bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

How ancient forest fragmentation and riparian connectivity generate high levels of genetic diversity in a microendemic Malagasy tree

Jordi Salmona[1*], Axel Dresen[1], Anicet E. Ranaivoson[1,2], Sophie Manzi[1], Barbara Le Pors[3], Cynthia Hong-Wa[4], Jacqueline Razanatsoa[5], Nicole V. Andriaholinirina[2], Solofonirina Rasoloharijaona[2], Marie-Elodie Vavitsara[2], Guillaume Besnard[1*]

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

[2] Faculté des Sciences, Université de Mahajanga, BP 652 401, Mahajanga, Madagascar

[3] Instituto Gulbenkian de Ciênca, Rua da Quinta Grande, 6, P-2780-156 Oeiras, Portugal

[4] Claude E. Phillips Herbarium, Delaware State University, 1200 N. Dupont Hwy, Dover, DE 19901-2277, USA

[5] Herbier, Département Flore, Parc Botanique et Zoologique de Tsimbazaza, BP 4096, Antananarivo - 101, Madagascar

*Corresponding authors:

Jordi Salmona: jordi.salmona@gmail.com

Guillaume Besnard: guillaume.besnard@univ-tlse3.fr

Number of figures: 4, all figures should be published in color.

Number of tables: 2

The Supporting Information comprises 10 method and 3 result supporting paragraphs, 10 tables, 24 figures and associated references.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

Understanding landscape changes is central to predicting evolutionary trajectories and defining conservation practices. While human-driven deforestation is intense throughout Madagascar, exceptions in areas like the Loky-Manambato region (North) raise questions. Such regions also harbor a rich and endemic flora, whose evolutionary origin remains poorly understood. We assessed the genetic diversity of an endangered microendemic Malagasy olive species (Noronhia spinifolia Hong-Wa) to better understand the vegetation dynamic in the Loky-Manambato region and its influence on past evolutionary processes. We characterized 72 individuals sampled across eight forests through nuclear and mitochondrial restriction associated sequencing data (RADseq) and chloroplast microsatellites (cpSSR). Combined population and landscape genetics analyses indicate that N. spinifolia diversity is largely explained by the current forest cover, highlighting a long-standing habitat mosaic in the region. This sustains a major and long-term role of riparian corridors in maintaining connectivity across those antique mosaic-habitats, calling for the study of organismal interactions that promote gene flow.

Key words: Habitat loss and fragmentation, habitat mosaic, Landscape genetics, Malagasy olive, Mitochondrial DNA, gene flow, connectivity, cpSSR, RADseq, Madagascar.

Running head:

Ancient mosaic landscape in Madagascar

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Introduction

One way to offset rapid anthropogenic habitat destruction and fragmentation, the primary causes of declines in global biodiversity (Fahrig, 2003; Goudie, 2018; Lindenmayer & Fischer, 2013), is to preserve habitat connectivity (Haddad et al., 2015). Such an approach is, however, insufficient by itself and requires better knowledge of species dispersal strategies (Gardner et al., 2018; Lebuhn et al., 2015; Sutherland et al., 2004), which remain poorly understood, particularly in tropical biodiversity hotspots. This typically requires understanding species diversity, their dynamic, behavior and interactions across rapidly changing landscapes (Pressey et al., 2007). Since the genetic diversity of a species is the result of both biological processes (e.g. mutation, drift, gene-flow), and effects of the environment (e.g. landscape heterogeneity), it retains the signatures of habitat and landscape changes (Balkenhol et al., 2016; Salmona, Heller, Lascoux, et al., 2017). The latter can be efficiently inferred from genetic data although assessing the relative and confounding effects of complex landscape dynamics (habitat loss, fragmentation, barriers emergence, etc.) is notoriously challenging (Beichman et al., 2018; Nater et al., 2015; Salmona, Heller, Lascoux, et al., 2017; Salmona, Heller, Quéméré, et al., 2017). Measuring the effects of landscape changes, structure, and components on genetic diversity of organisms thus remains a major issue in conservation biology.

Madagascar is considered as one of the hottest biodiversity hotspots (Myers et al., 2000). It undergoes high rates of deforestation [~40-50% area since the 1950s], resulting in habitats loss and fragmentation, and ultimately biodiversity loss (Harper et al., 2007; Vieilledent et al., 2018). Anthropogenic deforestations since human settlement on the island have long been assumed to have generated the Malagasy grasslands. However, their Miocene origin (Bond et al., 2008; Hackel et al., 2018; Salmona et al., 2020; Solofondranohatra et al., 2018; Vorontsova et al.,

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

2016) points to the antiquity of Madagascar's habitat mosaics, fueling a hot debate on the causes and age of its habitat fragmentation (Godfrey & Crowley, 2016; Joseph & Seymour, 2020, 2021). Meanwhile, it is a place where connectivity preservation would be a very impactful conservation approach (Morelli et al., 2020). Yet, an understanding of the nature of the fragmentation, whether natural or human-induced and/or ancient or recent, is essential to better assess how species are molded by landscape changes, and to consequently factor their survival strategies in an effective conservation plan.

The Loky-Manambato (LM) region in northern Madagascar, which features perplexingly mild deforestation (Quéméré et al., 2012; Salmona, Heller, Quéméré, et al., 2017), shows well-characterized matrix of forests and open-habitats, a diversity of putative barriers to gene flow, as well as high levels of endemism across living kingdoms (Goodman et al., 2018; Goodman & Wilmé, 2006). It rose as a small-scale model-region to investigate the effects of landscape changes, structure, and components on species' genetic diversity. For instance, the forest-matrix was identified among the major landscape components shaping genetic diversity across all species studied in the LM region (Aleixo-Pais et al., 2019; Quéméré et al., 2010; Rakotoarisoa et al., 2013a; Sgarlata et al., 2018; Tang et al., 2020). Contrastingly the Manankolana River, showed a dominant effect on *Propithecus tattersalli* Simmons, not recovered in other species. Aside mammals, contributions on other taxa, such as plants, are crucial to draw taxonomically-broad generalities regarding the history, connectiveness, and conservation value of the current LM landscape.

By their very nature, native tree species are at the forefront of deforestation and landscape changes. Trees, therefore, represent putatively good models for landscape genetics studies in fragmented habitats, despite their long generation time. However, only a few studies have used

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the genetic diversity of Malagasy plant populations (Andrianoelina et al., 2009; Gardiner et al., 2017; Salmona et al., 2020) to infer landscape dynamics and inform conservation strategies. The Malagasy olives (genus *Noronhia* Stadt. ex Thouars), with a high number of taxa and a high microendemism rate, are among the major components of Madagascar forests and of the LM region in particular (Hong-Wa, 2016; Hong-Wa & Besnard, 2014). Among them, the Malagasy spiny olive (*Noronhia spinifolia* Hong-Wa) is mostly endemic to the dry to sub-humid forests of the LM region and is of high conservation concern due to its narrow range. Nonetheless, *N. spinifolia* is quasi-ubiquitous in the LM region such that its genetic makeup is expected to reflect the signals of the processes operating on its populations across time and space. Moreover, being narrowly distributed, it may hold relatively low genetic diversity (Kimura, 1983) and suffer from inbreeding depression due to recent population collapse. Although its pollen and seed dispersal have yet to be studied, *N. spinifolia*'s flower and fruit morphology suggests insect pollination and animal-mediated dispersal of fruits (see below). *Noronhia spinifolia*, therefore, represents an excellent model to better understand Malagasy olives' ecology and offers a case study to define appropriate action for dry-forests plant conservation in northern Madagascar.

In such sexually-reproducing plants, dispersal occurs by two means: via haploid male gametes in pollen, and via diploid embryos in seeds. Without field data, population and landscape genetics offer an alternative way to estimate effective dispersal (Balkenhol et al., 2016; Holderegger et al., 2010). In particular, the combined use of complementary maternally and biparentally inherited genetic data [respectively from chloroplast or mitochondrial genomes (cpDNA or mtDNA) and the nuclear genome (nDNA)] allows disentangling, to a certain level, the relative contribution of seed and pollen dispersals in gene flow (Cruzan & Hendrickson, 2020). For instance, the congeneric *N. lowryi* Hong-Wa exhibited contrasting near-panmixia nuclear and

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

strong chloroplast genetic structures suggesting a long and short distance dispersal of pollen and seeds, respectively (Salmona et al., 2020). While progresses in sequencing technologies facilitated the generation of such genetic data for non-model organisms (Allendorf et al., 2010), recent advances in spatially explicit analyses also unlocked our ability to estimate the effect of numerous collinear landscape features on genetic diversity (Balkenhol et al., 2016; Prunier et al., 2017). Furthermore, although the limited number of tested alternative landscape hypotheses long relied on prior knowledge or expert opinions, recent approaches iterating around a large panel of resistance values (Graves et al., 2013) or optimizing resistance surfaces (Peterman, 2018; Peterman & Pope, 2021), widened the potential to identifying relevant landscape components while optimizing their cost values from the genetic data itself.

The overall goal of our study is to assess the effects of landscape on *N. spinifolia*'s genetic diversity and population structure to inform conservation strategies. We used genomic data from recently collected specimens of *N. spinifolia* across most of its range, the LM region. We first tested whether its restricted geographic distribution resulted in a low genetic diversity, as expected under a neutral model (Kimura, 1983), or remained relatively high as for co-distributed primates [*P. tattersalli* and *Microcebus tavaratra* Rasoloarison, Goodman & Ganzhorn (Aleixo-Pais et al., 2019; Quéméré et al., 2010)]. We then measured the effect of landscape components on biparentally and maternally inherited genetic diversity, to investigate patterns of pollen and seed dispersals. We compared its dispersal patterns with those of a congeneric species from the High Plateau [*N. lowryi* (Salmona *et al.*, 2020)] that exhibits near-panmixia on pollen-dispersed genes, but very short seed dispersal, and with those of co-distributed mammal taxa (abovementioned), whose dispersal is mainly limited by non-forest habitats and rivers. The little knowledge about its pollen and seed dispersal agents does not allow making strong predictions,

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

except that dispersal will depend on the vectors and on their use of the landscape. We also examined whether the relative stability of the forest cover in the past 70 years (Quéméré et al., 2012) is reflected in *N. spinifolia* genetic makeup, comparing the effect of recent and historical forest covers on gene flow, as a proxy for the temporality of its habitat loss and fragmentation. Finally, we present the application of our work to the conservation of the LM region forest network.

Material and methods

Study region

The Loky-Manambato (LM) region (Daraina; Fig. 1) is a biogeographical transition zone between dry deciduous and humid forests (Goodman & Wilmé, 2006), stretched over ~2,500km^2 between the Loky and Manambato Rivers. This region is crossed by the relatively shallow Manankolana River, bordered by riparian forests along most of its course, and by a national dirt road (Fig. 1). It consists of a dozen major forest patches, ranging from 8.3 to 62.5 km^2 and totaling ~360 km^2 (Goodman et al., 2018), separated by ~0.5-10 km of human-altered grasslands, dry scrub, agricultural lands and riparian corridors. Most forests are situated at low- to mid-elevations and mostly consist of dry deciduous vegetation. In contrast, some mountain forests (Binara and Antsahabe, plus Bobankora to a lower extent) are covered by a gradient of dry deciduous, transition, humid and ericoid vegetation (Gautier et al., 2006). Despite sustained grassland fires, slash-and-burn agriculture and charcoal production, as well as exploitation of wood, gold and sapphires (Fanamby, 2010; Goodman *et al.*, 2018), deforestation rate in the LM region is still relatively low (Quéméré *et al.*, 2012) compared with those of eastern and southwestern Madagascar (Vieilledent et al., 2018), likely stemming from its remoteness,

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

difficult accessibility and climate. However, to mitigate the threats, the LM region progressively became managed as a protected area by the Malagasy NGO "Fanamby" since 2005 (Fanamby, 2010; Goodman *et al.*, 2018).

