semi-deciduous forest zone of Ghana. Despite a global overlap of the reproduction periods of the two species, blooming was reported to start later for *M. excelsa* than for *M. regia* (Nyong'o et al. 1994). We also know that the onset of flowering in female trees coincides with the moment when most male trees have

reached blooming maturity and flowering lasts only a few weeks in *Milicia* (Nyong'o et al. 1994; Daïnou et al. 2012). Therefore, some late bloomers of *M. regia* female trees may be pollinated by pure males of *M. excelsa*, but the reverse mating is very unlikely because the pollen stock of *M. regia* would be absent or limited when females of *M. excelsa* start blooming. Hence, most hybrids should be fathered by *M. excelsa*. As maternal inheritance was shown to affect flowering in many plant species (e.g. Roach and Wulff 1987; Galloway et al. 2009), it is then plausible that F1 individuals may initiate flowering at the same moment as *M. regia*, explaining why hybrids backcross preferentially with *M. regia*. The phenological comparison of the two species by Nyong'o et al. (1994) is the only available information to explain the prezygotic barrier between these taxa. Unfortunately, Nyong'o et al. (1994) did not quantify the gap of flowering time amongst the two species, and no other work addressed this issue in *Milicia*. One may note that the existence of a clear delay of the onset of flowering time amongst species in the same habitat, if confirmed, may be a sign of environmental adaptation that would indicate a substantial degree of divergence in *Milicia*. Competition and incompatibility in gamete formation may also act as a barrier to interspecific hybridisation, as shown for instance by controlled crosses using mixed pollen loads in the Louisiana irises (Carney et al. 1994), but this phenomenon is rarely unidirectional (Carney et al. 1994; Olrik and Kjaer 2007; Lepais et al. 2013). Finally, postzygotic barriers may also contribute in explaining this asymmetric hybridisation pattern.

Relative performance of the different types of markers and analysis approaches in the study taxa

Studies have shown that microsatellites are able to detect first and advanced generations of hybrids when they exist (Barilani et al. 2007; Quintela et al. 2010). However, homoplasy, especially in closely related species, and null alleles can limit the ability of microsatellites to identify hybrids. Moreover, the identification of interspecific hybrids can be challenging when the genetic divergence between the taxa is mainly due to variation in allele frequency rather than to the occurrence of diagnostic alleles (Rhymer and Simberloff 1996). Using simulated microsatellite genotypes with no clear diagnostic alleles for the hybridising taxa, Vähä and Primmer (2006) proposed that a minimum of 12 loci are needed when $F_{ST} = 0.21$ between parental populations. In the case of *Milicia*, F_{ST} at SSR loci between pure populations of *M. regia* and *M. excelsa* was 0.18 (result not shown), with the advantage that we observed some diagnostic alleles in the set of seven loci (Table S1 in Online Resource). However, this is clearly insufficient for the HIest approach, which requires dozens of diagnostic markers (Fitzpatrick 2012). For that approach, our SNP data set was more informative. Vähä and Primmer (2006) reported that NewHybrids and Structure performed similarly well in detecting hybrids but that most of the backcrossed individuals were assigned to the parental populations unless the number of loci was increased to 48. Our findings confirmed that report. Misclassification of individuals from backcrosses was important, highlighting the need to use more loci when advanced generations of hybrids are present. However, similar with Vähä and Primmer (2006), we also observed that NewHybrids marginally outperformed Structure (but also Tess) in detecting backcrossed samples.

SNPs also have a proven record of their capacity to detect hybrid classes (e.g. Marsden et al. 2011; Fogelqvist et al. 2015). NewHybrids, Tess and HIest perfectly converged in assigning individuals to either the purebred or hybrid groups using the SNP loci. Therefore, we are tempted to assume that our set of 62 SNPs was more powerful in identifying hybrids than our seven SSR loci, for which the degree of congruence between the two programs was not excellent (Table 1). SSRs detected an introgression in both directions, whereas SNPs found introgression events towards *M. regia* only. As the sequences of the single-copy gene *At103* did not reveal any backcross to *M. excelsa* either, this again tends to show that the SNP loci employed here may be more powerful than the SSRs in hybridisation studies. In addition, none of the individuals identified as hybrids by SSRs exclusively (using NewHybrids and Tess) was confirmed as such by the *At103* haploweb (Table S2 in Online Resource). Future investigations should use uniparentally transmitted chloroplast DNA of hybrids to definitely confirm the direction of introgression in West African species of *Milicia*.

