



HAL
open science

Microsatellite-assisted identification and comparative population genetics of Malagasy olive species (*Noronhia* spp., Oleaceae)

Océane Gorrilliot, Cynthia Hong-Wa, Franck Rakotonasolo, Guillaume Besnard

► To cite this version:

Océane Gorrilliot, Cynthia Hong-Wa, Franck Rakotonasolo, Guillaume Besnard. Microsatellite-assisted identification and comparative population genetics of Malagasy olive species (*Noronhia* spp., Oleaceae). *Botany Letters*, In press, 10.1080/23818107.2021.1912636 . hal-03388612

HAL Id: hal-03388612

<https://hal.science/hal-03388612>

Submitted on 20 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Microsatellite-assisted identification and comparative** 2 **population genetics of Malagasy olive species (*Noronhia*** 3 **spp., Oleaceae)**

4
5 Océane Gorrilliot¹, Cynthia Hong-Wa², Franck Rakotonasolo^{3,4}, Guillaume Besnard^{1*}

6
7 ¹ CNRS-UPS-IRD, UMR5174, Laboratoire Évolution & Diversité Biologique, Université
8 Paul Sabatier, 118 route de Narbonne, 31062 Toulouse, France

9 ² Claude E. Phillips Herbarium, Delaware State University, 1200 N. Dupont Hwy, Dover, DE
10 19901-2277, USA

11 ³ Herbar, Département Flore, Parc Botanique et Zoologique de Tsimbazaza, BP4096,
12 Antananarivo - 101, Madagascar

13 ⁴ Kew Madagascar Conservation Centre, Lot II J 131 Ambodivoanjo, Ivandry, Antananarivo,
14 Madagascar

15
16 *Corresponding author. E-mail: guillaume.besnard@univ-tlse3.fr

17
18 **Running title:** Microsatellite genotyping of *Noronhia*

19 20 **ABSTRACT**

21 An exhaustive inventory of Madagascar's native flora is necessary to optimize its
22 conservation, particularly for highly diversified endemic groups of plants whose many
23 habitats are threatened with short-term loss. The Malagasy olive (*Noronhia* spp., Oleaceae),
24 with at least 80 species, is recognized as one of the highly diversified tree genera of the SW
25 Indian Ocean Flora. Field identification of *Noronhia* species is, however, often a challenge
26 and numerous sterile samples need posterior identification. As molecular tools may assist this
27 task, we tested cross-species amplification of 13 nuclear microsatellite loci (i.e. SSRs) among
28 a large sample of *Noronhia* taxa (60 of the 82 described species in Madagascar). Eight species
29 were also investigated at the population level with 13 nuclear and 15 plastid SSRs. A high
30 cross-species transferability of markers was demonstrated, although three nuclear loci showed
31 evidence of null alleles in some species. Based on the amount of genetic diversity observed
32 within eight diploid species, we showed that putative polyploids could be efficiently detected
33 using nuclear SSRs. Our markers thus allowed revealing one polyploid taxon, but their
34 relatively low frequency suggests that whole-genome duplication was not a major process
35 involved in the high diversification of *Noronhia* as previously hypothesized. Finally, plastid
36 lineages as well as genetic clusters identified with nuclear SSRs allowed distinguishing all six
37 species collected at Andohahela and Marojejy. These results demonstrate that microsatellites
38 can assist *Noronhia* identification in their diversity hotspots, as well as reveal contrasting
39 genetic diversity patterns (such as isolation by distance) between sympatric taxa.

40
41 **KEYWORDS:** Biodiversity inventory and conservation – IBD – Madagascar – *Noronhia* –
42 Polyploidy – Population genetics – Species identification – *tsilaitra*.

43 Introduction

44 Biodiversity inventory is an essential prerequisite for investigating species distribution
45 patterns and diversification processes, and such information is essential to define conservation
46 action in biodiversity hotspots (Willis, Gillson, and Knapp 2007). However, species
47 identification based on phenotypic criteria can be very challenging in highly diversified
48 groups and species complexes (Freudenstein et al. 2017), such that developing molecular
49 tools is often required to assist this task (Duminil and Di Michele 2009). Yet, DNA barcoding
50 approaches have also proven to be sometime limited in discriminating closely related taxa that
51 still exchange genes (e.g. Singh et al. 2012; Caetano Wyler and Naciri 2016), and thus
52 population genetics may be more efficient in uncovering cryptic species at the local scale
53 (Wang et al. 2014; McKendrick et al. 2017; Maduna et al. 2017). Furthermore,
54 polyploidization, as a major speciation process in plants, still needs to be better documented
55 in hyper-diversified tropical groups for which chromosome counts or cellular DNA content
56 measurements by flow cytometry are difficult to obtain on a comprehensive taxon sampling
57 (Husband, Baldwin, and Suda 2013), especially from herbarium samples (Viruel et al. 2019).
58 The use of highly variable cross-species markers (such as microsatellites or Simple Sequence
59 Repeats, SSRs), which can inform on the ploidy level of individuals (e.g. García-Verdugo et
60 al. 2009; Feng et al. 2018; Liang et al. 2019; Díaz Martínez, Boedeker, and Zuccarello 2020;
61 Hagl et al. 2021), represents a more integrative alternative approach.

62 In the Malagasy biodiversity hotspot, floristic inventory is still a matter of intense
63 research, and the taxonomy of numerous highly diversified groups needs re-evaluation, while
64 many habitats are threatened with short-term loss (Ganzhorn et al. 2008; Callmander et al.
65 2011; Goyder et al. 2017). The Malagasy olive (*Noronhia* Stadtm. ex Thouars, Oleaceae),
66 commonly named “*tsilaitra*”, has radiated since the Oligocene, resulting in at least 82 species
67 (Hong-Wa and Besnard 2013, 2014; Hong-Wa 2016). *Noronhia* is now recognized as one of
68 the highly diversified tree genera of the SW Indian Ocean Flora, although less diverse than
69 groups like *Dombeya* Cav. or *Diospyros* L. (Skema 2012; Linan et al. 2019). It has colonized
70 various types of habitats in Madagascar, such as coastal forests and rainforests (mainly in the
71 East and North), dry deciduous forests (in the North and West), and central High Plateau
72 gallery forests and savannas (Perrier de la Bâthie 1949, 1952; Hong-Wa 2016). The reasons
73 for the *Noronhia* diversification remain unknown, although an adaptive radiation involving
74 recurrent hybridizations and possible polyploidizations has been suggested (Hong-Wa and
75 Besnard 2013, 2014). Field identification of “*tsilaitra*” species is often a challenge, especially
76 in regions where closely related taxa locally co-exist; such as in dry and rain forests from the
77 North to the South-East (Hong-Wa and Besnard 2014). For instance, at least 25 taxa have
78 been recently recorded in the remnant forests of the Loky-Manambato, for a region of ca.
79 1000 km² (Daraina, North-East Madagascar; G. Besnard, C. Hong-Wa, J. Salmona, and A.E.
80 Ranaivoson, pers. obs.). Taxon identification is also made difficult in the field when flowers
81 and fruits are absent (especially during the dry season), or impossible to collect (e.g. trees of
82 the canopy). During our field collections, numerous sterile samples were thus not clearly
83 assigned to a species, needing posterior identification assisted by genetic markers. The
84 barcoding approach (based on plastid and nuclear ribosomal DNA) has however showed some
85 limitations to delineate species due to low nucleotide sequence divergence (Hong-Wa and
86 Besnard 2014).

87 Population genetics studies were recently carried out on *Noronhia* spp. to evaluate the
88 impact of landscape variables on their demography (Salmona, Dresen et al. 2020; Salmona,
89 Olofsson et al. 2020), but these studies have only focused on species that are very easily
90 identified (i.e. *N. lowryi* Hong-Wa and *N. spinifolia* Hong-Wa) and probably reproductively
91 well isolated from their congeners. More investigations are still necessary to document the

92 past demography and processes involved in the diversification of *Noronhia* species (e.g. local
93 adaptation, reproductive isolation, and recurrent admixture). Before tackling such questions, a
94 preliminary characterization of their genome diversity may help to identify species complexes
95 and/or polyploids, and then assist in designing the most appropriate research approaches.
96 Here, we thus first tested cross-species amplification of nuclear SSR markers recently
97 developed (Salmona, Olofsson et al. 2020) among a large sample of *Noronhia* taxa (ca. three-
98 quarters of the 82 described species in Madagascar). Eight species were also investigated at
99 the population level to compare their allelic richness and genetic diversity, as well as to test
100 for the presence of null alleles or allelic dropout. Besides, the potential utility of such multi-
101 allelic loci in determining the ploidy level of species was assessed. Nuclear and plastid SSR
102 data were finally used in two distant locations (Andohahela in the South East, and Marojejy in
103 the North East) to test their species discrimination power and their potential utility for
104 comparative studies of population genetics.

105

106 **Material and methods**

107 *Plant sampling*

108 To test the cross-species transferability of SSR makers, we first analysed 76 *Noronhia* taxa
109 from the entire range of the genus in Madagascar (Figure 1; Table S1). For each individual, a
110 voucher specimen has been deposited in at least a herbarium (Table S1). Most Malagasy
111 species (60 taxa) were represented by a single individual, whereas nine accessions remained
112 unidentified. In this initial sample, we decided to analyse four individuals of *N. louvelii*
113 H. Perrier (from three locations) to test the possible occurrence of different ploidy levels
114 within this taxon (see below). In addition, only three individuals of two species were collected
115 outside Madagascar: *N. peglerae* (C.H. Wright) Hong-Wa & Besnard from South Africa, and
116 *N. emarginata* (Lam.) Stadtm. ex Thouars from the Reunion Island and Hawaii. This latter
117 has been translocated from Madagascar to archipelagos of the Indian and Pacific Oceans as
118 well as in the Caribbean region (CABI 2020). In its new range, it is often cultivated as an
119 ornamental or vanilla tutor, but it can naturalize and is locally considered invasive (CABI
120 2020).