Study species

Noronhia spinifolia (Oleaceae) is a small-sized, understory tree that is easily distinguished from other *Noronhia* species by its narrow linear leaves with a spiny tip. The plant has cream-white, urceolate, small (< 7 mm long), and hermaphroditic flowers, as well as small (< 10 mm long) and drupaceous fruits that have a thin mesocarp and a rather crustaceous endocarp. Flowering and fruiting typically occur from October to May, during the rainy season. Flower and fruit characteristics, along with observational accounts, suggest insect pollination [e.g. bees] and animal dispersal [e.g. birds, lemurs, rodents] (Hong-Wa, 2016), which may be short-distanced depending on the nature and behavior of the vectors (e.g. forest-dwellers). It is microendemic to northern Madagascar, mainly found in the LM region except for one record from further north in Montagne des Français, and is reported mainly in semi-deciduous forests of low elevation, mostly on alkaline substrate [e.g. limestone, calc-alkaline rocks]. *Noronhia spinifolia* has been assigned a preliminary conservation status of "Endangered" due to threats to its habitat (Hong-Wa, 2016).

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

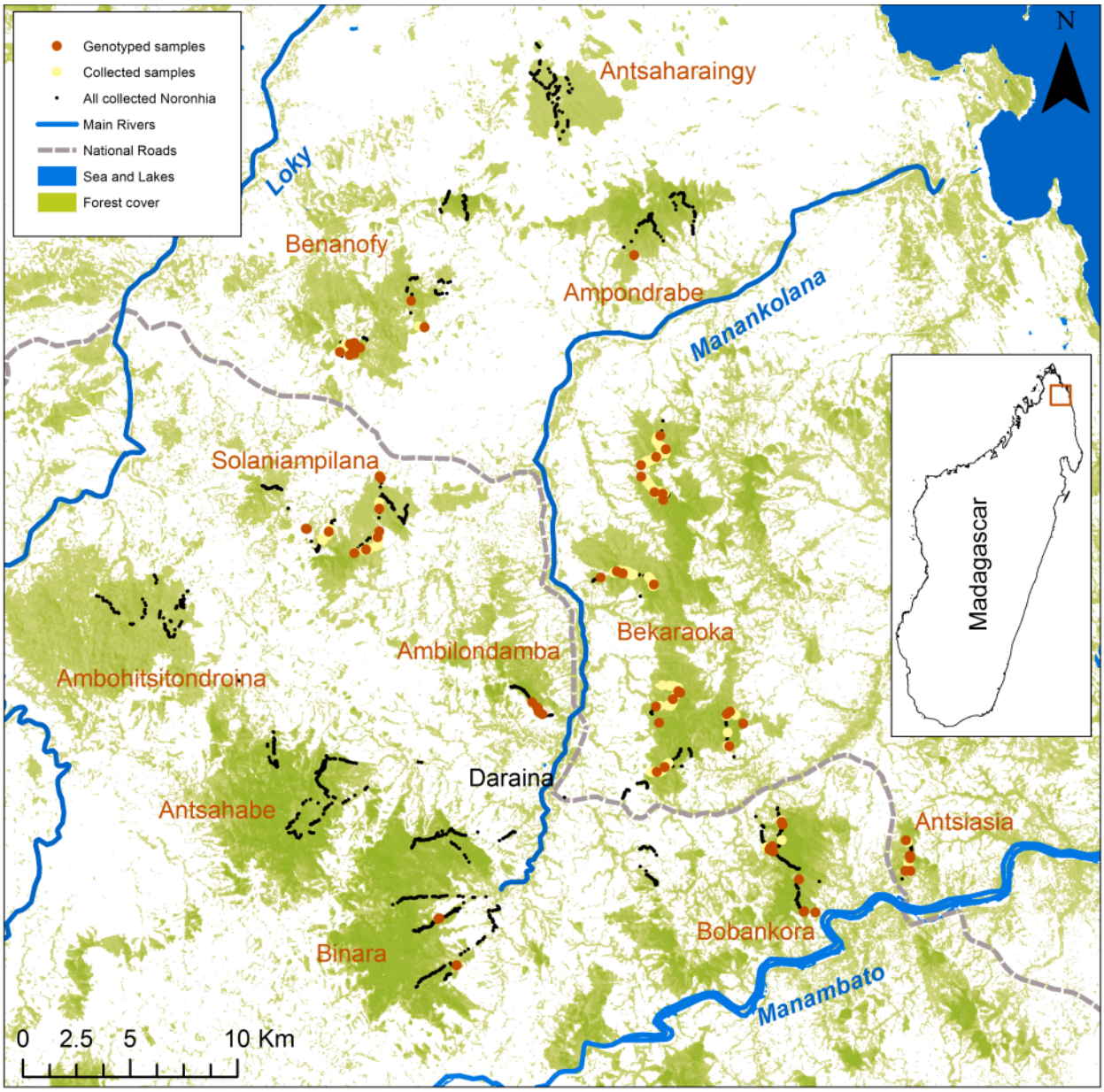


Figure 1: Map of *Noronhia spinifolia* sampling in the Loky-Manambato (LM) region.

The small black points represent samples collected for all *Noronhia* species (ca. 30 distinct taxa) and illustrate the survey effort conducted in the region. The yellow and red dots represent *N. spinifolia* samples, with the red dots corresponding to samples included in our genomic analyses. The forest cover is adapted from Hansen *et al.* (2013). Pixels with less than 30% tree cover are represented in white to illustrate the presence of riparian forests along streams of the LM region.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Plant sampling

To sample *N. spinifolia* populations, we surveyed all major forests of the LM region (Fig. 1) in 2017 and 2018, during the dry season (July-September), and used topography (altitude and shape) as a sampling guide to maximize the representation of all landscape features. Most surveys (black dots in Fig. 1) started from the forest edge at low altitude towards the forest core at higher elevation. We identified *Noronhia* species based on tree characteristics, leaf morphology and tissue structure, and collected leaf samples of 220 *N. spinifolia* trees - which tends to occur in patches - preserved in silica gel for DNA conservation. We prioritized fully-grown mature tree sampling because much of the density-dependent mortality takes place before maturity in trees, and their effective population size contributing to the genetic diversity is thus closer to the actual adult census size than to the size of the entire population including young trees and seedlings (Dodd et al., 1999; Petit & Hampe, 2006). Therefore, the regional patterns of diversity are expected to be better represented by adult samples. For each tree, we systematically recorded its height, diameter and reproductive state, as well as its geographical coordinates (GPS) and elevation. For all forests, at least one specimen voucher was prepared and deposited at the herbarium of the Parc Botanique et Zoologique de Tsimbazaza (TAN).

Laboratory procedures

DNA extraction, organellar and nuclear genotyping

We extracted DNA from 137 samples of *N. spinifolia* using the Qiagen Biosprint DNA Plant kit, followed by quality control procedures ensuring high quality genomic DNA. We subsequently genotyped 72 high DNA quality samples (Fig. 1, Methods S1); a cost-effective subsampling that nonetheless maximizes geographic and altitudinal representation, tends to limit clustered

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

sampling, and retains samples from all geographic distances, and that also prioritizes reproductively mature and fully-grown trees with a targeted sequencing depth >15×. Using a two-pronged approach, we genotyped 15 chloroplast microsatellites (cpSSR) and one mitochondrial microsatellite (mtSSR), originally developed on *Olea europaea* L. (Table S1, Methods S2, S3; Besnard *et al*., 2011), and also used restriction associated DNA sequencing (RADseq; generating data from the biparentally inherited nuclear genome and the mitogenome; Methods S4). RADseq consists in sequencing regions neighboring restriction sites, to obtain homologous sequences across individuals, spread across the genome, at a reasonable cost (Andrews et al., 2016; Baird et al., 2008).

Data processing

Organellar loci, nuclear ploidy and de-novo assembly of the nuclear loci catalog

After ad-hoc demultiplexing and cleaning of reads (Methods S4), we screened the organellar genomes using bwa-mem sequence alignment (Li, 2013) to the *N. clarinerva* Hong-Wa mitogenome and *N. spinifolia* plastome (MW202230 and MT081057, respectively; Methods S5). We identified ten mitochondrial *Sbf*I RAD (mtRAD) loci *in silico*, from which haplotypes were called using ANGSD v0.92 (Korneliussen et al., 2014; Nielsen et al., 2012), based on their highest effective base depth (Wang et al., 2013). Conversely, no cpDNA RAD locus was recovered, confirming *in silico* analyses (Methods S5). Organellar genomes are supposedly inherited from the mother in the olive tribe, and consequently their polymorphism should be closely linked (Besnard et al., 2002; Van de Paer et al., 2018). As we cannot directly verify this assumption on progenies of *N. spinifolia*, we here investigated the linkage disequilibrium [LD]

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

among organellar markers using the package *poppr* [Method S2 (Agapow & Burt, 2001; Kamvar et al., 2014, 2015)].

A catalog of nuclear tags (loci) was *de-novo* optimized (Methods S6) by iterating around the core parameters of Stacks (Rochette et al., 2019) to maximize the amount of available biological information (Paris et al., 2017). The final catalog was further cleaned (Methods S6) for exogenous contaminants using DeconSeq (Schmieder & Edwards, 2011) and endogenous orthologs using MUMmer (Kurtz et al., 2004). To enable ploidy-adapted analyses (catalog assembly and variants genotyping), ploidy was inspected using minor allele frequency plots and statistically confirmed using nQuire (Weiß et al., 2018).

RADseq genotyping

We used two fundamentally distinct genotyping approaches to ensure the robustness of our results: single nucleotide polymorphism (SNPs) called in Stacks, and genotype likelihoods (GLs) estimated with ANGSD (Methods S7). GLs retain information about uncertainty in base calls, which alleviates some issues associated with RADseq data such as unevenness in sequencing depth and allele drop-outs (Heller et al., 2021; Pedersen et al., 2018; Warmuth & Ellegren, 2019).

Population and landscape genetics

We conducted complementary analyses to assess the effect of landscape components on the genetic diversity of *N. spinifolia*. We first investigated the raw patterns of genetic diversity and structure without priors to describe the major trends and build hypotheses. Then, using univariate approaches under an isolation-by-resistance model (IBR; McRae, 2006), we assessed the effect of each landscape component, iterating through their cost (resistance or conductance) and

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

resolution (i.e. granularity or pixel size). Finally, we optimized the composite landscape resistance surface using a gradient based approach to assess the relative contribution of selected landscape components to *N. spinifolia*'s genetic connectivity across the area.

Genetic diversity

To assess *N. spinifolia*'s genetic diversity, its distribution in space and among genomic compartments, we estimated several diversity metrics. We estimated both the individual and the within-forest-averaged expected heterozygosity (H_E) according to Fumagalli (2013), as well as the proportion of heterozygous genotypes (H_O), from nuclear genotype likelihoods (GL), based on unfolded site frequency spectra estimated in ANGSD. We further estimated organellar diversity (h), the probability that two haplotypes are different (Nei, 1987).

Population structure

To understand how *N. spinifolia*'s population is structured across the LM region, we first used forest-based metrics and naive multivariate approaches. We assessed the level of genetic differentiation among forests with Reynolds' weighted F_{ST} (Reynolds et al., 1983) from GL inferred in ANGSD. We explored the genetic structure of our study system through naive clustering analyses (Methods S8), based on ANGSD GLs using NgsAdmix v32 (Skotte et al., 2013) and on Stacks called genotypes using ADMIXTURE v1.3.0 (Alexander et al., 2009), and with a principal component analysis (PCA) from GLs with PCAngsd. We estimated the level of organellar genetic differentiation among forests with Nei's weighted F_{ST} (Nei, 1973), Nei's D (Nei, 1973), and Edwards' chord distance [Edwards' D (Edwards, 1971)], based on frequency of haplotypes found in each forest and appropriate for organellar material haplotypic data (Grasty et al., 2020; Kohrn et al., 2017), using the R packages *hierfstat* and *adegenet*. We also investigated

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the phylogenetic structure of organellar DNA data using minimum spanning networks of genetic distances (see below) constructed with the R package *poppr* (Kamvar et al., 2015).