SNPs and *At103* sequences were not perfectly congruent. *At103* suggested that four individuals amongst the 12 hybrids (based on the SNP outcomes) may be purebreds of *M. regia* instead of hybrids. We expect the findings from the SNPs to be more reliable because they are based on many loci scattered across the genome, whereas *At103* is only one particular locus. A single locus of 406 bp obviously cannot reveal all possibilities of recombination in a genome. In fact, only F1 hybrids are expected to be always detectable at *At103* because F2 hybrids and backcrosses have 50% chance of inheriting haplotypes from a single species at a given locus. It is also important to note that all four hybrids detected using *At103* sequences were length-variant heterozygotes with chromatograms presenting numerous double peaks (Flot et al. 2006). As a result, their forward and reverse chromatograms could not be assembled easily and were initially discarded, which would have led to a significant bias in our analyses by suggesting that hybridisation does not occur at this locus. This emphasises how proper scrutiny of chromatogram pairs is crucial in order to detect length-variant heterozygotes and distinguish them from sequencing failures in studies aimed to detect hybrids using nuclear sequence markers.

Conclusive remarks

Overall, hybridisation appears quite common in the contact zone of *Milicia* populations, especially taking into account that the

proportion of hybrids was probably underestimated by our approaches. The interspecific hybrids, especially the F1 individuals, seem fertile although ascertaining that it will require additional direct observations in the field: the pollen viability of the two reciprocal F1 hybrids, seed rain and germination tests of hybrids vs. purebreds. The fitness of the seedlings and saplings of these groups should be also compared. As both species are wind-pollinated and seeds of *Milicia* are dispersed by flying animals (bats and birds) that can drop them far away from the mother trees, local introgression may occur more frequently than expected. A delay in flowering time between species and between sexes, combined with a maternal inheritance of flowering time, might explain the absence of introgression towards *M. excelsa*, although more data are needed to test this hypothesis. Our work also indicates that SNPs are probably more reliable in investigating hybridisation patterns than SSRs. However, using only one type of markers for such a study is not recommended; instead, one should combine various types of loci and approaches to obtain reliable insights on hybridisation and introgression issues.

Acknowledgments The development of the SNP markers was carried out with the financial support of the International Tropical Timber Organization (ITTO) through the project PD 620/11 Rev.1 (M): "Development and implementation of species identification and timber tracking in Africa with DNA fingerprints and stable isotopes". SNP genotyping was performed at the Genomic and Sequencing Facility of Bordeaux (grants from the Conseil Regional d'Aquitaine No. 20030304002FA and 20040305003FA, the European Union, FEDER No. 2003227 and "Investissements d'avenir, Convention attributive d'aide" No. ANR-10-EQPX-16-01). SSR genotyping and DNA sequencing were funded by the Belgian Fund for Scientific Research (F.R.S-FNRS, grant T0163.13). We would like also to thank the Associate Editor Felix Gugerli, two anonymous reviewers and Armel S. L. Donkpegan for their comments, corrections and suggestions that highly contributed in improving the manuscript.

Data archiving statement The phased *Milicia* sequences of the nuclear gene *At103* were deposited in GenBank under the accession numbers MF541241-MF541316. The *At103* chromatograms of the length-variant heterozygotes were deposited in Dryad together with the nuclear SSR and SNP data at http://dx.doi.org/10.5061/dryad.m457d.