121 To test the utility of our markers for population genetics and assess their power at
122 discriminating species at the local scale, we then characterized *Noronhia* individuals from
123 four distant sites (Table S2): Andohahela NP in the South East (38 individuals), Marojejy NP
124 in North East (74 ind.), Ankafoobe on the northern High Plateau (10 ind.), and Itremo-Ibity on
125 the central High Plateau (73 genotypes already analysed in Salmona, Olofsson et al. 2020).
126 For each sampled tree, we collected a leaf in silica gel for DNA extraction, recorded GPS
127 coordinates, and made a herbarium specimen for later species identification. Species sampling
128 on these sites is detailed as follows: (i) Trees collected in Andohahela were assigned to six
129 forest species (Figure S1): *N. aff. boivinii* Dubard (18 ind.), *N. grandifolia* H. Perrier (7 ind.),
130 and four undetermined species: hereafter named sp1_Ando (8 ind.), sp2_Ando (2 ind.),
131 sp3_Ando (2 ind.), and sp4_Ando (1 ind.); (ii) Samples collected in Marojejy were assigned
132 to six distinct forest species (Figure S1): *N. verticillata* H. Perrier (33 ind.), *N. humblotiana*
133 (H. Perrier) Hong-Wa (26 ind.), *N. disjuncta* Hong-Wa (8 ind.), *N. louvelii* (5 ind.), and two
134 undetermined species (hereafter named sp1_Marobj and sp2_Marobj; one individual each). Note
135 that two species were collected in Marojejy along parallel transects of ca. 3.5 km:
136 *N. verticillata*, a riparian species found along the river; and *N. humblotiana*, an understory
137 forest species collected along the path at ca. 10-50 m from the river; (iii) Lastly, all trees

138 collected in Ankafobe were assigned to *N. brevituba* H. Perrier (gallery forest species), and
139 (iv) those from Itremo-Ibity to *N. lowryi* (savanna shrub).

140

141 ***DNA extraction and microsatellite genotyping***

142 Total genomic DNA was extracted from ~5 mg of dried leaves. Each sample was ground in a
143 2-mL tube containing three tungsten beads with a TissueLyser (Qiagen). The BioSprint 15
144 DNA Plant Kit (Qiagen) was then used to extract DNA, which was eluted in 200 μ L AE
145 buffer. Among the 14 loci previously developed (Salmona, Olofsson et al. 2020), 13 were
146 retained, after Nor-14 was discarded due to amplification failure in initial tests on ten species
147 (i.e. *N. aminae* Hong-Wa, *N. brevituba*, *N. capuronii* Bosser, *N. clarinerva* Hong-Wa,
148 *N. disjuncta*, *N. humblotiana*, *N. intermedia* Hong-Wa, *N. sp1_Ando*, *N. spinifolia*, and
149 *N. verticillata*). This locus also showed a significant excess of homozygosity on *N. lowryi*
150 suggesting the presence of null alleles in this species (Salmona, Olofsson et al. 2020). As
151 developing cross-species loci can be a challenge in some plant groups (e.g. Kpatènon et al.
152 2020), their transferability can be maximized by designing primers from conserved genomic
153 regions (e.g. exons of genes), as it has been done for the 13 selected loci (see Salmona,
154 Olofsson et al. 2020). Nuclear SSRs were amplified by PCR following the methods described
155 in Salmona, Olofsson et al. (2020). In order to check the potential utility of cytoplasmic
156 markers for species identification at the local scale, all individuals collected in Marojejy and
157 Andohahela were also analysed with 15 cpSSR loci (loci cp_5, 11, 12, 21, 22, 24, 27, 29, 33,
158 36, 41, 43, 44, 50, and 51) as described in Besnard et al. (2011). All PCR products were
159 separated by capillary electrophoresis on an ABI Prism 3730 DNA Analyzer (Applied
160 Biosystems), and allele size scoring was done with Geneious v9.0.5 (Kearse et al. 2012).

161

162 ***Cross-amplification of SSR loci and survey of ploidy among Noronhia species***

163 Cross-amplification of the 13 nuclear SSR loci was first tested on 76 *Noronhia* taxa. The
164 maximum number of alleles per locus was recorded to estimate the minimum ploidy of each
165 individual.

166 Loci cross-amplification was then investigated among and within species at four
167 locations (i.e. Marojejy, Andohahela, Ankafobe, and Itremo-Ibity). The eight species analysed
168 at the population level (with at least seven individuals) were used to ascertain cross-species
169 transferability of loci both for PCR amplification and for polymorphism records. These
170 species were *a posteriori* considered as diploid since a maximum of two alleles per locus were
171 recorded on all individuals (see results, below). For these taxa, we individually reported the
172 number of alleles per locus (N_a), and then assessed the overall genetic diversity per locus
173 using allelic richness (A_R ; Hurlbert 1971) with the RStudio package *hierfstat* (R Development
174 Core Team 2014; Goudet et al. 1996), and using the observed (H_o) and expected (H_E)
175 heterozygosities with Genetix v4.05 (Belkhir et al. 2004). For each locus, in order to detect
176 loci with putative null alleles or allelic dropout, departure from Hardy–Weinberg proportion
177 was estimated with a Weir & Cockerham Fixation Index (F_{IS} , Weir and Cockerham 1984),
178 and a χ^2 exact test (hereafter HW) based on 10,000 Monte Carlo Markov Chain (MCMC)
179 batches and iterations per batch using Genepop v4.7.5 (Raymond and Rousset 1995). As the
180 power at estimating species ploidy based on the maximum number of alleles at a locus
181 depends on the number and frequency of alleles found at each locus within a species (Besnard
182 and Baali-Cherif 2009), this has to be taken with caution, especially when only one individual
183 is analysed. To test the power of the method, we estimated the probability P_{3x} that a triploid

184 individual present in a population of a diploid species is revealed with our loci (i.e., at least
185 one locus showing three distinct alleles in its genotype; Besnard and Baali-Cherif 2009).

186

187 ***Genetic structure and species delimitation at the local scale***

188 We then investigated the species discrimination power of SSR data and their potential utility
189 for comparative studies of population genetics. We first analysed cpDNA haplotype
190 relationships between species locally collected at two distant sites (Andohahela and Marojejy)
191 using reduced-median networks constructed with Network v5 (Bandelt, Forster, and Röhl
192 1999). Patterns of genetic variance were then examined with a principal component analysis
193 (PCA) using the package ade4 implemented in RStudio (Dray and Dufour 2007). These
194 analyses were based on nuclear SSR allele occurrences to test the discrimination of species at
195 Andohahela and Marojejy. We further assessed the potential presence of nuclear genetic
196 structure at the local scale using the snapclust clustering approach of the adegenet RStudio
197 package (Tonkin-Hill et al. 2019; Jombart 2008). We used the *K*-means method (for a number
198 of clusters *K* ranging from one to ten at the local scale, or fifteen at a global scale) and
199 selected the most appropriate *K* value using the BIC (Bayesian information criterion) statistics
200 (Schwarz 1978) over 20 iterations per *K* value. This analysis was done after excluding species
201 with few samples (one to five genotypes; see results) that may not bring enough information
202 to identify their own genetic cluster and could also bias clustering inferences of other taxa.

203 To assess the potential effect of distance on gene flow, we finally compared patterns of
204 isolation by distance (IBD) in Marojejy, where *N. verticillata* (33 individuals) and
205 *N. humblotiana* (26) were collected along parallel transects of ca. 3.5 km (Figure S1). The
206 comparison of a riparian and an understory forest species on close parallel transects may
207 reveal contrasting patterns of genetic IBD due to varying contribution of dispersal vectors
208 (e.g. wind and facilitated movement of insect pollinators in riparian habitats vs. limitation of
209 insect movement in the understory forest). To compare IBD patterns of the two species, the
210 pairwise kinship coefficients of Loiselle (Loiselle et al. 1995) and Ritland (Ritland 1996)
211 were estimated at several geographic distance classes (i.e., 0.1; 0.2; 0.3; 0.4; 0.5; 0.6; 0.7; 0.8;
212 0.9; 1; 1.5; 2; 2.5; 3; 3.5 km) in both species with SPAGeDi v1.5 (Hardy and Vekemans
213 2002). *P*-values for 2-sided tests of regression analysis on logarithmic geographic distances
214 for each species were calculated with 10,000 permutations with the same software.

215

216 **Results**

217 ***Nuclear SSR cross-species transferability and ploidy survey among Noronhia***

218 The SSR dataset for cross-species amplification among 76 *Noronhia* accessions is provided in
219 Table S1. Most samples were successfully amplified on the 13 loci, but very few
220 amplification failures were observed for two accessions on Nor-08 (GB_437 and MSV_320),
221 and on Nor-09 (GB_1185 and CHW_579). PCR amplification of these loci was independently
222 repeated three times on these samples to ensure the observation. All individuals analysed
223 showed a maximum of two alleles at each locus (i.e. diploid profile), except three of the four
224 samples of *N. louvelii* (from Analamazaotra and Ambodivohitra) that has three or four alleles
225 on four to 11 loci (i.e. tetraploid profiles; Table S1). On average 7.07 (± 2.19) loci were
226 heterozygous in the accessions from Madagascar. In contrast, the two non-native
227 *N. emarginata* individuals collected at the Reunion Island and Hawaii were both homozygous
228 on all loci, sharing the same genetic profile.