Isolation by distance

To assess how the geographic distance alone explains the above-mentioned among-individual genetic relationships [distance or relatedness], we investigated patterns of isolation by distance [IBD (Slatkin, 1993; Wright, 1943)]. We used Mantel tests (Mantel, 1967) between individual geographic and genetic distances (Methods S9). IBD may be limited to a certain scale and disrupted at larger scales (e.g. Keller & Holderegger, 2013; Van Strien *et al.*, 2015; Cayuela *et al.*, 2019), due to effective dispersal over recent or longer evolutionary time frames (Cruzan & Hendrickson, 2020). To assess the scale at which IBD is maximized, we compared data subsets upper-bonded by their maximum geographic distance (S) (Methods S9), and which represent the cumulative effect of all data included below S. In addition, to evaluate how effective dispersal is distributed geographically, we build Mantel correlograms (Oden & Sokal, 1986; Sokal, 1986), using the *ecodist* R package (Goslee & Urban, 2007).

We used the Bruvo's genetic distances (Bruvo et al., 2004) to consider the specific mutation model of the cpSSR data. Additionally, we used Nei's genetic distance [Nei's *D* (Nei, 1973)], appropriate for small scale study of organellar material haplotypic data [(Kohrn et al., 2017) Methods S3]. Finally, we estimated the covariance of nuclear RADseq GLs (Meisner & Albrechtsen, 2018), in PCAnsgd, as such metrics based on principal components analysis have been shown to be particularly fitted for landscape genetics (Shirk et al., 2017).

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Isolation by resistance

Landscapes are rarely homogeneous, and gene flow may be limited or facilitated by its components. To assess the cost (i.e. resistance/conductance) associated with effective dispersal through each landscape feature, we used an IBR approach (McRae, 2006). As detailed below, we assessed the unique and common contributions of each landscape variable, at varying cost and resolution (i.e. granularity or pixel resolution), for two movement models [Least Cost Path (Adriaensen et al., 2003) and Circuit Theory (McRae & Beier, 2007)], to the individual genetic distances of *N. spinifolia*, for each of the three genetic compartments (cpDNA, mtDNA, nDNA).

Landscape variables, cost and resolution

As *N. spinifolia* was recently described and occurs in a remote area (Hong-Wa, 2016), we had little prior knowledge on the landscape variables that may affect pollen and seed dispersal. We therefore assessed the effect of most available landscape variables (Table 1; Methods S10). To test if the genetic diversity of old trees may be better explained by past forest cover, we used forest cover data from 1953, 1973, and 2000s (Hansen et al., 2013; Vieilledent et al., 2018). Although strong priors associating a landscape component to a particular cost (resistance or conductance) may be available for well-studied species (e.g. Dellicour *et al.*, 2019; Quéméré *et al.*, 2010), landscape variables and their associated cost are often chosen almost arbitrarily when little or no data are available (Beier et al., 2008, 2011). To identify the variable-cost associations that matter for our study system, we iteratively tested 14 conductance-resistance values (Methods S10). Similarly, organisms do not necessarily perceive each environmental component at the same resolution [or granularity, expressed as pixel size (Chubaty et al., 2020; Galpern et al., 2012; Galpern & Manseau, 2013)]. To identify the variable-cost-granularity relevant for *N. spinifolia*, we tested four pixel resolutions (Methods S10).

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 1: Landscape variables.

Variable	Variable code	Type[a]	nRAD Z-value[b]	mt-cp Z-value[b]
Rivers	Rivers	D	nr	nr
Streams	Streams	D	nr	nr
Roads	Roads	D	nr	nr
Trails	Trails	D	nr	nr
Altitude / elevation	Alt	C	-16.56	-9.24
Slope	Slope	C	16.29	7.03
November wind speed	Wind_November / wind11	C	6.52	nr
Percent tree cover	%_tree_cov / ptc	C	16.73	12.21
Percent tree cover, discrete	%_tree_cov_dis	D	nr	nr
Forest cover ~2000	Veg_2000 / fc2000	D	nr	nr
Forest cover ~1973	Veg_1973 / fc1973	D	nr	nr
Forest cover ~1953	Veg_1953 / fc1953	D	nr	nr

Abbreviations. [a]D = Discrete, C = Continuous; [b]nr = not retained.

Movement model and analytical flow

To determine which dispersal model best applies to *N. spinifolia*, we used both the Least Cost Path (LCP) and the Circuit Theory (CT). These two approaches, respectively, consider the least cost trajectory and the cost of all possible trajectories (McRae & Beier, 2007). We computed landscape distances using the R package *gdistance* (van Etten, 2012).

We used a two-step procedure to assess the landscape components' relative contributions to the spatial structure of *N. spinifolia*'s genetic diversity. First we selected landscape components, as well as their best fitting cost, resolution, and movement model. We estimated the correlation between geographic or landscape distance and genetic matrices (i.e. Landscape variables and Genetic distances as described above) using Mantel tests (Mantel, 1967) in the R Package *vegan* (Dixon, 2003). We retained variables showing a better fit (R^2) than IBD, exhibiting sensitivity to cost values (i.e. variables with a fixed fit across all cost values were discarded), and investigated

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

their sensitivity to cost, movement model, and resolution. Second, we optimized the composite landscape conductance surface using the gradient based approach implemented in *radish* (Peterman & Pope, 2021) to assess the contribution of landscape variables, using a log linear conductance and a maximum likelihood population-effect parameterization of a mixed effects model [MLPE (Clarke et al., 2002; van Strien et al., 2012)]. MLPE accommodates the nonindependence of pairwise genetic and environmental distance and was assessed as the most accurate regression-based approach when conducting landscape genetics model selection in individual-based analyses (Shirk et al., 2018). Because the first and second step models are relatively distinct, we assessed the contribution of the variables retained from the first procedure step, as well as variables required to test specific dispersion hypotheses (wind, slope, streams network, etc.). To limit the analytical computation load, we hierarchically build models of increasing complexity (with one to four variables) combining the variables showing the best fit (AIC), and a significant effect at the previous hierarchical level. Altogether, we compared 32 organellar and 59 nuclear optimized composite landscape conductance models (Tables S2 and S3).

Results

Species occurrence

We sampled *N. spinifolia* in eight of the 11 surveyed major forests of the LM region (Fig. 1). The species occurs from low to medium elevation, between 87 and 505 m, but with strong discrepancies among forests (Fig. S1). While it was mainly recorded in dry forests, it was surprisingly found in dry to wet transition forests at medium elevation (451-505 m) in Binara. Furthermore, the species was not found in three major forest patches of the LM region - namely

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Antsahabe, Ambohitsitondroina and Antsaharaingy - despite (*i*) large prospection efforts in these forests, and (*ii*) apparently similar habitat as the neighboring forests harboring the species (Fig. 1).

Organellar genotyping, nuclear ploidy and catalog construction

Of the 15 chloroplast microsatellites, 14 showed polymorphism (Table S4), and allowed distinguishing 55 chlorotype profiles among 72 trees (Results S1). The ten mitochondrial RAD loci (mtRAD) allowed identifying 11 SNPs (Results S1; Table S5). The combination of mtRADs and the mtSSR locus permits the identification of 15 mitotypes among 72 trees (Table 2). Contrasting LD patterns were found within and among organellar markers (Fig. S2; Results S1), suggesting the occurrence of recombination in the mitogenome possibly due to occasional paternal leaks of mitochondria.

Individual-based minor allele frequency profiles displayed unimodal diploid patterns, confirmed by nQuire analyses, and echoing the low frequency of polyploid in the genus *Noronhia* (Gorrilliot et al., 2021). The nuclear catalog parameter space exploration iterating around the core parameters for Stacks [i.e. m – the minimum number of reads required to build a stack, M – the maximum number of differences between stacks of an individual allowed when building a locus; and N – the maximum number of differences between loci of multiple individuals allowed when building a loci] allowed selecting values ($m = 4$, $M = 5$, $N = 8$) that offer a trade-off between the coverage, loci number, and SNP number, while limiting the number of paralogs and the presence of contaminants (Results S2). The SNP-calling procedure showing low ability to recover the genetic makeup of *N. spinifolia* (when compared to the GL-based procedure; Figs S3-S8), we therefore limited its use to preliminary analyses (ADMIXTURE & genetic distances) and proceeded with the GL-based procedure for downstream analyses.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Genetic diversity

Chloroplast microsatellites revealed a relatively high genetic diversity with only two chlorotypes shared by individuals from more than one forest, resulting in a high probability that two randomly sampled haplotypes are different (h = 0.99) and a mean allelic richness (A_r; estimated for five individuals) of 2.41 (Table 2). Consequently, most forests showed an extremely high cpSSR genetic diversity (h > 0.92) with the exception of Binara that appeared slightly less diverse (h = 0.73; Table 2). A relatively high mitotype diversity was also revealed [h = 0.85 (ranging from 0.66 to 0.97 per forest), A_r = 2.12]. Contrastingly, most forest patches exhibit moderately high levels of nuclear diversity with H_E values ranging from 4.53 to 6.52 x 10^{-3}, with discrepancies within and among forests (Tables 2 and S1; Fig. S9). This diversity is not homogeneously distributed in space, and higher levels of genetic diversity seemingly occur in certain areas such as Solaniampilana (Fig. S10). Furthermore, genetic diversity does not seem influenced by altitude (Fig. S11).

Table 2: Chloroplast, mitochondrial and nuclear summary statistics.

	cpSSR				mtRAD				nRAD	
Forests	**N**	n_h	H	A_r	**N**	n_h	H	A_r	**N**	H_E
Ambilondamba	6	5	0.98	2.22	5	4	0.97	2.16	5	0.00510
Ampondrabe	1	1	-	-	1	1	-	-	1	-
Antsiasia	6	4	0.92	2.67	6	3	0.81	2.14	6	0.00563
Bekaraoka	25	19	0.99	2.38	23	5	0.66	2.04	23	0.00652
Benanofy	11	8	0.94	2.39	11	4	0.78	2.26	11	0.00559
Binara	5	2	0.73	2.36	5	2	0.73	2.05	5	0.00453
Bobankora	11	10	0.99	2.45	11	3	0.73	2.04	11	0.00624
Solaniampilana	10	8	0.97	2.37	10	5	0.87	2.17	10	0.00624
Total / Mean	75	55	0.99	2.41	72	15	0.85	2.12	72	0.00676

N = number of analyzed individuals; n_h = number of haplotypes; h = haplotype diversity; A_r: allelic richness (estimated for five individuals), H_E = expected heterozygosity.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Population structure

The chloroplast and mitochondrial data both revealed substantial differentiation among forests (F_{ST} estimates ranging from 0.040 to 0.393 for cpSSRs; and 0.005 to 0.661 for mtRADs). As expected, a strong differentiation was also observed when combining cpDNA and mtDNA data (F_{ST} estimates ranging from 0.101 to 0.401; Table S6). The Solaniampilana and Benanofy forests, the farthest apart from other forest patches, formed an easily distinguishable haplotype group for both mtDNA and cpDNA (Figs S12-S13), while other more adjacent forests showed more limited divergence among them. Haplotype networks based on cpSSR and/or mtRAD data also revealed that one maternal lineage is unique to Solaniampilana and Benanofy (Fig. 2). Furthermore, the geographic Euclidean distances showed low, but highly significant, power at explaining genetic distances among individuals (R^2 [cpSSR]: 11.7%; R^2 [mtRAD]: 20.7%; and R^2 [cpSSR + mtRAD]: 21.3%; Figs S8, S14, S15; Results S3) and supported by the Mantel correlogram positive and significant values within 10-12 km classes (Fig. S16).