References

Addo-Fordjour P, Obeng S, Anning A, Addo M (2009) Floristic composition, structure and natural regeneration in a moist semi-deciduous forest following anthropogenic disturbances and plant invasion. Int J Biodivers Conserv 1:021–037

Anderson EC, Thompson EA (2002) A model-based method for identifying species hybrids using multilocus genetic data. Genetics 160: 1217–1229

Bandelt H-J, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16:37–48

Barilani M, Sfougaris A, Giannakopoulos A, Mucci N, Tabarroni C, Randi E (2007) Detecting introgressive hybridisation in rock partridge populations (*Alectoris graeca*) in Greece through Bayesian admixture analyses of multilocus genotypes. Conserv Genet 8:343–354

Blanc-Jolivet C, Kersten B, Daïnou K, Hardy OJ, Guichoux E, Delcamp A, Degen B (2017) Development of nuclear SNP markers for genetic tracking of Iroko, *Milicia excelsa* and *Milicia regia*. Conserv Genet Resour. doi: 10.1007/s12686-017-0716-2

Carney SE, Cruzan MB, Arnold ML (1994) Reproductive interactions between hybridizing irises: analyses of pollen-tube growth and fertilization success. Am J Bot 81:1169–1175

Chen C, Durand E, Forbes F, François O (2007) Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study. Mol Ecol Notes 7:747–756

Chevalier A (1912) Novitates florae africanae—Urticaceae. B Soc Bot Fr 58:207–211

Chevalier A (1917) Les végétaux utiles d'Afrique tropicale française—La forêt et les bois du Gabon. Challamel, Paris

Craft KJ, Ashley MV, Koenig WD (2002) Limited hybridization between *Quercus lobata* and *Quercus douglasii* (Fagaceae) in a mixed stand in central coastal California. Am J Bot 89:1792–1798

Daïnou K, Bizoux JP, Doucet JL, Mahy G, Hardy OJ, Heuertz M (2010) Forest refugia revisited: nSSRs and cpDNA sequences support historical isolation in a wide-spread African tree with high colonization capacity, *Milicia excelsa* (Moraceae). Mol Ecol 19:4462–4477

Daïnou K, Laurenty E, Mahy G, Hardy OJ, Brostaux Y, Tagg N, Doucet JL (2012) Phenological patterns in a natural population of a tropical timber tree species, *Milicia excelsa* (Moraceae): evidence of isolation by time and its interaction with feeding strategies of dispersers. Am J Bot 99:1453–1463

Daïnou K, Mahy G, Duminil J, Dick CW, Doucet J-L, Donkpégan ASL, Pluijgers M, Sinsin B, Lejeune P, Hardy OJ (2014) Speciation slowing down in widespread and long-living tree taxa: insights from the tropical timber tree genus *Milicia* (Moraceae). Heredity 113:74–85

Daïnou K, Blanc-Jolivet C, Degen B, Kimani P, Ndiade Bourobou D, Donkpégan ASL, Tosso F, Kaymak E, Bourland N, Doucet J-L, Hardy OJ (2016) Revealing hidden species diversity in closely related species using nuclear SNPs, SSRs and DNA sequences—a case study in the tree genus *Milicia*. BMC Evol Biol 16:259

DeFaveri J, Viitaniemi H, Leder E, Merilä J (2013) Characterizing genic and nongenic molecular markers: comparison of microsatellites and SNPs. Mol Ecol Resour 13:377–392

Doyle JJ (1995) The irrelevance of allele tree topologies for species delimitation, and a non-topological alternative. Syst Bot 20:574–588

Edwards CE, Soltis DE, Soltis PS (2008) Using patterns of genetic structure based on microsatellite loci to test hypotheses of current hybridization, ancient hybridization and incomplete lineage sorting in *Conradina* (Lamiaceae). Mol Ecol 17:5157–5174

Fehrer J, Gemeinholzer B, Chrtek J Jr, Brautigam S (2007) Incongruent plastid and nuclear DNA phylogenies reveal ancient intergeneric hybridization in Pilosella hawkweeds (*Hieracium*, Cichorieae, Asteraceae). Mol Phylogenet Evol 42:347–361

Fitzpatrick BM (2012) Estimating ancestry and heterozygosity of hybrids using molecular markers. BMC Evol Biol 12:131

Flot J-F (2007) Champuru 1.0: a computer software for unraveling mixtures of two DNA sequences of unequal lengths. Mol Ecol Notes 7: 974–977