229 SSR polymorphisms were then investigated among and within species at four locations.
230 Microsatellite loci were successfully amplified on all individuals (195 trees; Table S2). First,
231 only the five *N. louvelii* samples from Marojejy again showed more than three alleles on four
232 to ten loci (polyploid profile). Although a maximum of four alleles per locus was scored on
233 four individuals, one sample (MSV_508) showed five alleles of similar intensity on one locus
234 (Nor-10). We then compared genetic diversity patterns among the eight species with a diploid
235 profile and for which at least seven individuals were sampled (Table 1). Allelic polymorphism
236 was revealed on all loci in *N. lowryi* and *N. aff. boivinii*, whereas one to three loci were
237 monomorphic in the six remaining taxa. Nor-03 was the most variable locus in six species,
238 with a maximum of 21 alleles in *N. lowryi* and *N. humblotiana* (Table 1). In contrast, Nor-12
239 and Nor-13 were the least variable loci, each reaching a maximum of four alleles in one
240 species (*N. humblotiana* and *N. aff. boivinii*, respectively). A homozygosity excess coupled
241 with a significant departure from Hardy-Weinberg equilibrium was revealed on Nor-08 in
242 *N. lowryi* and *N. aff. boivinii*, and Nor-02 in *N. aff. boivinii* (Table 1). These observations may
243 result from the presence of null alleles as already observed on Nor-08 with the PCR failure on
244 two species (see above). For Nor-02, the observed allele size range is particularly wide among
245 *Noronhia* species (i.e. 210-288 bp; Tables S1 and S2) and allelic dropout could also explain
246 homozygosity excess. These loci should thus be used with caution because frequent null
247 alleles may bias the estimation of some parameters (i.e. P_{3x} ; see below). A significant
248 departure from Hardy-Weinberg equilibrium was also observed on Nor-01 in *N. humblotiana*,
249 despite no homozygosity excess, but this should result from a strong structure in the
250 population (e.g. Walhund effect; see below). Finally, allelic richness (computed for seven
251 individuals) ranged from 3.03 to 5.54, respectively in *N. verticillata* and *N. humblotiana*.

252 All individuals, except those assigned to *N. louvelii*, showed a diploid profile (see
253 above), but the maximum number of alleles on a dozen loci does not necessarily reflect the
254 ploidy level. We thus estimated the probability that a triploid genotype would be detected, P_{3x}
255 (i.e. observation of at least a locus with three alleles), within the eight diploid species sampled
256 at the population level. This parameter was first assessed with all loci, but we then excluded
257 loci with putative null alleles in at least one species for a more conservative test (Table 2). P_{3x}
258 computed on the ten most reliable loci was still relatively high (89.2 to 99.5%) in all species,
259 meaning that a triploid individual has a great probability of being detected on at least one
260 locus considering the levels of genetic diversity in the eight studied diploid species.

261

262 ***Potential use of SSRs for comparative population genetics and species identification***

263 We then investigated genetic differentiation between *Noronhia* species at two localities:
264 Andohahela and Marojejy. Plastid SSR data support taxa distinction with very few
265 intraspecific polymorphisms (Figure 2; Table S3). PCAs based on nuclear SSR alleles also
266 show that all species with at least five individuals were clearly distinguished from each other
267 at both sites (Figures 3 and 4). In addition, other species, represented only by either a single
268 individual or a pair, were also discriminated from the most frequently collected taxa on the
269 two first axes of the PCAs, supporting their status as distinct species. The analysis of one or
270 two individuals per species is, however, not enough to reveal intra- and inter-specific
271 variability that may allow distinguishing them in a multivariate analysis.

272 Then, we excluded taxa with less than seven samples for which a Bayesian clustering
273 analysis on nuclear SSR data could not be performed. This approach revealed a distinct
274 genetic cluster for each of the three species collected at Andohahela (Figure 3C) and Marojejy
275 (Figure 4D). All individuals were assigned to a specific cluster with a probability equal to 1

276 (the probability that an individual belongs to another cluster was $< 10^{-11}$ in Marojejy and
277 $< 10^{-15}$ in Andohahela). No evidence of admixture was thus observed. Similarly, the global
278 clustering analysis of the eight species retained for local-scale analyses (including those from
279 Ankafohe and Itremo-Ibity) also strongly supported them as part of distinct genetic clusters
280 (Figure S2).

281 Patterns of IBD were finally compared between *N. humblotiana* and *N. verticillata*, two
282 species collected in Marojejy on parallel transects along the path close to the river course
283 (Figures 5, S3, S4 and S5). The understory species (*N. humblotiana*) shows a stronger and
284 highly significant IBD pattern (both for Ritland and Loiselle kinship coefficients, p -value
285 $< 10^{-4}$) than the riparian species (*N. verticillata*; p -value = 0.0303 and 0.0475, for Ritland and
286 Loiselle kinship coefficients, respectively).

287

288 Discussion

289 *Cross-species nuclear SSRs revealed a low frequency of polyploids in Noronhia*

290 A relatively high cross-species transferability of our nuclear SSR markers was demonstrated
291 among a representative sample of *Noronhia* (60 of the 82 described species in Madagascar;
292 Hong-Wa 2016). Three loci should, however, be used with caution since PCR failure and/or
293 homozygosity excess were revealed in some species suggesting the presence of null alleles
294 (Nor-08 and Nor-09) or allelic dropout (Nor-02). All loci have also revealed intra-specific
295 polymorphisms in several species demonstrating their potential utility for comparative
296 population genetics among *Noronhia*. Only the two invasive individuals of *N. emarginata*
297 from the Reunion Island and Hawaii were genetically identical and homozygous on all tested
298 microsatellites, suggesting a strong bottleneck associated with their establishment into their
299 new range (i.e., Seychelles, Mascarenes, French Polynesia, Cook Islands, Hawaii, Caribbean,
300 and Bermuda; CABI 2020).

301 At least ten loci revealed polymorphisms within the eight diploid species analysed at the
302 population level. The chance of detecting a non-diploid genotype present in these populations
303 on at least one locus is high as shown by the probability of triploid detection ($P_{3x} > 0.89$). This
304 test is quite conservative given that the probability of detecting a higher level of ploidy
305 (tetraploid and more) is expected to be far greater (but extremely complicated to estimate).
306 The diploid profile of most *Noronhia* accessions (i.e. 257/263) thus supports that polyploidy
307 is not frequent in this genus. However, P_{3x} is not a formal test because it does not allow
308 directly estimating the probability of a polyploid species to be detected based on a single SSR
309 genotype, since this probability will depend on the genetic diversity present within a given
310 taxon (e.g. recent autopolyploids should be more difficult to detect). Non-diploid evidence
311 based on one or two loci should be also avoided due to possible locus duplication events.
312 Consequently, we acknowledge that species ploidy should be determined with high
313 confidence based on the genetic profile of several loci from several individuals of a
314 population (e.g. García-Verdugo et al. 2009; Feng et al. 2018; Hagl et al. 2021). Our data thus
315 allowed us to unambiguously reveal eight polyploid individuals in one taxon (*N. louvelii*)
316 sampled in rainforests of two distant places (i.e. six from Marojejy-Ambodivohitra, and two
317 from Analamazaotra). The diploid profile of one *N. louvelii* individual (RZC_157) from
318 another location, however, calls for additional studies on a more comprehensive sample to test
319 that this taxonomical denomination, which shows a wide morphological variation in
320 herbarium collections, could include several cryptic species with different ploidy levels. Non-
321 diploid individuals of *N. louvelii* are probably tetraploid although one of them showed five

322 alleles at one locus. Abnormal chromosome segregation in polyploids is frequent and could
323 also explain irregularities linked to aneuploid formation (as already observed, for instance, in
324 progenies of the hexaploid ‘Moroccan olive’; Besnard et al. 2008). Although our results
325 confirm the existence of Malagasy olive polyploids, the relatively low frequency of non-
326 diploid profiles suggests that whole-genome duplication was not a major process involved in
327 *Noronhia* diversification, as previously hypothesized (Hong-Wa and Besnard 2014).
328 Supporting this statement, further ongoing population analyses in northern Madagascar (i.e.,
329 Loky-Manambato and Montagne d’Ambre) following the same approach on more than 700
330 individuals belonging to more than 30 *Noronhia* species have not revealed any polyploid taxa
331 yet (G. Besnard, unpublished data). Ploidy verification of *Noronhia* individuals and taxa
332 should however remain a prerequisite step when developing population genomics, which aims
333 to infer their past demographic history or document hybridization patterns.

334

335 ***Comparative population genetics of co-occurring Noronhia species***

336 Diploid species collected at the population level were then clearly distinguished with
337 multivariate and Bayesian clustering analyses of nuclear SSR data. In addition, at the two
338 locations where multiple *Noronhia* species were collected (Marojejy and Andohahela), both
339 nuclear and plastid microsatellite data confirmed taxa delineation, thus proving their potential
340 utility in assisting *Noronhia* identification. The high transferability of markers also allowed
341 comparative investigation on the genetic diversity among species (Table 1), as well as
342 comparison of their genetic structure (as shown, for instance, in related species of
343 Brassicaceae; Clauss, Cobban, Mitchell-Olds 2012). To illustrate this purpose, we showed
344 that two co-occurring *Noronhia* species collected on a rainforest transect harbour contrasted
345 patterns of genetic diversity (Table 1). In addition, a highly significant IBD was revealed in
346 the understory species (*N. humblotiana*) in contrast to the riparian species (*N. verticillata*),
347 which has a rather dendritic distribution inside the forest along the edges of rivers. This
348 observation suggests that gene flow in *N. humblotiana* may be reduced by the forest cover,
349 probably as a consequence of a limited movement of insect pollinators and reduced wind flow
350 in the understory. Conversely, more wind and facilitated movement of insect pollinators in the
351 open riparian habitat as well as seed dispersal via the river may allow gene flow to a longer
352 distance inside the forest and reduce the IBD pattern in *N. verticillata*. This also means that
353 connectivity between trees within and among *Noronhia* populations should depend on
354 numerous factors (i.e., population density, habitat opening, pollination and fruit dispersal
355 vectors; Salmona, Dresen et al. 2020; Salmona, Olofsson et al. 2020) and may considerably
356 affect genetic diversity patterns that could be contrasted between congeneric species, even
357 from very close habitats. Such knowledge may also be highly relevant when defining
358 conservation practices of protected forests. Finally, strong IBD could reduce the power of
359 markers at delimitating taxa when comparing individuals of the same species from distant
360 locations, thus calling for additional studies.