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

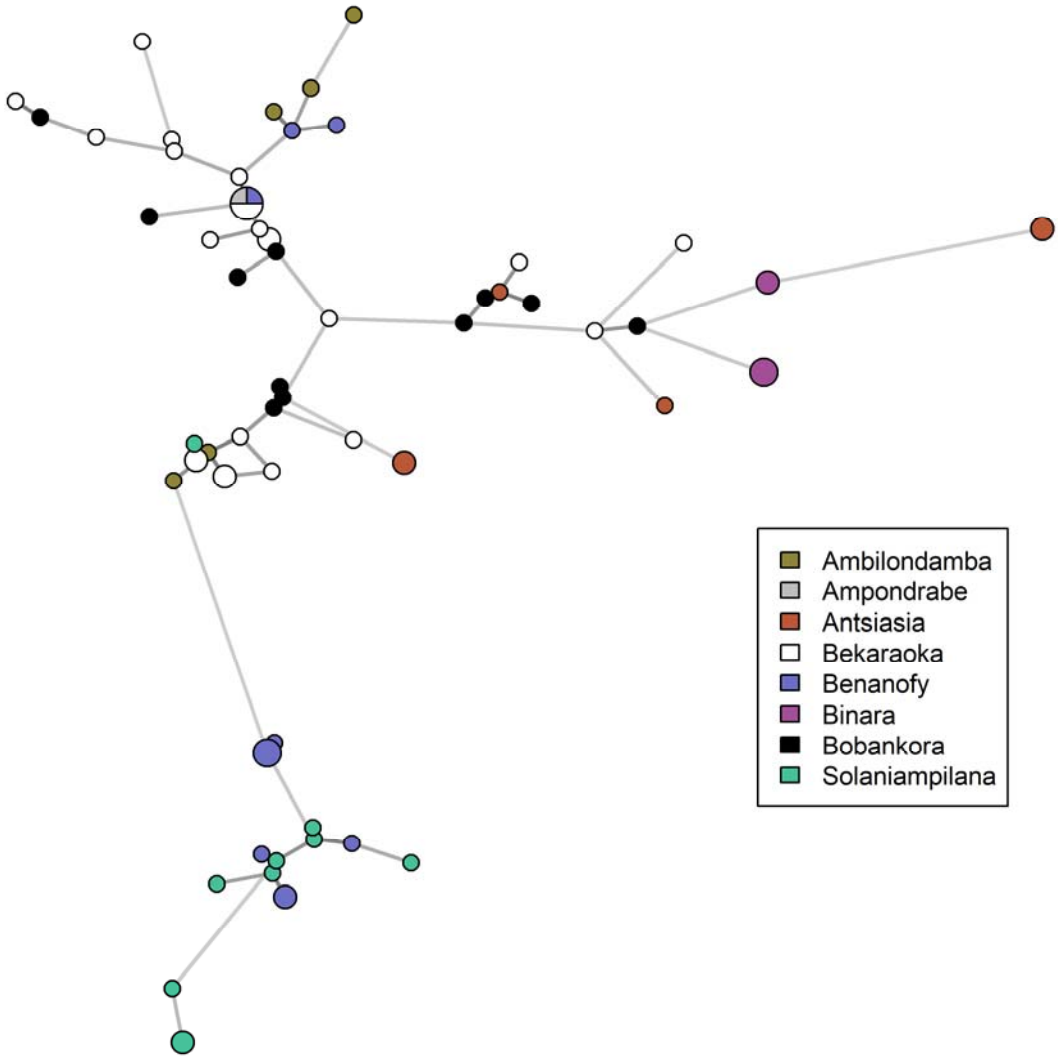


Figure 2: Organellar DNA haplotype network of *Noronhia spinifolia*.

Line length and gray scale are proportional to the Bruvo's cpDNA + Manhattan mtDNA combined genetic distances between distinct organellar haplotypes. Pie chart size is proportional to the occurrence number of a given haplotype. All edges of equal weight are represented. Distances among haplotypes are represented both through longer edges and the gray scale. The network highlights the huge organellar DNA diversity in *N. spinifolia*, with only one haplotype shared by individuals from at least two forests. It further shows a limited spatial structure, with, for instance, haplotypes from Solaniampilana and Benanofy grouping together at the bottom of the network.

F_{ST} estimates based on nuclear markers (Table S7) ranged from 0.089 to 0.210, indicating that most forests are differentiated from each other. However, we found no strong structure in sub-populations, with no particular support for number of clusters >1, both for GL- and SNP-based

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

analyses (Figs S3, S4). Instead, we found a clear northwest-southeast signal of continuous genetic differentiation across space, through GL-based PCA (First axis, ~15% of the variance explained; Fig. S17), clustering (Figs 3, S5, S6), and IBD analyses (Figs S8, S15). The observed continuous structure is well illustrated by the clustering structure for K = 3 that shows admixed patterns at all sampling sites (Fig. 3). We found a clear IBD signal explaining up to 56.6% of the among-individuals nuclear GL covariance (Fig. S15), further supported by the Mantel correlogram positive and significant values within 14 km classes (Fig. S16).

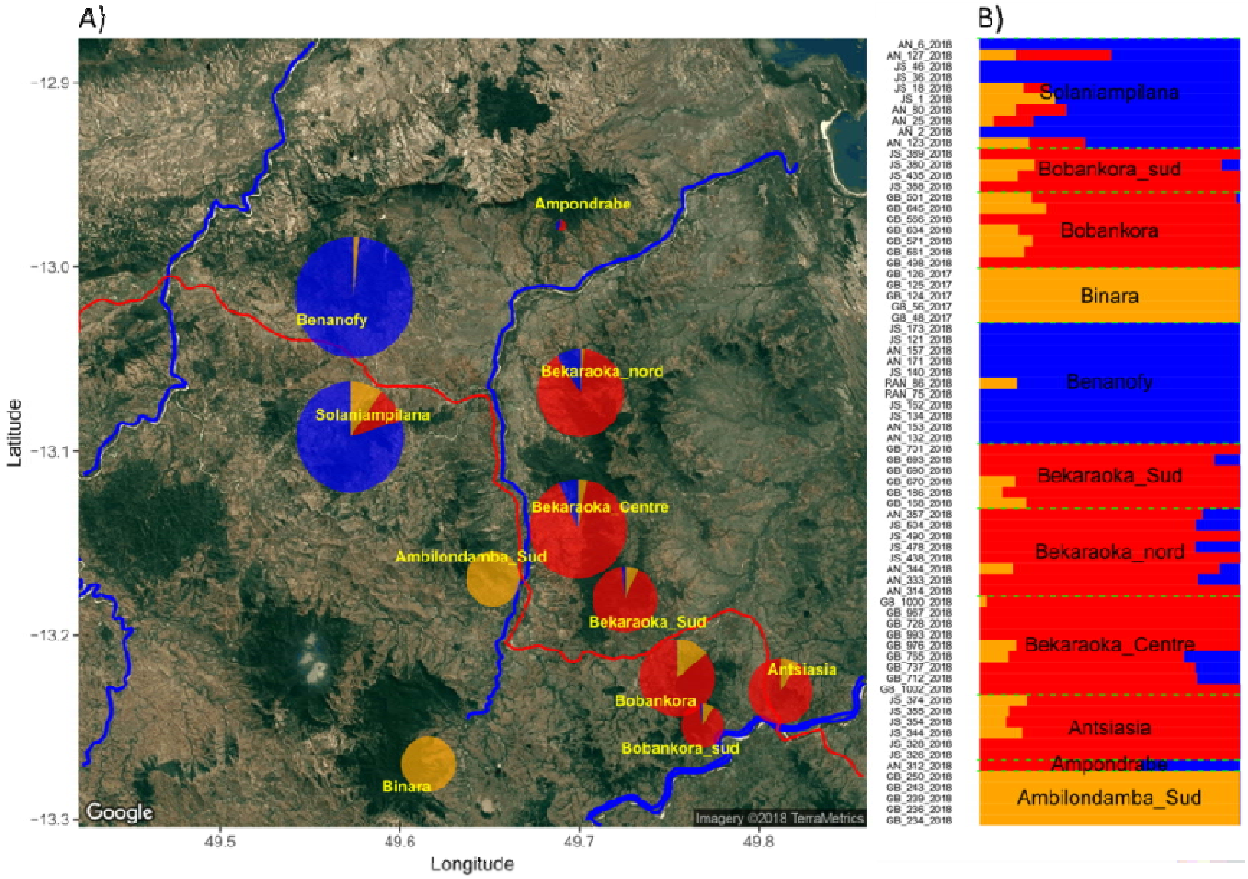


Figure 3: Spatial genetic structure of *Noronhia spinifolia* in the Loky-Manambato region. NgsAdmix ancestry proportions (for K = 3 genetic clusters) represented either **(A)** spatially by sampling site, or **(B)** per individual. Size of pie charts (in A) is proportional to the number of samples per site. Pie shares represent the sums of individual ancestry proportions that are shown

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

in B. Results are arbitrarily represented for $K = 3$, according to the likelihood and delta$K$ results in Fig. S8, because this K value best illustrates the continuous pattern of structure inferred using ngsAdmix and other approaches.

Landscape genetics

The univariate Mantel tests of genetic and landscape distances (IBR) showed lower linear fit for cpDNA (R^2 max ~0.14) than for mtDNA (R^2 max ~0.38) and nDNA (R^2 max ~0.90). Among the four vegetation layers, the continuous and discrete percent tree cover layer always exhibited the highest fit for conductance values at high resolution with cpDNA, mtDNA and nDNA individual-based genetic distances (R^2 = 0.14; 0.38 and 0.90, respectively; Figs S18-S21). Altogether, the parameter space exploration reveals a strong effect of all forest cover layers, whereas some other variables (i.e., rivers, roads and slope) may have subtle lower effects too. To build multivariate models, we retained in priority landscape variables showing a better fit (R^2) than the null model (IBD), and exhibiting sensitivity to cost values (e.g. % forest cover).
The hierarchically-built composite landscape conductance surface of increasing complexity confirmed that recent forest cover layers – the percent of tree cover, and the forest cover in 1973 to a lower extant – showed high univariate fit and log-collinearity with nuclear and organellar genetic data (Fig. S22). In addition, among monthly wind speed, that of November showed the best fit and highest collinearity with nuclear data (Fig. S23). The best fitting composite landscape conductance surface optimization for nuclear data (Table S3) retained the percent of tree cover (Hansen et al., 2013), the elevation (altitude), the wind speed in November and the slope (Table 1, Fig. 4). Similarly, the model best fitting the organellar data (Table S2) retained the percent of tree cover, the elevation and the slope (Fig. S24). For both nuclear and organellar data, the percent of tree cover always displayed the highest positive z-value for all covariates (Table 1),

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

whereas the elevation was the only variable exhibiting negative relationship with genetic distances (Table 1, Figs 4 and S24).