Flot J-F (2010) SeqPHASE: a web tool for interconverting PHASE input/output files and FASTA sequence alignments. Mol Ecol Resour 10: 162–166

Flot JF, Tillier A, Samadi S, Tillier S (2006) Phase determination from direct sequencing of length-variable DNA regions. Mol Ecol Notes 6:627–630

Flot JF, Couloux A, Tillier S (2010) Haplowebs as a graphical tool for delimiting species: a revival of Doyle's "field for recombination" approach and its application to the coral genus Pocillopora in Clipperton. BMC Evol Biol 10:372

Fogelqvist J, Verkhozina AV, Katyshev AI, Pucholt P, Dixelius C, Rönnberg-Wästljung AC, Lascoux M, Berlin S (2015) Genetic and morphological evidence for introgression between three species of willows. BMC Evol Biol 15:1

Freeland JR, Kirk H, Petersen SD (2011) Molecular ecology. Wiley, Oxford

Galloway LF, Etterson JR, McGlothlin JW (2009) Contribution of direct and maternal genetic effects to life-history evolution. New Phytol 183:826–838

Gompert Z, Buerkle CA (2016) What, if anything, are hybrids: enduring truths and challenges associated with population structure and gene flow. Evol Appl 9:909–923

Granevitze Z, David L, Twito T, Weigend S, Feldman M, Hillel J (2014) Phylogenetic resolution power of microsatellites and various single-nucleotide polymorphism types assessed in 10 divergent chicken populations. Anim Genet 45:87–95

Hardy OJ, Vekemans X (2002) Spagedi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Mol Ecol Notes 2:618–620

Hawthorne WD (1995) Ecological profile of Ghanaian forest trees. Oxford Forestry Institute, Oxford

Hegarty MJ, Hiscock SJ (2005) Hybrid speciation in plants: new insights from molecular studies. New Phytol 165:411–423

Jakobsson M, Rosenberg NA (2007) Clumpp: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23:1801–1806

Joker D (2002) *Milicia excelsa* seed leaflet 63. Available at http://curis.ku.dk/ws/files/20648316/milicia_excelsa_63_int_001.pdf

Lamer JT, Ruebush BC, Arbieva ZH, McClelland MA, Epifanio JM, Sass GG (2015) Diagnostic SNPs reveal widespread introgressive hybridization between introduced bighead and silver carp in the Mississippi river basin. Mol Ecol 24:3931–3943

Lepais O, Petit RJ, Guichoux E, Lavabre JE, Alberto F, Kremer A, Gerber S (2009) Species relative abundance and direction of introgression in oaks. Mol Ecol 18:2228–2242

Lepais O, Roussel G, Hubert F, Kremer A, Gerber S (2013) Strength and variability of postmating reproductive isolating barriers between four European white oak species. Tree Genet Genomes 9:841–853

Levin DA, Francisco-Ortega J, Jansen RK (1996) Hybridization and the extinction of rare plant species. Conserv Biol 10:10–16

Ley AC, Hardy OJ (2017) Hybridization and asymmetric introgression after secondary contact in two tropical African climber species, *Haumania danckelmaniana* and *Haumania liebrechtsiana* (Marantaceae). Int J Plant Sci 178:421–430

Li M, Wunder J, Bissoli G, Scarponi E, Gazzani S, Barbaro E, Saedler H, Varotto C (2008) Development of COS genes as universally amplifiable markers for phylogenetic reconstructions of closely related plant species. Cladistics 24:727–745

Mallet J (2005) Hybridization as an invasion of the genome. Trends Ecol Evol 20:229–237

Marsden CD et al (2011) Asymmetric introgression between the M and S forms of the malaria vector, *Anopheles gambiae*, maintains divergence despite extensive hybridization. Mol Ecol 20:4983–4994

Nichols JD, Agurgo FB, Agyeman VK, Wagner MR, Cobbinah JR (1998) Distribution and abundance of *Milicia* species in Ghana. Ghana J Forest 6:1–7

Nielsen EE, Bach LA, Kotlicki P (2006) hybridlab (version 1.0): a program for generating simulated hybrids from population samples. Mol Ecol Notes 6:971–973