361

362 **Conclusions and perspectives**

363 Our study shows how microsatellite markers can be used in comparative population genetics
364 studies of *Noronhia* spp. Genetic characterization of dozens of species occurring in two
365 hotspots of the genus (i.e., Loky-Manambato and Montagne d’Ambre) is in progress and will
366 help us to determine their ploidy as well as assist their recognition and compare their
367 diversity. The possibility that species complexes are present in these areas deserves special

368 attention, and such preliminary analyses may be essential to identify them and define the most
369 appropriate sampling for tackling questions on their reproductive isolation with genomic data.

370 Finally, the extremely low heterozygosity revealed on *N. emarginata* is particularly
371 unusual in *Noronhia*. As this taxon is quite easy to sample in its invasive range and its floral-
372 to-fruit development biology relatively well described (Rohwer 1995), as opposed to all other
373 species of the genus, we consider it as a choice model for developing *Noronhia* genomics.
374 Therefore, the two accessions from the Reunion Island and Hawaii are currently under
375 sequencing to assemble a reference genome of the genus that could be used for further
376 population genomics of any *Noronhia* species.

377

378 **Acknowledgments**

379 We thank the Direction Générale du Ministère de l'Environnement et des Forêts de
380 Madagascar, Madagascar's Ad Hoc Committee for Fauna and Flora, and Organizational
381 Committee for Environmental Research (CAFF/CORE) for permission to perform this study
382 (permit number: [232/11]/MEF/SG/DGF/DCB.SAP/SCB), and for their support. We thank
383 the local communities of Andohahela, Ankafobe and Marojejy for their warm reception and
384 support. We are grateful to the Get-Plage sequencing platform for genotyping services. We
385 thank particularly S. Fourre, M. A. Naranjo Arcos, A. Staquet, U. Suescun, C. Verbeke, and
386 C. Vignes for lab assistance, as well as the KMCC team, J. Razanatsoa, Anicet, Barbara, Bat,
387 Jordi, and Mino for their enthusiasm during field trips. We also thank J. Corrège and B.
388 Fougerson for their support in informatics. Lab work was funded through an ERA-NET
389 BiodivERsA project: INFRAGECO (Inference, Fragmentation, Genomics, and Conservation,
390 ANR-16-EBI3-0014). We also thank the LABEX TULIP (ANR-10-LABX-0041) and CEBA
391 (ANR-10-LABX-25-01).

392

393 **Notes on contributors**

394 *Océane Gorrilliot* obtained an MSc degree in Ecology and Evolution at the Toulouse
395 University (Paul Sabatier) in 2020. Contribution: Performed data analyses and co-wrote the
396 paper.

397 *Cynthia Hong-Wa* is a curator at Claude E. Phillips Herbarium. Contribution: Secured and
398 identified plant material, and edited the final text.

399 *Franck Rakotonasolo* is a curator at the Parc Botanique & Zoologique de Tsimbazaza and the
400 Kew Madagascar Conservation Centre. Contribution: Secured plant material and obtained
401 research permit.

402 *Guillaume Besnard* is a researcher at the French Centre National de la Recherche
403 Scientifique. He focuses on the biogeography of various plant and animal groups.
404 Contribution: Conceived the study, performed lab work and data analyses, and wrote the
405 paper with the help of all authors.

406

407

408 **References**

- 409 Bandelt, H. J., P. Forster, and A. Röhl. 1999. "Median-joining networks for inferring intraspecific phylogenies."
410 *Molecular Biology and Evolution* 16: 37–48.
- 411 Belkhir, K., P. Borsa, L. Chikhi, N. Raufaste, and F. Bonhomme. 2004. "Genetix 4.05, logiciel sous Windows
412 TM pour la génétique des populations." Laboratoire Genome, Populations, Interactions, CNRS UMR 5000,
413 Université de Montpellier II, Montpellier, France. URL:<http://www.genetix.univmontp2.fr/genetix/intro.htm>
- 414 Besnard, G., C. García-Verdugo, R. Rubio de Casas, U. A. Treier, N. Galland, and P. Vargas. 2008. "Polyploidy
415 in the olive complex (*Olea europaea* L.): evidence from flow cytometry and nuclear microsatellite analyses."
416 *Annals of Botany* 101: 25–30.
- 417 Besnard, G., and D. Baali-Cherif. 2009. "Coexistence of diploids and triploids in a Saharan relict olive: evidence
418 from nuclear microsatellite and flow cytometry analyses." *Comptes Rendus – Biologies* 332: 1115–1120.
- 419 Besnard, G., P. Hernández, B. Khadari, G. Dorado, and V. Savolainen. 2011. "Genomic profiling of plastid
420 DNA variation in the Mediterranean olive tree." *BMC Plant Biology* 11: 80.
- 421 CABI. 2020. "*Noronhia emarginata* (Madagascar olive)." In *Invasive Species Compendium*. Wallingford, UK:
422 CAB International. [<https://www.cabi.org/isc/datasheet/19708471>]
- 423 Caetano Wyler, S., and Y. Naciri. 2016. "Evolutionary histories determine DNA barcoding success in vascular
424 plants: seven case studies using intraspecific broad sampling of closely related species." *BMC Evolutionary
425 Biology* 16: 103.
- 426 Callmänder, M. W., P. B. Phillipson, G. E. Schatz, S. Andriambololonera, M. Rabarimanarivo, N. Rakotonirina,
427 J. Raharimampionona, C. Chatelain, L. Gautier, and P. P. II Lowry. 2011. "The endemic and non-endemic
428 vascular flora of Madagascar updated." *Plant Ecology and Evolution* 144: 121–125.
- 429 Clauss, M. J., H. Cobban, and T. Mitchell-Olds. 2002. "Cross-species microsatellite markers for elucidating
430 population genetic structure in *Arabidopsis* and *Arabis* (Brassicaceae)." *Molecular Ecology* 11: 591–601.
- 431 Díaz Martínez, S., C. Boedeker, and G. C. Zuccarello. 2020. "Microsatellite design for species delimitation and
432 insights into ploidy for the Lake Baikal Cladophoraceae species flock." *Phycologia* 59: 355–364.
- 433 Dray, S., and A. Dufour. 2007. "The ade4 package: implementing the duality diagram for ecologists." *Journal of
434 Statistical Software* 22: 1–20.
- 435 Duminil, J., and M. Di Michele. 2009. "Plant species delimitation: A comparison of morphological and
436 molecular markers." *Plant Biosystems* 143: 528–542.
- 437 Feng, B., S. V. Yi, M. Zhang, and X. Zhou. 2018. "Development of novel EST-SSR markers for ploidy
438 identification based on de novo transcriptome assembly for *Misgurnus anguillicaudatus*." *PLoS ONE* 13:
439 e0195829.
- 440 Freudenstein, J. V., M. B. Broe, R. A. Folk, and B. T. Sinn. 2017. "Biodiversity and the species concept:
441 Lineages are not enough." *Systematic Biology* 66: 644–656.
- 442 Ganzhorn, J. U., P. P. II Lowry, G. E. Schatz, and S. Sommer. 2008. "The biodiversity of Madagascar: one of the
443 world's hottest hotspots on its way out." *Oryx* 35: 346–348.
- 444 García-Verdugo, C., M. F. Fay, C. Granado-Yela, R. Rubio de Casas, L. Balaguer, G. Besnard, and P. Vargas.
445 2009. "Genetic diversity and differentiation processes in the ploidy series of *Olea europaea*: a multiscale
446 approach from subspecies to insular populations." *Molecular Ecology* 18: 454–467.
- 447 Goudet, J., M. Raymond, T. de Meeüs, and F. Rousset. 1996. "Testing differentiation in diploid populations."
448 *Genetics* 144: 1933–1940.
- 449 Goyder, D., W. J. Baker, G. Besnard, J. Dransfield, L. Gardiner, J. Moat, D. Rabehevitra, L. Rajaovelona, S.
450 Rakotoarisoa, F. Rakotonasolo, H. Ralimanana, T. Randriamboavonjy, J. Razanatsoa, V. Sarasan, M. S.
451 Vorontsova, P. Wilkin, and S. Cable. 2017. "Country focus - status of knowledge of Madagascan plants." In
452 *State of the World's Plants 2017*, edited by K. J. Willis, Chap. 6, 36–41. Kew: Royal Botanic Gardens.
- 453 Hagl, A., R. Gargiulo, M. F. Fay, C. Solofondranohatra, J. Salmona, U. Suescun, N. Rakotomalala, C. E. R.
454 Lehmann, G. Besnard, A. S. T. Papadopoulos, and M. S. Vorontsova. 2021. "Geographical structure of
455 genetic diversity in *Loudezia simplex* (Poaceae) in Madagascar and South Africa." *Botanical Journal of the
456 Linnean Society*, doi: 10.1093/botlinnean/boaa98.
- 457 Hardy, O. J., and X. Vekemans. 2002. "SPAGeDi: a versatile computer program to analyse spatial genetic
458 structure at the individual or population levels." *Molecular Ecology Notes* 2: 618–620.
- 459 Hong-Wa, C. 2016. "A taxonomic revision of the genus *Noronhia* Stadtm. ex Thouars (Oleaceae) in Madagascar
460 and the Comoro Islands." In *Boissiera*, edited by M. W. Callmänder and P. P. II Lowry, Vol. 70. Geneva:
461 Conservatoire et Jardin Botanique.