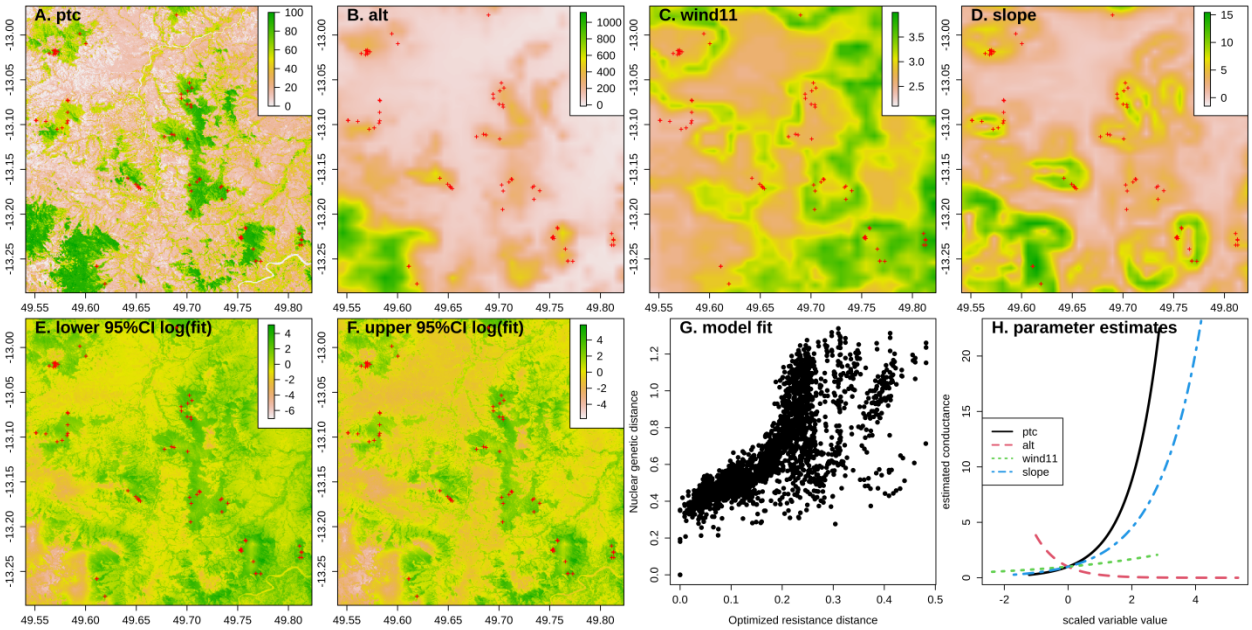


Figure 4: Landscape contribution to nuclear gene flow in *Noronhia spinifolia*.
The figure shows the four landscape variables (**A-D**) likely contributing to *N. spinifolia* nuclear gene flow (pollen and seeds), the confidence interval of their fitted composite conductance (**E-F**), the model fit relationship between the optimized resistance distance and the nuclear genetic (**G**), and the relative effects of the four retained variables (**H**). It illustrates a major conducting effect of the current forest cover (ptc: percent of tree cover) on the connectivity of *N. spinifolia*, and it further shows a positive effect of the wind speed in November [wind11] and a composite effect of the topography (positive effect of slope and negative effect of altitude [alt]).

DISCUSSION

With a comprehensive and extensive sampling of *Noronhia spinifolia* in its core distribution area, and leveraging the rare combination of nuclear and mitochondrial RADseq data with cpDNA microsatellites, this study allowed us to reveal a strong effect of forest cover on gene flow in a patchy habitat in northern Madagascar. We not only report a surprisingly high

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

organellar genetic diversity unevenly distributed in space, but also found that GL-based approaches were able to recover much more information than SNP-calling approaches in our model species. Moreover, the iterative parameter space exploration and the optimization of resistance surface allowed identifying outstanding landscape variables with a strong effect on the connectivity of *N. spinifolia*. Finally, we show that recent forest cover better explains the genetic structure of *N. spinifolia* than more ancient ones.

***Noronhia spinifolia*, a highly diverse Malagasy microendemic**

Our analyses first exhibit unexpectedly high organellar genetic diversity (h_{cp} = 0.99; 55 chlorotypes for 72 individuals, and h_{mt} = 0.85; 15 mitotypes) in a microendemic Malagasy tree species. Such cpDNA diversity is tremendously higher than that of another microendemic congener of the High Plateau (*N. lowryi*) when using the same 15 cpSSR loci [6 haplotypes in 77 individuals; h_{cp} = 0.58 (Salmona *et al.*, 2020)]. More surprisingly, more cpDNA haplotypes and diversity were revealed in 72 *N. spinifolia* individuals than in 1263 wild olive trees (*Olea europaea*) from the whole Mediterranean basin [47 chlorotypes; h_{cp} = 0.35 (Besnard et al., 2013)] and thus across very different geographic scales (LM region = 2,500 km² *vs* Mediterranean basin = ~2.5 Million km²) and despite the use of more polymorphic cpSSRs (n = 35) in the Mediterranean olive. Similarly, the *N. spinifolia* mtDNA diversity is also higher than in the Mediterranean olive [4 mitotypes; h_{mt} = 0.58; (Besnard et al., 2002)], although comparable diversity levels have been revealed in other plant groups exhibiting large mitogenomes with high mutation rates as *Silene vulgaris* (Moench) Garcke in Central Europe [30 mitotypes; h_{mt} = 0.94; (Štorchová & Olson, 2004)]. Secondly, the moderately high levels of nuclear genomic diversity (H_E = 4.53-6.52 x 10^{-3}) are within the range of that estimated in poplar, pedunculate oak and Norway spruce populations across distribution ranges several order of magnitude larger (Chen et

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

al., 2019; Ma et al., 2018; Plomion et al., 2018). Nonetheless, our unexpectedly high species genetic diversity is also comparable to that of other range-restricted microendemics (e.g. Eliades *et al.*, 2011; Lanes *et al.*, 2018).

Although high standing genetic diversity is generally common in plants, including forest trees as our focal species, the relative importance of the multiple mechanisms generating and maintaining this diversity are still debated (Isabel et al., 2020; Petit & Hampe, 2006; Scotti et al., 2016). In *N. spinifolia*, several non-exclusive evolutionary mechanisms may explain such a high intraspecific genetic diversity. Firstly, it suggests that long-term maintenance of a large effective population size precluded significant genetic drift. Persistent connectivity between forest patches may have been key in this process, particularly during climatic fluctuations of the Late Quaternary that may have contributed to fragmenting habitat, as suggested for other species of the LM region (Salmona, Heller, Quéméré, et al., 2017; Sgarlata et al., 2019). Secondly, the genus *Noronhia* has extremely diversified in northern Madagascar (Hong-Wa, 2016), and about 30 taxa (including already-described species, new species to be described, potential hybrids, and cryptic species) have been recently recorded and sampled in the LM region (JS & GB, unpublished data). What caused such diversification remains unknown. But the co-occurrence of closely related taxa may offer some opportunities for hybridization events, which could have contributed to the increased genetic diversity in *N. spinifolia*. However, the cpSSR characterization of four sympatric/parapatric LM *Noronhia* (i.e. *N. candicans* H.Perrier, *N. clarinerva*, *N. crassinodis* H.Perrier and *N. intermedia* Hong-Wa; > 200 individuals), closely related to *N. spinifolia* (according to cpDNA and nDNA data; Salmona *et al.*, 2020), shows that these species have no shared chlorotype with our study model (GB, unpubl. data), thus suggesting that maternal introgression events to *N. spinifolia*, if any, may not be recent. Lastly, high mutation rate may

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

also contribute to the high genetic diversity in *N. spinifolia*. An obvious acceleration of the mitogenome evolutionary rate has been recently documented in the closely related species *N. candicans*, *N. clarinerva*, *N. intermedia* and *N. spinifolia*, with a high number of di- or tri-nucleotide mutations possibly reflecting frequent mtDNA recombination in this clade (Van de Paer, 2017), as also suggested by a lack of LD between some SNPs (Fig. S2). While accelerated mutation rate was missing on the plastome (Salmona et al., 2020), we are still lacking any evidence for the nuclear genome. Such accelerated evolutionary rate could result from relatively frequent and recurrent hybridization events in this group, promoting genomic instability (Fontdevila, 2005; Payseur & Rieseberg, 2016). Moreover, the strong linear relationship between geographic and genetic distance could preclude cryptic radiation (Pillon et al., 2014) and microgeographic adaptation (Scotti et al., 2016) as major drivers of the observed diversity. In conclusion, the surprisingly high genetic diversity calls for the identification of the evolutionary, ecological and/or molecular mechanisms underlying this peculiar pattern.

Landscape effects on the genetic diversity of *Noronhia spinifolia*

A strong continuous spatial structure

Beyond revealing surprisingly high levels of genetic diversity, our results also show complementary signals of a strong continuous structure in space (PCA, clustering and IBD), from both organelles and the nucleus, in contrast to generally expected incongruent patterns among genomes (Bianconi et al., 2020; Olofsson et al., 2019). While the northwest-southeast differentiation cline represented as much as ~15% of the variance of the PCA on nuclear data, the geographic Euclidean distance alone explained up to ~55% of the nuclear genetic variance using IBD tests. This strong pattern of nuclear genetic structure sharply contrasts with the

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

absence of nuclear spatial structure in the savanna olive tree, *N. lowryi* (Salmona et al., 2020). However, reported IBD patterns in trees show a wide range from low values in *Dalbergia monticola* Bosser & R.Rabev. across eastern Madagascar humid forests [R^2 = 0.18; (Andrianoelina et al., 2009)], or *Coffea mauritiana* Lam. in the Réunion Island [R^2 = 0.21; (Garot et al., 2019)], to high values in *Swietenia macrophylla* King in Central America [R^2 = 0.62; (McRae & Beier, 2007)]. Unexpectedly, this genetic structure was here extremely well explained by the vegetation cover (percent tree cover; mtDNA R^2 = 0.38; nDNA R^2 = 0.90), and to a lower extant by the topography (elevation and slope) and by the wind speed in November. Although strong landscape effects were also found in *S. macrophylla* (McRae & Beier, 2007), we report a unique evidence of a strong habitat effect explained predominantly by one landscape variable.

On seed-mediated gene flow: the organellar DNA testimony

Although organellar IBR patterns (Figs S15, S18-20, S24) suggest that seed-mediated gene flow is driven by forest cover, the recovered pattern was of lower intensity than for pollen-mediated gene flow (nDNA). Despite watershed networks being candidates for hydrochory, only slope (barochory) and altitude showed noticeable effect on seed dispersal. Similarly, the overall structures of organellar haplotype networks (Figs 2, S12-S13) are coherent with the geographic repartition of forests, and in line with the effect of the forest cover. These prevailing effects of forest cover, and the limited haplotype sharing between forests, suggest that seed dispersal may be relatively limited, primarily driven by slope (barochory) and performed by forest-dwelling animals (zoochory), especially those with limited and/or rare trans-forest movements, such as lemurs, rodents and territorial birds (Aleixo-Pais et al., 2019; Quéméré et al., 2010; Rakotoarisoa et al., 2013a; Sgarlata et al., 2018). Furthermore, the limited haplotype sharing between forests suggest that *N. spinifolia*'s organellar diversity is the result of mutation-drift rather than that of

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

dispersal-drift processes (Cruzan & Hendrickson, 2020). However, the networks also show multiple potential fluxes among forests, hence recalling the network complementarity to the IBR approach. Several non-exclusive interpretations can be invoked for explaining these patterns: (*i*) relevant landscape variables are not included or of low resolution (e.g. watershed networks and climatic variables); (*ii*) the cpDNA and mtDNA diversities are confounded by homoplasy, recombination, strong drift, long-term phylogenetic or demographic history; (*iii*) seed dispersal also results from infrequent seed ingestion by wide-ranging birds (or other vertebrates); and (*iv*) the sampling possibly limited resolution at the seed dispersal scale.