Njokuocha RC (2006) Airborne pollen grains in Nsukka. Nigeria Grana 45:73–80

Nyong'o RN, Cobbinah JR, Appiah-Kwarteng J (1994) Flowering and fruiting patterns in *Milicia excelsa* and *Milicia regia* Welw. Ghana J Forest 1:19–29

Oduro K, Duah-Gyamfi A, Acquah S, Agyeman V (2012) Ghana forest and wildlife handbook. Forestry Commission, Accra

Ofori DA, Cobbinah JR (2007) Integrated approach for conservation and management of genetic resources of Milicia species in West Africa. Forest Ecol Manag 238:1–6

Ofori DA, Swaine M, Leifert C, Cobbinah JR, Price AH (2001) Population genetic structure of *Milicia excelsa* characterized by using RAPD and nucleotide sequencing. Genet Resour Crop Ev 48:637–647

Olrik DC, Kjaer ED (2007) The reproductive success of a *Quercus petraea* × *Q. robur* F1-hybrid in back-crossing situations. Ann For Sci 64:37–45

Ouinsavi C, Sokpon N, Bousquet J, Newton CH, Khasa DP (2006) Novel microsatellite DNA markers for the threatened African endemic tree species, *Milicia excelsa* (Moraceae), and cross-species amplification in *Milicia regia*. Mol Ecol Notes 6:480–483

Papakostas S, Michaloudi E, Proios K, Brehm M, Verhage L, Rota J, Peña C, Stamou G, Pritchard VL, Fontaneto D, Declerck SAJ (2016) Integrative taxonomy recognizes evolutionary units despite widespread mitonuclear discordance: evidence from a rotifer cryptic species complex. Syst Biol 65:508–524

Pappoe A, Armah F, Quaye E, Kwakye P, Buxton G (2010) Composition and stand structure of a tropical moist semi-deciduous forest in Ghana. Int Res J Plant Sci 1:95–106

Pollegioni P, Olimpieri I, Woeste KE, Simoni G, Gras M, Malvolti ME (2012) Barriers to interspecific hybridization between *Juglans nigra* L. and *J. regia* L species. Tree Genet Genomes 9:291–305

Poorter L, Bongers F, Kouamé FN, Hawthorne WD (2004) Biodiversity of West African forests. CABI Publishing, Cambridge

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

Quintela M, Thulin C-G, Höglund J (2010) Detecting hybridization between willow grouse (*Lagopus lagopus*) and rock ptarmigan (*L. muta*) in central Sweden through Bayesian admixture analyses and mtDNA screening. Conserv Genet 11:557–569

Rieseberg LH, Wendel JF (1993) Introgression and its consequences in plants. In: Harrison RG (ed) Hybrid zones and the evolutionary process. Oxford University Press, Oxford, pp 70–109

Rhymer JM, Simberloff D (1996) Extinction by hybridization and introgression. Annu Rev Ecol Syst:83–109

Rieseberg LH, Carney SE (1998) Plant hybridization. New Phytol 140: 599–624

Roach DA, Wulff RD (1987) Maternal effects in plants. Annu Rev Ecol Syst 18:209–235

Rosenberg NA (2004) Distruct: a program for the graphical display of population structure. Mol Ecol Notes 4:137–138

Sites JW, Marshall JC (2004) Operational criteria for delimiting species. Annu Rev Ecol Evol S 35:199–227

Stephens M, Donnelly P (2003) A comparison of Bayesian methods for haplotype reconstruction from population genotype data. Am J Hum Genet 73:1162–1169

Tondeur G (1939) Monographie forestière du *Chlorophora excelsa* Benth. et Hook. Bulletin Agricole du Congo Belge 30:163–198

Twyford AD, Ennos RA (2012) Next-generation hybridization and introgression. Heredity 108:179–189

Vähä JP, Primmer CR (2006) Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. Mol Ecol 15:63–72

von Holdt BM, Pollinger JP, Earl DA, Parker HG, Ostrander EA, Wayne RK (2013) Identification of recent hybridization between gray wolves and domesticated dogs by SNP genotyping. Mamm Genome 24:80–88