- 462 Hong-Wa, C., and G. Besnard. 2013. "Intricate patterns of relationships in the olive family as inferred from
463 multi-locus plastid and nuclear DNA sequence analyses: a close-up on *Chionanthus* and *Noronhia*
464 (Oleaceae)." *Molecular Phylogenetics and Evolution* 67: 367–378.
- 465 Hong-Wa, C., and G. Besnard. 2014. "Diversification and species limits in the Madagascar olive." *Botanical*
466 *Journal of the Linnean Society* 174: 141–161.
- 467 Hurlbert, S. H. 1971. "The nonconcept of species diversity: a critique and alternative parameters." *Ecology* 52:
468 577–586.
- 469 Husband, B.C., S. J. Baldwin, and J. Suda. 2013. "The incidence of polyploidy in natural plant populations:
470 Major patterns and evolutionary processes." In *Plant Genome Diversity*, edited by J. Greilhuber, J. Doležal,
471 and J. F. Wendel, Vol. 2, 255–276. Vienna: Springer.
- 472 Jombart, T. 2008. "adeigenet: a R package for the multivariate analysis of genetic markers." *Bioinformatics* 24:
473 1403–1405.
- 474 Kpatènou, M. J., K. V. Salako, S. Santoni, L. Zekraoui, M. Latreille, C. Tollon-Cordet, C. Mariac, E. Jaligot, T.
475 Beulé, and K. Adéoti. 2020. "Transferability, development of simple sequence repeat (SSR) markers and
476 application to the analysis of genetic diversity and population structure of the African fan palm (*Borassus*
477 *aethiopum* Mart.) in Benin." *BMC Genetics* 21: 145.
- 478 Kearse, M., R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz,
479 C. Duran, T. Thierer, B. Ashton, P. Meintjes, and A. Drummond. 2012. "Geneious basic: an integrated and
480 extendable desktop software platform for the organization and analysis of sequence data." *Bioinformatics* 28:
481 1647–1649.
- 482 Liang, T, W. Sun, H. Ren, I. Ahmad, N. Vu, Maryam, and J. Huang. 2019. "Genetic diversity of *Ziziphus*
483 *mauritiana* germplasm based on SSR markers and ploidy level estimation." *Planta* 249: 1875–1887.
- 484 Linan, A. G., G. E. Schatz, P. P. II Lowry, A. Miller, and C. E. Edwards. 2019. "Ebony and the Mascarenes: the
485 evolutionary relationships and biogeography of *Diospyros* (Ebenaceae) in the western Indian Ocean."
486 *Botanical Journal of the Linnean Society* 190: 359–373.
- 487 Loiselle, B. A., V. A. Sork, J. Nason, and C. Graham. 1995. "Spatial genetic structure of a tropical understory
488 shrub, *Psychotria officinalis* (Rubiaceae)." *American Journal of Botany* 82: 1420–1425.
- 489 McKendrick, L., J. Provan, Ú. Fitzpatrick, M. J. F. Brown, T. E. Murray, E. Stolle, and R. J. Paxton. 2017.
490 "Microsatellite analysis supports the existence of three cryptic species within the bumble bee *Bombus*
491 *lucorum* sensu lato." *Conservation Genetics* 18: 573–584.
- 492 Maduna, S. N., C. Rossouw, C. da Silva, M. Soekoe, and Z. E. Bester-van der Merwe. 2017. "Species
493 identification and comparative population genetics of four coastal houndsharks based on novel NGS-mined
494 microsatellites." *Ecology and Evolution* 7: 1462–1486.
- 495 Perrier de la Bâthie H. 1949. "Révision des Oléacées de Madagascar et des Comores." *Mémoires de l'Institut*
496 *Scientifique de Madagascar, Série B, Biologie Végétale* 2: 275–310.
- 497 Perrier de la Bâthie H. 1952. "Oléacées. 166ème famille." In *Flore de Madagascar et des Comores*, edited by H.
498 Humbert. Paris: Muséum National d'Histoire Naturelle.
- 499 R Development Core Team. 2014. "R: A language and environment for statistical computing." Vienna: R
500 Foundation for Statistical Computing.
- 501 Raymond, M., and F. Rousset. 1995. "Genepop (version 1.2): population genetics software for exact tests and
502 ecumenicism." *Journal of Heredity* 86: 248–249.
- 503 Ritland, K. 1996. "Estimators for pairwise relatedness and individual inbreeding coefficients." *Genetical*
504 *Research* 67: 175–185.
- 505 Rohwer, J. G. 1995. "From flower to fruit in *Noronhia emarginata* (Lam.) Thouars (Oleaceae)." *Flora* 190: 35–
506 43.
- 507 Salmona, J., J. K. Olofsson, C. Hong-Wa, J. Razanatosoa, F. Rakotonasolo, H. Ralimanana, T.
508 Randriamboavonjy, U. Suescun, M. S. Vorontsova, and G. Besnard. 2020. "Late Miocene origin and recent
509 population collapse of the savanna Malagasy olive (*Noronhia lowryi*)." *Biological Journal of the Linnean*
510 *Society* 129: 227–243.
- 511 Salmona, J., A. Dresen, A. E. Ranaivoson, S. Manzi, B. Le Pors, C. Hong-Wa, J. Razanatosoa, N. V.
512 Andriaholinirina, S. Rasoloharijaona, M. E. Vavistara, and G. Besnard. 2020. "How ancient forest
513 fragmentation and riparian connectivity generate high levels of genetic diversity in a micro-endemic
514 Malagasy tree." *bioRxiv*, doi:10.1101/2020.11.25.394544.
- 515 Schwarz, G. E. 1978. "Estimating the dimension of a model." *Annals of Statistics* 6: 461–464.

- 516 Singh H. K., I. Parveen, S. Raghuvanshi, and S. B. Babbar. 2012. “The loci recommended as universal barcodes
517 for plants on the basis of floristic studies may not work with congeneric species as exemplified by DNA
518 barcoding of *Dendrobium* species.” *BMC Research Notes* 5: 42.
- 519 Skema, C. 2012. “Toward a new circumscription of *Dombeya* (Malvales: Dombeyaceae): A molecular
520 phylogenetic and morphological study of *Dombeya* of Madagascar and a new segregate genus, *Andringitra*.”
521 *Taxon* 61: 612–628.
- 522 Tonkin-Hill, G., J. A. Lees, S. D. Bentley, S. D. W. Frost, and J. Corander. 2019. “Fast hierarchical Bayesian
523 analysis of population structure.” *Nucleic Acids Research* 47: 5539–5549.
- 524 Viruel, J., M. Conejero, O. Hidalgo, L. Pokorny, R. F. Powell, F. Forest, M. B. Kantar, M. Soto Gomez, S. W.
525 Graham, B. Gravendeel, P. Wilkin, and I. J. Leitch. 2019. “A target capture-based method to estimate ploidy
526 from herbarium specimens.” *Frontiers in Plant Science* 10: 937.
- 527 Weir, B., and C. Cockerham. 1984. “Estimating *F*-statistics for the analysis of population structure.” *Evolution*
528 38: 1358–1370.
- 529 Wang, H., J. Yang, L. Boykin, Q. Y. Zhao, Y. J. Wang, S. S. Liu, and X. W. Wang. 2014. “Developing
530 conversed microsatellite markers and their implications in evolutionary analysis of the *Bemisia tabaci*
531 complex.” *Scientific Reports* 4: 6351.
- 532 Willis, K. J., L. Gillson, and S. Knapp. 2007. “Biodiversity hotspots through time: an introduction.”
533 *Philosophical Transactions of the Royal Society B: Biological Sciences* 362: 169–174.

534

535 **Data availability statement**

536 All samples information and genotyping data are provided in Supplemental Data.

537

538 **Supplemental data**

539 Supplemental Data may be found in the online version of this article at the publisher’s web-
540 site:

541 **Table S1.** Cross-species transferability of 13 nuclear SSR loci among 76 *Noronhia* accessions

542 **Table S2.** Full dataset of nuclear SSR genotypes on 266 individuals (for 13 loci)

543 **Table S3.** Plastid SSR (cpSSR) profiles of 112 individuals collected in Marojejy and
544 Andohahela NPs

545 **Figure S1.** Geographic distribution of sampled species and individuals at the two sites of (A)
546 Marojejy and (B) Andohahela

547 **Figure S2.** Inference of genetic structure among eight *Noronhia* species (with a minimum of
548 seven genotypes) from four distant locations

549 **Figure S3.** Pairwise kinship relationships according to the geographic distance in
550 *N. humblotiana*

551 **Figure S4.** Pairwise kinship relationships according to the geographic distance in
552 *N. verticillata*

553 **Figure S5.** Comparison of IBD patterns between two rainforest *Noronhia* species collected on
554 the same transect in the Marojejy NP (based on the Loiselle kinship coefficient)

Table 1. Summary of genetic diversity parameters among eight diploid *Noronhia* species assessed on 13 SSR loci. The number of alleles (N_a), allelic richness (A_R ; computed for seven individuals), observed and expected heterozygosities (H_O and H_E , respectively), F_{IS} fixation index (Weir and Cockerham 1984), and departure from Hardy–Weinberg proportion (HW; χ^2 p -value) are given. The number of characterized individuals is indicated in parentheses under each species name. As multiple tests were done for each species (i.e. between 10 and 13 polymorphic loci), departure from Hardy–Weinberg was considered significant when HW was inferior to 0.005 (in bold).