A deep forest cover effect on gene flow

Unlike organellar DNAs, nDNA diversity is deeply explained by the LM region forest cover (Fig. 4). While this partially confirms the effect of forest cover, elevation and slope on seed dispersal, since nDNA diversity is influenced by both seed and pollen movement, wind-mediated pollen dispersal is also supported. Indeed, November wind speed, which fits the flowering temporal window of *N. spinifolia* (Hong-wa, 2016), shows a subtle but significant effect (Fig. 4) suggesting that pollen dispersal or pollinator movements could be influenced by the wind regime. Combined, the strong forest cover effect, the slope, elevation and wind effects suggest that pollen dispersal could be mediated by forest-dwelling organisms with movements limited by non-forest environments and high elevation, but facilitated by slope and November wind speed. Insect-mediated pollen dispersal in *N. spinifolia* is also strongly suggested by its flower morphology and color (Hong-Wa, 2016). However, the currently limited knowledge of the Malagasy entomofauna and plant-pollinator networks prevents us from clearly identifying this species' forest-dwelling pollinators. Furthermore, our analyses are limited by (*i*) the lack of high-resolution wind directionality data, and (*ii*) their lack of integration (Fernández-López & Schliep,

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

2018; Kling & Ackerly, 2021) with composite landscape surface optimization approaches (Peterman, 2018; Peterman & Pope, 2021).

The antiquity of habitat mosaic in northern Madagascar

Our results further support a long-standing habitat mosaic in the LM region. First, the better fit of the recent forest cover (~2000s), compared to older vegetation cover (~1953, ~1973), suggests that the small forest changes that have occurred through this period (Quéméré et al., 2012) are unable to explain the genetic diversity of *N. spinifolia*. These mild landscape changes, mostly located at the forest patches edges (Quéméré et al., 2012), contrast with the high deforestation rates observed throughout Madagascar since the 1950s. They might be explained by the low human settlement density (Green & Sussman, 1990; Schüßler et al., 2020) and the dominant dry forest type (Hansen et al., 2013; Vieilledent et al., 2018) of the LM region. Under such high recent deforestation rates, a better fit of the recent forest cover layer would be very unlikely, even considering that its better resolution could positively bias its fit. Second, because we mostly genotyped fully-grown mature trees, and since the generation time of *Noronhia* is potentially long [>20-50 years; (Salmona et al., 2020)], the genetic diversity is expected to reflect ancient forest cover. The time lag for a particular landscape feature to imprint its effects in the genetic diversity of a species, has been little studied (Landguth et al., 2010; Mona et al., 2014). However, in *N. spinifolia*, based on the strength of the signal, the high level of diversity and of gene-flow, the re-shuffling of allele frequencies after fragmentation can be relatively long (tens to hundreds of generations), before harboring the signature of the new geographical pattern. The period with data on forest cover (1953-present) represents less than five generations, a too short period to erase the signal of previous population structure (or lack thereof). This suggests that the landscape changes leading to the current forest cover long predates the most ancient available

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

layer (1953). The strong genetic correlation with the recent forest cover is, therefore, sound evidence that the landscape of the LM region was relatively stable at least for the last century (i.e. when most of Madagascar's deforestation occurred), and possibly the last millennium. This result concurs with those of recent studies (Quéméré et al., 2012; Salmona et al., 2020) supporting a relative antiquity of habitat mosaic in northern Madagascar. Furthermore, both the high diversity of *N. spinifolia*, and its predominant distribution in low-elevation dry forest suggests that this habitat type may have been spatially, topographically, and temporally extensive in northern Madagascar, albeit frequently fragmented, as seemingly evidenced by a rare and likely relictual occurrence of the species in contemporary high-elevation humid forest (e.g. Binara) and similarly peculiar presence further north (e.g. Montagne des Français). To assess forest-cover changes over a larger time frame (e.g. the last ten or so millennia), inferences of *N. spinifolia*'s demography over time would be relevant (Beichman et al., 2018; Salmona, Heller, Lascoux, et al., 2017). Coupling these inferences, with that of short-generation grassland organisms, would also help clarifying the dynamics of fire-prone non-forest environments, through the succession of environmental changes that occurred during last millennia, namely the last-glacial-maximum, early human's colonization, the mid-Holocene transition, and the 1-Kya expansion of agro-pastoralism.

Further prospects and conservation implications

The power of coupling genomic data to landscape genetics allowed not only identifying major landscape components influencing effective dispersal, but also their respective effects on seed and pollen dispersal. This surprising result warrants further investigation using higher resolution landscape and environmental layers, not used, or not available to our study. In particular, it would benefit from the use of forest type, soil type, land use, and climate data of better

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

resolution. In addition, the wind effect has been inferred without considering its directionality. Recent landscape genetics analytical advances allowing directionality integration (Fernández-López & Schliep, 2018; Kling & Ackerly, 2021) are still limited by the lack of high-resolution directional data (e.g. wind), and of integration with composite landscape surface optimization approaches (Peterman, 2018; Peterman & Pope, 2021). These analyses may also benefit from the use of genetic distances reflecting the directional changes in diversity among pairs of samples. Recent methodological development contrasting with previous evaluations of the MLPE (Peterman et al., 2019; Shirk et al., 2018; Winiarski et al., 2020) showed that the Generalized Wishart has higher accuracy in identifying 2-variables composite landscape surface (Peterman & Pope, 2021). Although this is beyond the scope of our study, cross-validation approaches could assess this new optimization model, the effect of seemingly clustered sampling of *N. spinifolia*, or that of modifying the conductance function (Peterman, 2018). While our study clearly identifies that seed and pollen are dispersed by forest-dwelling organisms, it neither identifies these organisms nor does it clearly show that seed and pollen do still effectively disperse among forests. These questions could be tackled (i) by inferring pedigree data from high density population sampling, coupled with sampling of young trees and seedlings, (ii) using field survey of potential dispersers during flowering and fructification (e.g. camera tracking), and/or (iii) using metabarcoding approaches to assess the interaction network within the LM forests.

While our study confirms the biological importance of the LM region, which is known for its species richness and endemism across taxa (Goodman & Wilmé, 2006; Rakotondravony, 2006, 2009; Sgarlata et al., 2019), and more specifically for the genus *Noronhia* (Hong-Wa, 2016), our results also have several implications for biodiversity conservation in the region:

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

- First, they underscore the conservation value of the often-overlooked intraspecific genetic diversity, which is unexpectedly high in *N. spinifolia*.

- Second, this study highlights the importance of riparian forests of the LM region for their major role both as corridors connecting forest patches, which is supported by the fact that genetic diversity in *N. spinifolia* is explained by forest cover rather than geographical distance, and as vectors promoting the roles of vertebrates and insects on seed and pollen dispersal. Therefore, actively maintaining, protecting, and reforesting riparian and corridor forests, which are likely pivotal for the functional connectivity of *N. spinifolia* but also most native and endemic species of the LM region (Aleixo-Pais et al., 2019; Rakotoarisoa et al., 2013a; Sgarlata et al., 2018), remain critical conservation actions.

- Third, our study identifies the Binara forest as unique among the major forests of the LM region and in urgent need of deeper conservation focus. Indeed, our extensive forest survey allowed us to find and collect just a few *N. spinifolia* samples in this forest, where they were found only at unexpectedly higher altitude and wetter habitat (Fig. S1). Similarly, several other Malagasy olive species that are mostly distributed in dry forests [e.g. *N. ankaranensis* (H.Perrier) Hong-Wa, *N. candicans*, *N. christenseniana* Hong-Wa and *N. oblanceolata* H.Perrier; GB and JS unpublished data], were also found to occur only at higher altitude in the mountain evergreen forests of this region (e.g. Binara and Antsahabe). Altogether, this pattern, though unclear, echoes the peculiarities of these forests that likely acted as refugia for numerous taxa during drier periods (Goodman & Wilmé, 2006; Rakotoarisoa et al., 2013b; Raxworthy & Nussbaum, 1995; Sgarlata et al., 2019).

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Data accessibility

Raw RADseq data and RADseq mtDNA alignments have been deposited at the Short Read Archive (SRA) NCBI database under the reference PRJNA632767. Organellar microsatellite genotypes and mtRAD variants are available in Tables S9 and S10, respectively. All additional data, scripts and materials are available to readers at https://doi.org/10.5281/zenodo.7193354 .

Benefit sharing

Benefits Generated. Research collaborations were developed and strengthened with scientists from Madagascar, the country of origin of the genetic samples. One Malagasy student (AER) was trained during this work. All academic collaborators are included as co-authors. Three guides from the Loky-Manambato protected area protected area guides and cook association were trained to botanical work, as well as several additional guides in remote communities. The research results have been and will be repeatedly shared with the concerned communities, the managers of the protected area, and the broader scientific community. The research also addresses a priority concern, in this case the conservation of the focal organism, *N. spinifolia*, and by extent its habitats, the Loky-Manambato region and Northern Madagascar. More broadly, our group is committed to international scientific partnerships, as well as institutional capacity building.

Conflict of interest disclosure

The authors of this article declare no financial conflict of interest with its content.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Acknowledgments

We thank the Direction Générale du Ministère de l'Environnement et des Forêts de Madagascar, Madagascar's Ad Hoc Committee for Fauna and Flora, and Organizational Committee for Environmental Research (CAFF/CORE) for permission to perform this study (permits number: [49/17] & [127/18]/MEEF/SG/DGF/DSAP/SCB.Re) and for their support. We thank the local communities of Daraina and of the LM region for their warm reception and support. We thank E. Rasolondraibe and the many local guides and cooks for sharing their incomparable expertise and help in the field, *misaotra anareo jiaby*. We thank U. Suescun, C. Verbeke and M.A. Naranjo Arcos for lab assistance, L. Chikhi for comments on an earlier version of the draft. This work was mostly funded through an ERA-NET BiodivERsA project: INFRAGECO (Inference, Fragmentation, Genomics, and Conservation, ANR-16-EBI3-0014). We also thank the LABEX TULIP (ANR-10-LABX-0041) and CEBA (ANR-10-LABX-25-01), and the LIA BEEG-B (Laboratoire International Associé – Bioinformatics, Ecology, Evolution, Genomics and Behaviour, CNRS). We are grateful to the Get-Plage sequencing, Genotoul bioinformatics (Bioinfo Genotoul) and GenOuest bioinformatics platforms for sequencing services and providing computing resources. The fourth preprint version of this manuscript has been peer-reviewed and recommended by Peer Community In Evolutionary Biology (https://doi.org/10.24072/pci.evolbiol.100136).

Author Contribution

JS and GB designed the experiment. JS, AER, BLP, JR, CHW and GB were pivotal to field material collection and herbarium composition. JS, SM, and GB generated the genetic data. JS conducted bioinformatics and population genetic analyses. JS and AD conducted IBR analyses.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

JS and GB drafted a first version of the manuscript with significant input from CHW. All co-authors agreed with the last version of the manuscript.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

Adriaensen, F., Chardon, J. P., De Blust, G., Swinnen, E., Villalba, S., Gulinck, H., & Matthysen, E. (2003). The application of 'least-cost'modelling as a functional landscape model. *Landscape and Urban Planning*, *64*(4), 233–247.

Agapow, P.-M., & Burt, A. (2001). Indices of multilocus linkage disequilibrium. *Molecular Ecology Notes*, *1*(1–2), 101–102.

Aleixo-Pais, I., Salmona, J., Sgarlata, G. M., Rakotonanahary, A., Sousa, A. P., Parreira, B., Kun-Rodrigues, C., Ralantoharijaona, T., Jan, F., Rasolondraibe, E., Minhós, T., Zaonarivelo, J. R., Andriaholinirina, N. V., & Chikhi, L. (2019). The genetic structure of a mouse lemur living in a fragmented habitat in northern Madagascar. *Conservation Genetics*, *20*(2), 229–243.

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*(9), 1655–1664.

Allendorf, F. W., Hohenlohe, P. A., & Luikart, G. (2010). Genomics and the future of conservation genetics. *Nature Reviews Genetics*, *11*(10), 697–709.

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, *17*(2), 81–92.

Andrianoelina, O., Favreau, B., Ramamonjisoa, L., & Bouvet, J.-M. (2009). Small effect of fragmentation on the genetic diversity of *Dalbergia monticola*, an endangered tree species of the eastern forest of Madagascar, detected by chloroplast and nuclear microsatellites. *Annals of Botany*, *104*(6), 1231–1242.

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., Selker, E. U., Cresko, W. A., & Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One*, *3*(10), e3376.

Balkenhol, N., Cushman, S., Storfer, A., & Waits, L. (Eds.). (2016). *Landscape genetics: Concepts, methods, applications*. John Wiley & Sons.

Beichman, A. C., Huerta-Sanchez, E., & Lohmueller, K. E. (2018). Using genomic data to infer historic population dynamics of nonmodel organisms. *Annual Review of Ecology, Evolution, and Systematics*, *49*, 433–456.

Beier, P., Majka, D. R., & Spencer, W. D. (2008). Forks in the road: Choices in procedures for designing wildland linkages. *Conservation Biology*, *22*(4), 836–851.

Beier, P., Spencer, W., Baldwin, R. F., & McRae, B. H. (2011). Toward best practices for developing regional connectivity maps. *Conservation Biology*, *25*(5), 879–892.

Besnard, G., Hernández, P., Khadari, B., Dorado, G., & Savolainen, V. (2011). Genomic profiling of plastid DNA variation in the Mediterranean olive tree. *BMC Plant Biology*, *11*(1), 80.

Besnard, G., Khadari, B., Baradat, P., & Bervillé, A. (2002). Combination of chloroplast and mitochondrial DNA polymorphisms to study cytoplasm genetic differentiation in the olive complex (*Olea europaea* L.). *Theoretical and Applied Genetics*, *105*(1), 139–144.

Besnard, G., Khadari, B., Navascués, M., Fernández-Mazuecos, M., El Bakkali, A., Arrigo, N., Baali-Cherif, D., Brunini-Bronzini de Caraffa, V., Santoni, S., Vargas, P., & Savolainen, P. (2013). The complex history of the olive tree: From Late Quaternary diversification of Mediterranean lineages to primary domestication in the northern Levant. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1756), 20122833.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Bianconi, M. E., Dunning, L. T., Curran, E. V., Hidalgo, O., Powell, R. F., Mian, S., Leitch, I. J., Lundgren, M. R., Manzi, S., Vorontsova, M. S., Besnard, G., Osborne, C. P., Olofsson, J. K., & Christin, P.-A. (2020). Contrasted histories of organelle and nuclear genomes underlying physiological diversification in a grass species. *Proceedings of the Royal Society B: Biological Sciences*, *287*(1938), 20201960.

Bond, W. J., Silander Jr, J. A., Ranaivonasy, J., & Ratsirarson, J. (2008). The antiquity of Madagascar's grasslands and the rise of C_4 grassy biomes. *Journal of Biogeography*, *35*(10), 1743–1758.

Bruvo, R., Michiels, N. K., D'Souza, T. G., & Schulenburg, H. (2004). A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level. *Molecular Ecology*, *13*(7), 2101–2106.

Cayuela, H., Boualit, L., Laporte, M., Prunier, J. G., Preiss, F., Laurent, A., Foletti, F., Clobert, J., & Jacob, G. (2019). Kin-dependent dispersal influences relatedness and genetic structuring in a lek system. *Oecologia*, *191*(1), 97–112.

Chen, J., Li, L., Milesi, P., Jansson, G., Berlin, M., Karlsson, B., Aleksic, J., Vendramin, G. G., & Lascoux, M. (2019). Genomic data provide new insights on the demographic history and the extent of recent material transfers in Norway spruce. *Evolutionary Applications*, *12*(8), 1539–1551.

Chubaty, A. M., Galpern, P., & Doctolero, S. C. (2020). The r toolbox grainscape for modelling and visualizing landscape connectivity using spatially explicit networks. *Methods in Ecology and Evolution*, *11*(4), 591–595.

Clarke, R. T., Rothery, P., & Raybould, A. F. (2002). Confidence limits for regression relationships between distance matrices: Estimating gene flow with distance. *Journal of Agricultural, Biological, and Environmental Statistics*, *7*(3), 361–372.

Cruzan, M. B., & Hendrickson, E. C. (2020). Landscape genetics of plants: Challenges and opportunities. *Plant Communications*, *1*(6), 100100.

Dellicour, S., Prunier, J. G., Piry, S., Eloy, M.-C., Bertouille, S., Licoppe, A., Frantz, A. C., & Flamand, M.-C. (2019). Landscape genetic analyses of *Cervus elaphus* and *Sus scrofa*: Comparative study and analytical developments. *Heredity*, *123*(2), 228–241.

Dixon, P. (2003). VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science*, *14*(6), 927–930.

Dodd, M. E., Silvertown, J., & Chase, M. W. (1999). Phylogenetic analysis of trait evolution and species diversity variation among angiosperm families. *Evolution*, *53*(3), 732–744.

Edwards, A. W. F. (1971). Distances between populations on the basis of gene frequencies. *Biometrics*, 873–881.

Eliades, N.-G. H., Gailing, O., Leinemann, L., Fady, B., & Finkeldey, R. (2011). High genetic diversity and significant population structure in Cedrus brevifolia Henry, a narrow endemic Mediterranean tree from Cyprus. *Plant Systematics and Evolution*, *294*(3), 185–198.

Fahrig, L. (2003). Effects of habitat fragmentation on biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, *34*(1), 487–515.

Fanamby. (2010). Plan de gestion environnementale et de sauvegarde sociale (PGESS). Etude d'impact environnemental et social (EIES) de la Nouvelle Aire Protégée Loky-Manambato.

Fernández-López, J., & Schliep, K. (2018). rWind: Download, edit and include wind data in ecological and evolutionary analysis. *Ecography*, *42*(4), 804–810.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Fontdevila, A. (2005). Hybrid genome evolution by transposition. *Cytogenetic and Genome Research*, *110*(1–4), 49–55.

Fumagalli, M. (2013). Assessing the effect of sequencing depth and sample size in population genetics inferences. *PLoS One*, *8*(11), e79667.

Galpern, P., & Manseau, M. (2013). Modelling the influence of landscape connectivity on animal distribution: A functional grain approach. *Ecography*, *36*(9), 1004–1016.

Galpern, P., Manseau, M., & Wilson, P. (2012). Grains of connectivity: Analysis at multiple spatial scales in landscape genetics. *Molecular Ecology*, *21*(16), 3996–4009.

Gardiner, L. M., Rakotoarinivo, M., Rajaovelona, L. R., & Clubbe, C. (2017). Population genetics data help to guide the conservation of palm species with small population sizes and fragmented habitats in Madagascar. *PeerJ*, *5*, e3248.

Gardner, C. J., Waeber, P. O., Razafindratsima, O. H., & Wilmé, L. (2018). Decision complacency and conservation planning. *Conservation Biology*, *32*(6), 1469–1472.

Garot, E., Joët, T., Combes, M.-C., & Lashermes, P. (2019). Genetic diversity and population divergences of an indigenous tree (*Coffea mauritiana*) in Reunion Island: Role of climatic and geographical factors. *Heredity*, *122*(6), 833–847.

Gautier, L., Ranirison, P., Nusbaumer, L., & Wohlhauser, S. (2006). Aperçu des massifs forestiers de la région Loky-Manambato. *Inventaires de La Faune et de La Flore Du Nord de Madagascar Dans La Région Loky-Manambato, Analamerana et Andavakoera. Série Sciences Biologiques*, *23*, 81–99.

Godfrey, L. R., & Crowley, B. E. (2016). Madagascar's ephemeral palaeo-grazer guild: Who ate the ancient C_4 grasses? *Proceedings of the Royal Society B: Biological Sciences*, *283*(1834), 20160360.

Goodman, S. M., Raherilalao, M. J., & Wohlhauser, S. (2018). *The terrestrial protected areas of Madagascar: Their history, description, and biota* (Association Vahatra). University of Chicago Press.

Goodman, S. M., & Wilmé, L. (2006). Inventaires de la faune et de la flore du nord de Madagascar dans la région Loky-Manambato, Analamerana et Andavakoera. *Recherche Pour Le Dévelopement*, *23*, 1–238.

Gorrilliot, O., Hong-Wa, C., Rakotonasolo, F., & Besnard, G. (2021). Microsatellite-assisted identification and comparative population genetics of Malagasy olive species (*Noronhia* spp., Oleaceae). *Botany Letters*, *168*(4), 523–535.

Goslee, S. C., & Urban, D. L. (2007). The ecodist package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software*, *22*, 1–19.

Goudie, A. S. (2018). *Human impact on the natural environment: Past, present and future* (Eighth edition). Wiley-Blackwell.

Grasty, M. R., Thompson, P. G., Hendrickson, E. C., Pheil, A. E., & Cruzan, M. B. (2020). Fine-scale habitat heterogeneity and vole runways influence seed dispersal in Plagiobothrys nothofulvus. *American Journal of Botany*, *107*(3), 413–422.

Graves, T. A., Beier, P., & Royle, J. A. (2013). Current approaches using genetic distances produce poor estimates of landscape resistance to interindividual dispersal. *Molecular Ecology*, *22*(15), 3888–3903.

Green, G. M., & Sussman, R. W. (1990). Deforestation history of the eastern rain forests of Madagascar from satellite images. *Science*, *248*(4952), 212–215.

Hackel, J., Vorontsova, M. S., Nanjarisoa, O. P., Hall, R. C., Razanatsoa, J., Malakasi, P., & Besnard, G. (2018). Grass diversification in Madagascar: In situ radiation of two large C_3

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

shade clades and support for a Miocene to Pliocene origin of C_4 grassy biomes. *Journal of Biogeography*, *45*(4), 750–761.

Haddad, N. M., Brudvig, L. A., Clobert, J., Davies, K. F., Gonzalez, A., Holt, R. D., Lovejoy, T. E., Sexton, J. O., Austin, M. P., & Collins, C. D. (2015). Habitat fragmentation and its lasting impact on Earth's ecosystems. *Science Advances*, *1*(2), e1500052.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., & Loveland, T. R. (2013). High-resolution global maps of 21st-century forest cover change. *Science*, *342*(6160), 850–853.

Harper, G. J., Steininger, M. K., Tucker, C. J., Juhn, D., & Hawkins, F. (2007). Fifty years of deforestation and forest fragmentation in Madagascar. *Environmental Conservation*, *34*(4), 325–333.

Heller, R., Nursyifa, C., Garcia Erill, G., Salmona, J., Chikhi, L., Meisner, J., Korneliussen, T. S., & Albrechtsen, A. (2021). A reference-free approach to analyze non-model RADseq data using standard next generation sequencing toolkits. *Molecular Ecology Resources*, *21*(4), 1085–1097.

Holderegger, R., Buehler, D., Gugerli, F., & Manel, S. (2010). Landscape genetics of plants. *Trends in Plant Science*, *15*(12), 675–683.

Hong-Wa, C. (2016). *A taxonomic revision of the genus* Noronhia *Stadtm. Ex Thouars (Oleaceae) in Madagascar and the Comoro Islands* (M. Callmander & P. P. Lowry II, Eds.; *Boissiera*, Vol. 70). Conservatoire et Jardin Botanique de la ville de Genève.

Hong-Wa, C., & Besnard, G. (2014). Species limits and diversification in the Madagascar olive (*Noronhia*, Oleaceae). *Botanical Journal of the Linnean Society*, *174*(1), 141–161.

Isabel, N., Holliday, J. A., & Aitken, S. N. (2020). Forest genomics: Advancing climate adaptation, forest health, productivity, and conservation. *Evolutionary Applications*, *13*(1), 3–10.