Species	Locus	N_a	A_R	H_O	H_E	F_{IS}	HW	Species	Locus	N_a	A_R	H_O	H_E	F_{IS}	HW		
<i>N. lowryi</i> (73)	Nor-01	10	5.00	0.685	0.773	0.115	0.096	<i>N. verticillata</i> (33)	Nor-01	6	3.41	0.455	0.586	0.226	0.127		
	Nor-02	12	5.69	0.658	0.795	0.174	0.044		Nor-02	5	4.26	0.788	0.742	-0.063	0.908		
	Nor-03	21	8.89	0.877	0.916	0.043	0.548		Nor-03	14	8.26	0.970	0.907	-0.071	0.395		
	Nor-05	10	5.71	0.753	0.775	0.027	0.419		Nor-05	6	5.14	0.636	0.785	0.192	0.264		
	Nor-06	3	2.26	0.233	0.237	0.015	0.420		Nor-06	3	2.37	0.273	0.432	0.373	0.040		
	Nor-07	18	8.04	0.877	0.895	0.021	0.156		Nor-07	6	3.63	0.636	0.599	-0.064	0.638		
	Nor-08	11	5.69	0.630	0.803	0.216	<10⁻⁴		Nor-08	1	1.00	0.000	0.000	-	-		
	Nor-09	19	7.85	0.795	0.877	0.095	0.010		Nor-09	1	1.00	0.000	0.000	-	-		
	Nor-10	5	2.65	0.301	0.294	-0.026	0.338		Nor-10	2	1.98	0.303	0.339	0.109	0.606		
	Nor-11	4	1.78	0.123	0.119	-0.037	1.000		Nor-11	2	2.00	0.515	0.508	-0.015	1.000		
	Nor-12	3	2.25	0.438	0.430	-0.020	1.000		Nor-12	1	1.00	0.000	0.000	-	-		
	Nor-13	3	2.50	0.411	0.371	-0.108	0.827		Nor-13	3	2.38	0.515	0.470	-0.098	0.713		
	Nor-15	8	5.07	0.808	0.781	-0.035	0.626		Nor-15	3	2.92	0.546	0.604	0.098	0.388		
	Mean		9.77	4.88	0.584	0.620	0.037		0.422	Mean		4.08	3.03	0.434	0.459	0.057	0.508
	<i>N. humblotiana</i> (26)	Nor-01	5	3.73	0.731	0.714	-0.024		0.001	<i>N. disjuncta</i> (8)	Nor-01	6	5.87	0.625	0.867	0.293	0.005
Nor-02		18	8.95	0.885	0.919	0.039	0.040	Nor-02	6		5.50	0.500	0.675	0.273	0.306		
Nor-03		21	9.73	0.923	0.928	0.006	0.652	Nor-03	11		9.98	1.000	0.933	-0.077	0.564		
Nor-05		13	7.17	0.846	0.812	-0.043	0.775	Nor-05	5		4.83	0.875	0.808	-0.089	0.949		
Nor-06		14	8.56	0.962	0.912	-0.056	0.651	Nor-06	4		3.86	0.625	0.525	-0.207	1.000		
Nor-07		7	4.51	0.654	0.560	-0.172	0.964	Nor-07	6		5.74	0.750	0.817	0.087	0.574		
Nor-08		7	3.70	0.346	0.631	0.457	0.005	Nor-08	9		8.36	0.750	0.908	0.185	0.054		
Nor-09		11	7.80	0.962	0.896	-0.075	0.543	Nor-09	3		2.88	0.500	0.592	0.164	1.000		
Nor-10		12	7.34	0.808	0.875	0.078	0.625	Nor-10	3		2.88	0.750	0.575	-0.333	0.273		
Nor-11		1	1.00	0.000	0.000	-	-	Nor-11	3		2.87	0.250	0.342	0.282	0.200		
Nor-12		4	3.08	0.615	0.604	-0.019	0.673	Nor-12	1		1.00	0.000	0.000	-	-		
Nor-13		2	1.98	0.269	0.340	0.212	0.288	Nor-13	1		1.00	0.000	0.000	-	-		
Nor-15		8	4.42	0.615	0.556	-0.110	0.667	Nor-15	8		7.37	0.625	0.842	0.271	0.105		
Mean			9.46	5.54	0.663	0.673	0.024	0.490	Mean			5.08	4.79	0.558	0.606	0.077	0.457

Table 1, end.

Species	Locus	N_a	A_R	H_O	H_E	F_{IS}	HW	Species	Locus	N_a	A_R	H_O	H_E	F_{IS}	HW		
<i>N. brevituba</i> (10)	Nor-01	3	3.00	0.600	0.700	0.150	0.758	<i>N. sp1_Ando</i> (8)	Nor-01	7	6.63	0.875	0.875	0.000	0.621		
	Nor-02	6	5.54	0.900	0.826	-0.095	0.893		Nor-02	5	4.86	0.875	0.775	-0.140	0.652		
	Nor-03	7	6.28	0.700	0.853	0.187	0.314		Nor-03	5	4.63	0.625	0.650	0.041	1.000		
	Nor-05	6	5.24	0.800	0.747	-0.075	0.978		Nor-05	6	5.73	0.750	0.817	0.087	0.943		
	Nor-06	4	3.68	0.700	0.668	-0.050	0.576		Nor-06	6	5.62	0.750	0.767	0.023	0.504		
	Nor-07	4	3.40	0.300	0.537	0.455	0.108		Nor-07	7	6.62	0.625	0.867	0.293	0.238		
	Nor-08	4	3.99	1.000	0.774	-0.314	0.706		Nor-08	4	3.88	0.875	0.750	-0.181	0.680		
	Nor-09	3	2.98	0.800	0.584	-0.398	0.333		Nor-09	5	4.63	0.750	0.683	-0.105	1.000		
	Nor-10	3	2.70	0.400	0.563	0.301	0.307		Nor-10	5	4.87	1.000	0.800	-0.273	1.000		
	Nor-11	1	1.00	0.000	0.000	-	-		Nor-11	4	3.87	0.500	0.675	0.273	0.478		
	Nor-12	2	2.00	0.400	0.337	-0.200	1.000		Nor-12	2	2.00	0.375	0.458	0.192	1.000		
	Nor-13	3	2.70	0.300	0.563	0.481	0.173		Nor-13	2	1.88	0.125	0.125	0.000	1.000		
	Nor-15	4	3.90	0.600	0.721	0.176	0.709		Nor-15	2	1.99	0.250	0.233	-0.077	1.000		
	Mean		3.84	3.57	0.577	0.606	0.051		0.571	Mean		4.62	4.40	0.644	0.652	0.011	0.760
	<i>N. aff. boivinii</i> (18)	Nor-01	7	5.78	0.833	0.830	-0.004		0.439	<i>N. grandifolia</i> (7)	Nor-01	3	3.00	0.143	0.539	0.750	0.021
Nor-02		11	8.30	0.611	0.911	0.336	0.001	Nor-02	4		4.00	0.571	0.714	0.213	0.734		
Nor-03		17	9.53	0.833	0.925	0.102	0.493	Nor-03	5		5.00	0.286	0.824	0.671	0.007		
Nor-05		8	5.26	0.556	0.744	0.259	0.123	Nor-05	4		4.00	0.857	0.736	-0.180	0.712		
Nor-06		8	5.45	0.889	0.749	-0.193	0.750	Nor-06	3		3.00	0.571	0.473	-0.231	1.000		
Nor-07		7	3.82	0.444	0.397	-0.124	1.000	Nor-07	3		3.00	0.286	0.539	0.489	0.105		
Nor-08		9	6.92	0.278	0.870	0.687	<10⁻⁴	Nor-08	1		1.00	0.000	0.000	-	-		
Nor-09		9	6.85	0.944	0.868	-0.091	0.569	Nor-09	7		7.00	0.714	0.890	0.211	0.383		
Nor-10		10	6.98	0.944	0.856	-0.107	0.940	Nor-10	4		4.00	0.857	0.780	-0.108	0.190		
Nor-11		3	1.78	0.111	0.110	-0.015	1.000	Nor-11	7		7.00	1.000	0.846	-0.200	0.958		
Nor-12		3	2.02	0.167	0.160	-0.041	1.000	Nor-12	1		1.00	0.000	0.000	-	-		
Nor-13		4	3.98	0.722	0.770	0.064	0.785	Nor-13	2		2.00	0.143	0.143	-0.000	1.000		
Nor-15		6	4.18	0.500	0.594	0.162	0.166	Nor-15	6		6.00	0.857	0.802	-0.075	0.744		
Mean			7.85	5.45	0.603	0.676	0.080	0.559	Mean			3.85	3.85	0.484	0.560	0.140	0.485

Table 2. Probability of detecting a triploid individual in eight diploid species by using individual SSR locus (p_{3x}) or their combination (P_{3x} ; Besnard and Baali-Cherif 2009). The number of characterized individuals is indicated in parentheses under each species name. P_{3x} was computed for each species by combining information from the 13 loci or 10 loci without evidence of null allele. Probability per locus: $p_{3x} = 3! \sum p_i p_j p_k$ (with p_i , p_j and p_k , the frequencies of three distinct alleles i , j , and k); Probability combining information from x loci: $P_{3x} = 1 - \prod (1 - p_{3x})$.

Locus	<i>N. lowryi</i> (73)	<i>N. verticillata</i> (33)	<i>N. humblotiana</i> (26)	<i>N. disjuncta</i> (8)	<i>N. brevitiba</i> (10)	<i>N. sp1_Ando</i> (8)	<i>N. aff. boivinii</i> (18)	<i>N. grandifolia</i> (7)
Nor-01	0.432	0.161	0.291	0.513	0.221	0.533	0.509	0.079
<i>Nor-02</i>	<i>0.482</i>	<i>0.360</i>	<i>0.730</i>	<i>0.270</i>	<i>0.461</i>	<i>0.354</i>	<i>0.689</i>	<i>0.245</i>
Nor-03	0.751	0.707	0.757	0.668	0.516	0.217	0.729	0.415
Nor-05	0.462	0.443	0.528	0.413	0.348	0.428	0.376	0.276
Nor-06	0.022	0.033	0.712	0.135	0.216	0.360	0.387	0.066
Nor-07	0.699	0.187	0.209	0.434	0.111	0.519	0.099	0.079
<i>Nor-08</i>	<i>0.494</i>	<i>0.000</i>	<i>0.186</i>	<i>0.609</i>	<i>0.354</i>	<i>0.293</i>	<i>0.594</i>	<i>0.000</i>
<i>Nor-09</i>	<i>0.665</i>	<i>0.000</i>	<i>0.672</i>	<i>0.082</i>	<i>0.135</i>	<i>0.234</i>	<i>0.591</i>	<i>0.547</i>
Nor-10	0.044	0.000	0.625	0.079	0.067	0.390	0.572	0.337
Nor-11	0.007	0.000	0.000	0.038	0.000	0.214	0.004	0.478
Nor-12	0.024	0.000	0.143	0.000	0.000	0.000	0.008	0.000
Nor-13	0.044	0.037	0.000	0.000	0.067	0.000	0.373	0.000
Nor-15	0.445	0.153	0.200	0.492	0.278	0.000	0.204	0.411
P_{3x} (13)	99.90	94.39	99.97	99.45	96.74	98.92	99.95	97.65
P_{3x} (10)*	98.90	91.23	99.52	97.91	89.19	96.91	99.03	93.13

* P_{3x} (10) computed on ten loci after excluding Nor-02, Nor-08 and Nor-09 (in italic), for which PCR failure and/or significant homozygosity excess were observed in at least one species (Table 1; Supporting Information, Table S1).

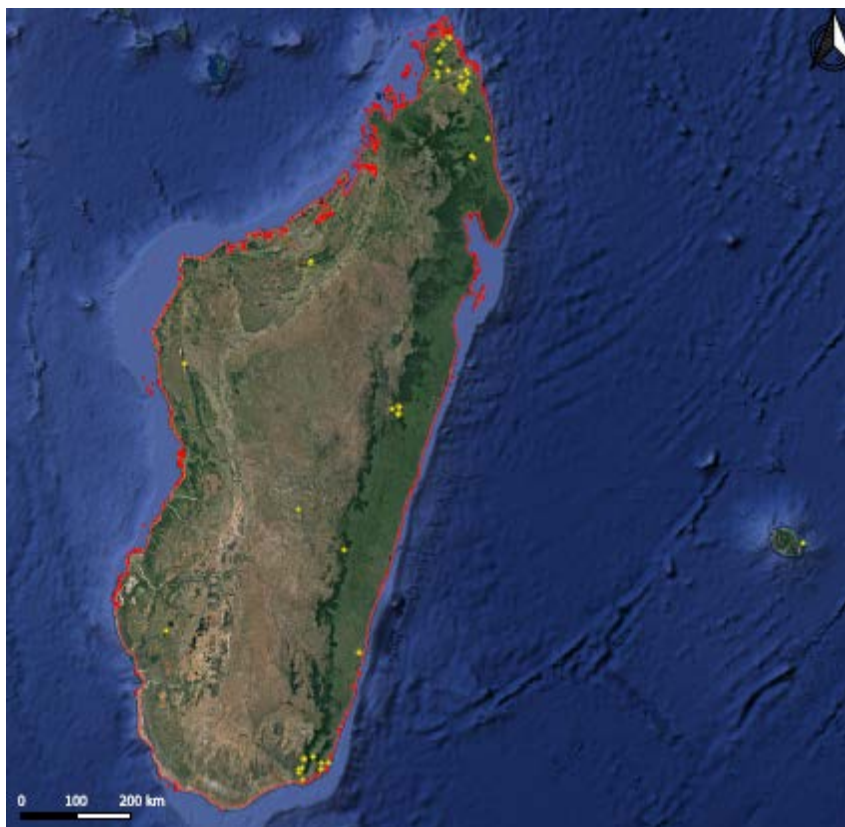


Figure 1. Geographic origin of the 74 samples of the Malagasy floristic region used for cross-species amplification tests of nuclear SSRs. Each yellow dot represents a sample. North and South-East were the most sampled areas and also correspond to regions with the highest species diversity and endemism for *Noronhia* (Hong-Wa 2016). Map was drawn with the help of B. Fougeron.

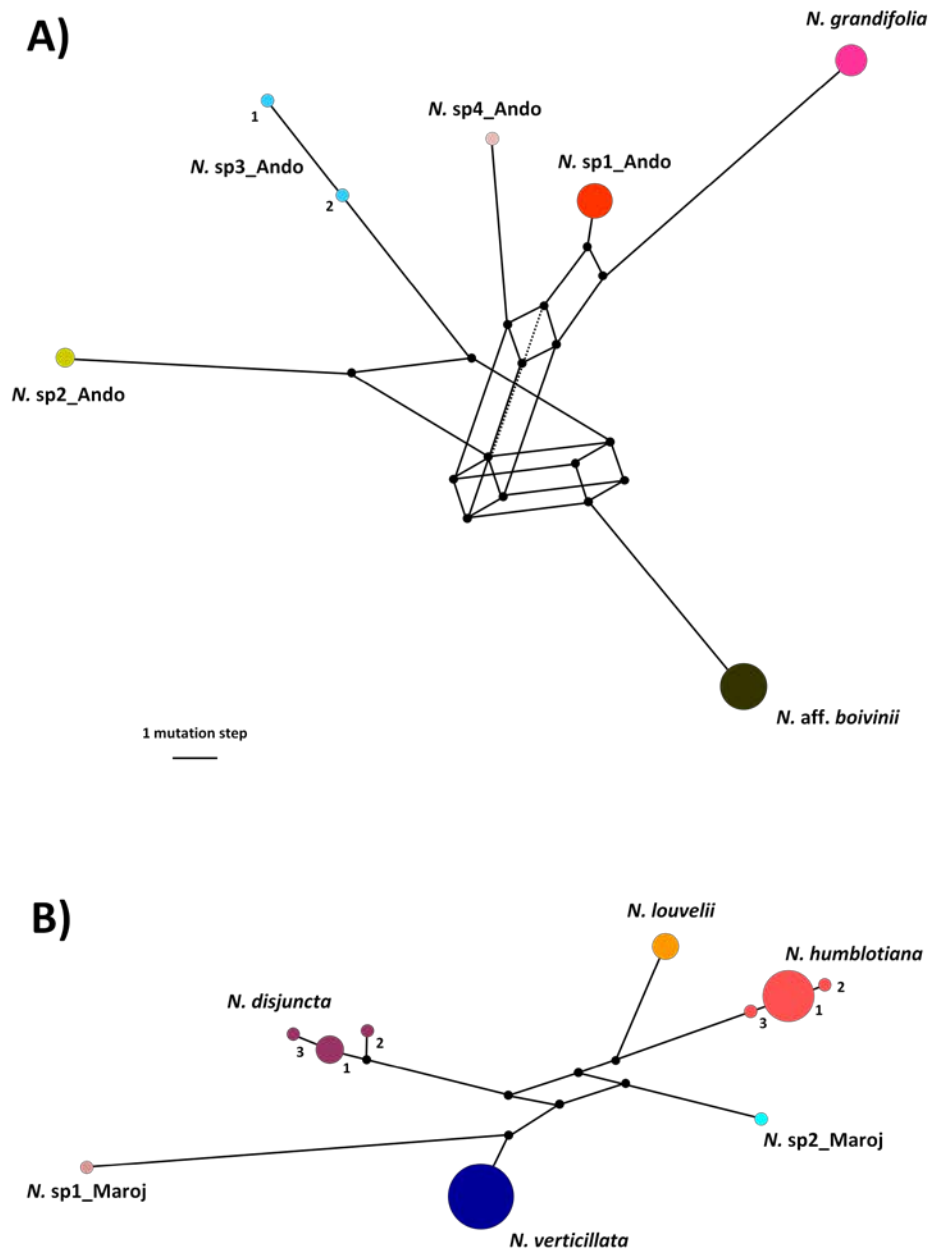


Figure 2. Phylogenetic relationships between the six species collected at Andohahela (A) and the six species collected at Marojejy (B) were analysed with a reduced-median network (Bandelt, Forster, and Röhl 1999) based on cpSSR data (Table S3). At the two sites, these analyses distinguished six plastid lineages (with very few intra-lineage polymorphisms), confirming taxa distinction.

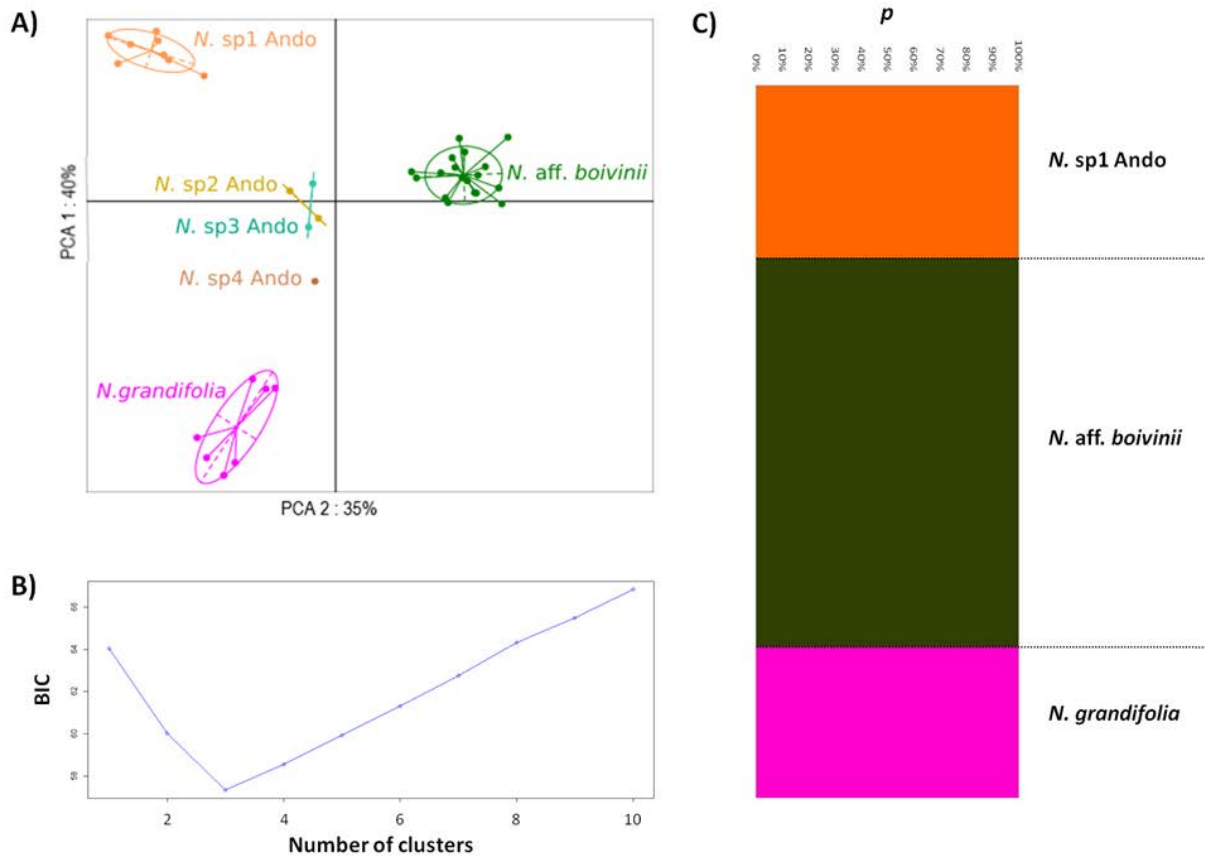


Figure 3. Species discrimination based on nuclear SSR data in Andohahela NP: (A) Principal component analysis (PCA) on all sampled individuals based on SSR allele presence; (B-C) Inference of genetic structure using Bayesian simulations with snapclust on Rstudio (Tonkin-Hill et al. 2019; Jombart 2008) based on SSR genotypes (with at least seven individuals): (B) Bayesian information criterion (BIC; Schwarz 1978) over 20 iterations per K value, and (C) Barplot of the snapclust analysis based on the best K value ($K = 3$) according to BIC (best K is reached when BIC stops to decrease and begins to increase). The percentage (p) of assignment of each individual to the clusters averaged over 20 iterations is shown. Each individual is represented by a horizontal bar.

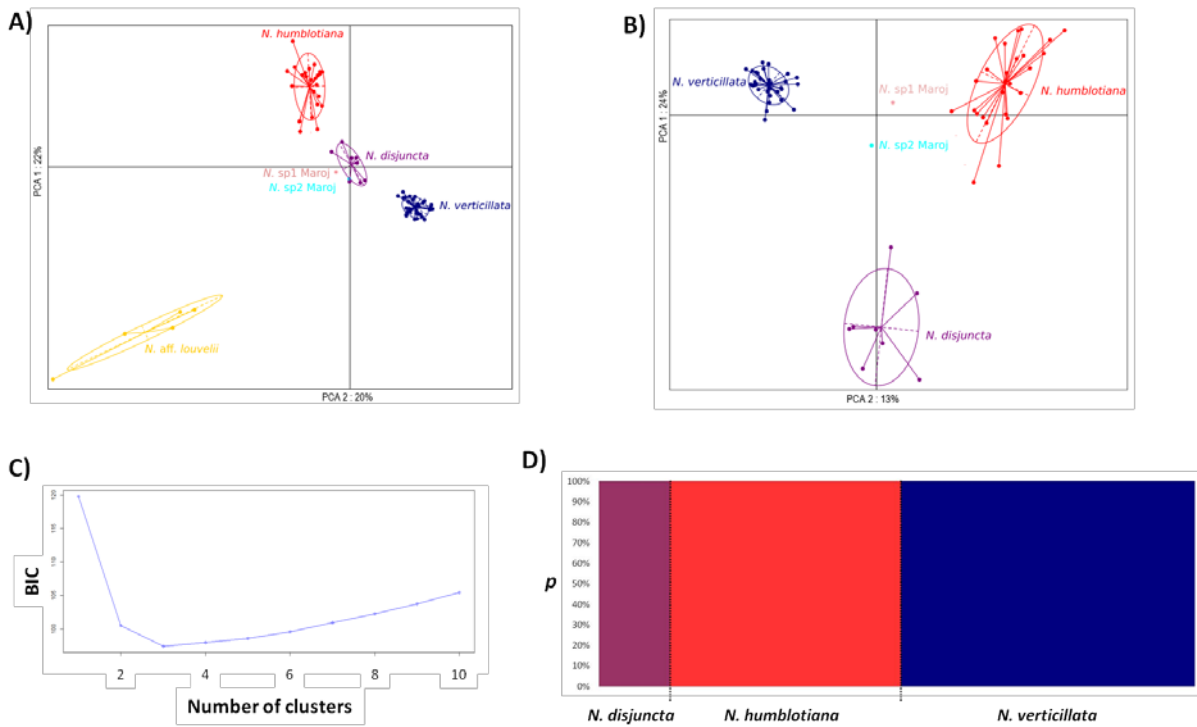


Figure 4. Species discrimination based on nuclear SSR data in Marojejy NP: (A) Principal component analysis (PCA) on all taxa based on SSR allele presence; (B) PCA after excluding *N. louvelii* (five polyploid individuals); (C-D) Inference of genetic structure using Bayesian simulations with snapclust on Rstudio (Tonkin-Hill et al. 2019; Jombart 2008) based on SSR genotypes of three species (with at least eight individuals): (C) Bayesian information criterion (BIC; Schwarz 1978) over 20 iterations per K value, and (D) Barplot of the snapclust analysis based on the best K value ($K = 3$) according to BIC (best K is reached when BIC stops to decrease and begins to increase). The percentage (p) of assignment of each individual to the clusters averaged over 20 iterations is shown. Each individual is represented by a vertical bar.

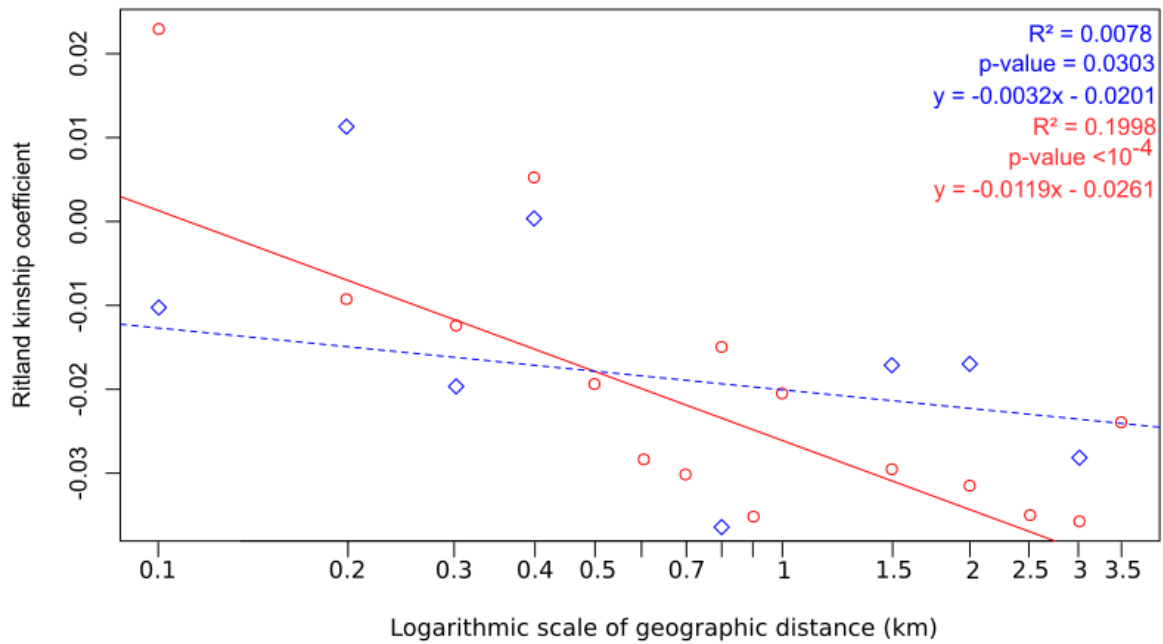


Figure 5. Comparison of IBD patterns between two rainforest *Noronhia* species collected on the same transect in the Marojejy NP. Mean kinship relationship (Ritland 1996) according to the geographic distance in *N. verticillata* (Blue line and diamonds) and *N. humblotiana* (Red line and circles). The R-square and p-value statistics are given in the upper right. For each distance class, a mean pairwise kinship coefficient is given. Individual pairwise kinship estimates are given Figures S3 and S4.