Joseph, G. S., & Seymour, C. L. (2020). Madagascan highlands: Originally woodland and forest containing endemic grasses, not grazing-adapted grassland. *Proceedings of the Royal Society B: Biological Sciences*, *287*(1937), 20201956.

Joseph, G. S., & Seymour, C. L. (2021). The unlikely 'antiquity of Madagascar's grasslands': Disproportionately forest□limited endemic fauna support anthropogenic transformation from woodland. *Journal of Biogeography*, *48*(8), 2111–2115.

Kamvar, Z. N., Brooks, J. C., & Grünwald, N. J. (2015). Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Frontiers in Genetics*, *6*, 208.

Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, *2*, e281.

Keller, D., & Holderegger, R. (2013). Damselflies use different movement strategies for short- and long-distance dispersal. *Insect Conservation and Diversity*, *6*(5), 590–597.

Kimura, M. (1983). *The neutral theory of molecular evolution*. Cambridge University Press.

Kling, M. M., & Ackerly, D. D. (2021). Global wind patterns shape genetic differentiation, asymmetric gene flow, and genetic diversity in trees. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(17), e2017317118.

Kohrn, B. F., Persinger, J. M., & Cruzan, M. B. (2017). An efficient pipeline to generate data for studies in plastid population genomics and phylogeography. *Applications in Plant Sciences*, *5*(11), 1700053.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of next generation sequencing data. *BMC Bioinformatics*, *15*(1), 356.

Kurtz, S., Phillippy, A., Delcher, A. L., Smoot, M., Shumway, M., Antonescu, C., & Salzberg, S. L. (2004). Versatile and open software for comparing large genomes. *Genome Biology*, *5*(2), R12.

Landguth, E. L., Cushman, S. A., Schwartz, M. K., McKelvey, K. S., Murphy, M., & Luikart, G. (2010). Quantifying the lag time to detect barriers in landscape genetics. *Molecular Ecology*, *19*(19), 4179–4191.

Lanes, É. C., Pope, N. S., Alves, R., Carvalho Filho, N. M., Giannini, T. C., Giulietti, A. M., Imperatriz-Fonseca, V. L., Monteiro, W., Oliveira, G., & Silva, A. R. (2018). Landscape genomic conservation assessment of a narrow-endemic and a widespread morning glory from Amazonian savannas. *Frontiers in Plant Science*, *9*, 532.

Lebuhn, G., Droege, S., Connor, E. F., Gemmill-Herren, B., Potts, S. G., Minckley, R. L., Jean, R. P., Kula, E., Roubik, D. W., & Wright, K. W. (2015). Evidence-based conservation: Reply to Tepedino et al. *Conservation Biology*, *29*(1), 283–285.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *ArXiv Preprint ArXiv:1303.3997*.

Lindenmayer, D. B., & Fischer, J. (2013). *Habitat fragmentation and landscape change: An ecological and conservation synthesis*. Island Press.

Ma, T., Wang, K., Hu, Q., Xi, Z., Wan, D., Wang, Q., Feng, J., Jiang, D., Ahani, H., & Abbott, R. J. (2018). Ancient polymorphisms and divergence hitchhiking contribute to genomic islands of divergence within a poplar species complex. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(2), E236–E243.

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research*, *27*(2), 209–220.

McRae, B. H. (2006). Isolation by resistance. *Evolution*, *60*(8), 1551–1561.

McRae, B. H., & Beier, P. (2007). Circuit theory predicts gene flow in plant and animal populations. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(50), 19885–19890.

Meisner, J., & Albrechtsen, A. (2018). Inferring population structure and admixture proportions in low-depth NGS data. *Genetics*, *210*(2), 719–731.

Mona, S., Ray, N., Arenas, M., & Excoffier, L. (2014). Genetic consequences of habitat fragmentation during a range expansion. *Heredity*, *112*(3), 291–299.

Morelli, T. L., Smith, A. B., Mancini, A. N., Balko, E. A., Borgerson, C., Dolch, R., Farris, Z., Federman, S., Golden, C. D., & Holmes, S. M. (2020). The fate of Madagascar's rainforest habitat. *Nature Climate Change*, *10*(1), 89–96.

Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, *403*(6772), 853–858.

Nater, A., Greminger, M. P., Arora, N., Schaik, C. P., Goossens, B., Singleton, I., Verschoor, E. J., Warren, K. S., & Krützen, M. (2015). Reconstructing the demographic history of orang-utans using Approximate Bayesian Computation. *Molecular Ecology*, *24*(2), 310–327.

Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America*, *70*(12), 3321–3323.

Nei, M. (1987). *Molecular Evolutionary Genetics*. Columbia University Press.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Nielsen, R., Korneliussen, T., Albrechtsen, A., Li, Y., & Wang, J. (2012). SNP calling, genotype calling, and sample allele frequency estimation from new-generation sequencing data. *PLoS One*, *7*(7), e37558.

Oden, N. L., & Sokal, R. R. (1986). Directional autocorrelation: An extension of spatial correlograms to two dimensions. *Systematic Zoology*, *35*(4), 608–617.

Olofsson, J. K., Dunning, L. T., Lundgren, M. R., Barton, H. J., Thompson, J., Cuff, N., Ariyarathne, M., Yakandawala, D., Sotelo, G., Zeng, K., Osborne, C. P., Nosil, P., & Christin, P.-A. (2019). Population-specific selection on standing variation generated by lateral gene transfers in a grass. *Current Biology*, *29*(22), 3921–3927.

Paris, J. R., Stevens, J. R., & Catchen, J. M. (2017). Lost in parameter space: A road map for stacks. *Methods in Ecology and Evolution*, *8*(10), 1360–1373.

Payseur, B. A., & Rieseberg, L. H. (2016). A genomic perspective on hybridization and speciation. *Molecular Ecology*, *25*(11), 2337–2360.

Pedersen, C.-E. T., Albrechtsen, A., Etter, P. D., Johnson, E. A., Orlando, L., Chikhi, L., Siegismund, H. R., & Heller, R. (2018). A southern African origin and cryptic structure in the highly mobile plains zebra. *Nature Ecology & Evolution*, *1*(3), 491–498.

Peterman, W. E. (2018). ResistanceGA: An R package for the optimization of resistance surfaces using genetic algorithms. *Methods in Ecology and Evolution*, *9*(6), 1638–1647.

Peterman, W. E., & Pope, N. S. (2021). The use and misuse of regression models in landscape genetic analyses. *Molecular Ecology*, *30*(1), 37–47.

Peterman, W. E., Winiarski, K. J., Moore, C. E., Carvalho, C. da S., Gilbert, A. L., & Spear, S. F. (2019). A comparison of popular approaches to optimize landscape resistance surfaces. *Landscape Ecology*, *34*(9), 2197–2208.

Petit, R. J., & Hampe, A. (2006). Some evolutionary consequences of being a tree. *Annual Review of Ecology, Evolution, and Systematics*, *37*, 187–214.

Pillon, Y., Hopkins, H. C., Rigault, F., Jaffré, T., & Stacy, E. A. (2014). Cryptic adaptive radiation in tropical forest trees in New Caledonia. *New Phytologist*, *202*(2), 521–530.

Plomion, C., Aury, J.-M., Amselem, J., Leroy, T., Murat, F., Duplessis, S., Faye, S., Francillonne, N., Labadie, K., & Le Provost, G. (2018). Oak genome reveals facets of long lifespan. *Nature Plants*, *4*(7), 440–452.

Pressey, R. L., Cabeza, M., Watts, M. E., Cowling, R. M., & Wilson, K. A. (2007). Conservation planning in a changing world. *Trends in Ecology & Evolution*, *22*(11), 583–592.

Prunier, J. G., Colyn, M., Legendre, X., & Flamand, M.-C. (2017). Regression commonality analyses on hierarchical genetic distances. *Ecography*, *40*(12), 1412–1425.

Quéméré, E., Amelot, X., Pierson, J., Crouau-Roy, B., & Chikhi, L. (2012). Genetic data suggest a natural prehuman origin of open habitats in northern Madagascar and question the deforestation narrative in this region. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(32), 13028–13033.

Quéméré, E., Crouau-Roy, B., Rabarivola, C., Louis, E. E., & Chikhi, L. (2010). Landscape genetics of an endangered lemur (*Propithecus tattersalli*) within its entire fragmented range. *Molecular Ecology*, *19*(8), 1606–1621.

Rakotoarisoa, J.-E., Raheriarisena, M., & Goodman, S. M. (2013a). A phylogeographic study of the endemic rodent *Eliurus carletoni* (Rodentia: Nesomyinae) in an ecological transition zone of northern Madagascar. *Journal of Heredity*, *104*(1), 23–35.

Rakotoarisoa, J.-E., Raheriarisena, M., & Goodman, S. M. (2013b). Late Quaternary climatic vegetational shifts in an ecological transition zone of northern Madagascar: Insights from

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

genetic analyses of two endemic rodent species. *Journal of Evolutionary Biology*, *26*(5), 1019–1034.

Rakotondravony, H. A. (2006). Patterns de la diversité des reptiles et amphibiens de la région de Loky-Manambato. *Inventaires de La Faune et de La Flore Du Nord de Madagascar Dans La Région Loky-Manambato, Analamerana et Andavakoera. Série Sciences Biologiques*, *23*, 101–148.

Rakotondravony, H. A. (2009). Aspects de la conservation des reptiles et des amphibiens dans la région de Daraina. *Madagascar Conservation & Development*, *1*(1), 15–18.

Raxworthy, C. J., & Nussbaum, R. A. (1995). Systematics, speciation and biogeography of the dwarf chameleons (*Brookesia*; Reptilia, Squamata, Chamaeleontidae) of northern Madagascar. *Journal of Zoology*, *235*(4), 525–558.

Reynolds, J., Weir, B. S., & Cockerham, C. C. (1983). Estimation of the coancestry coefficient: Basis for a short-term genetic distance. *Genetics*, *105*(3), 767–779.

Rochette, N. C., Rivera-Colón, A. G., & Catchen, J. M. (2019). Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology*, *28*(21), 4737–4754.

Salmona, J., Heller, R., Lascoux, M., & Shafer, A. (2017). Inferring demographic history using genomic data. In O. Rajora (Ed.), *Population Genomics* (pp. 511–537). Springer.

Salmona, J., Heller, R., Quéméré, E., & Chikhi, L. (2017). Climate change and human colonization triggered habitat loss and fragmentation in Madagascar. *Molecular Ecology*, *26*(19), 5203–5222.

Salmona, J., Olofsson, J. K., Hong-Wa, C., Razanatsoa, J., Rakotonasolo, F., Ralimanana, H., Randriamboavonjy, T., Suescun, U., Vorontsova, M. S., & Besnard, G. (2020). Late Miocene origin and recent population collapse of the Malagasy savanna olive tree (*Noronhia lowryi*). *Biological Journal of the Linnean Society*, *129*(1), 227–243.

Schmieder, R., & Edwards, R. (2011). Fast identification and removal of sequence contamination from genomic and metagenomic datasets. *PLoS One*, *6*(3), e17288.

Schüßler, D., Mantilla-Contreras, J., Stadtmann, R., Ratsimbazafy, J. H., & Radespiel, U. (2020). Identification of crucial stepping stone habitats for biodiversity conservation in northeastern Madagascar using remote sensing and comparative predictive modeling. *Biodiversity and Conservation*, *29*(7), 2161–2184.

Scotti, I., González-Martínez, S. C., Budde, K. B., & Lalagüe, H. (2016). Fifty years of genetic studies: What to make of the large amounts of variation found within populations? *Annals of Forest Science*, *73*(1), 69–75.

Sgarlata, G. M., Salmona, J., Aleixo-Pais, I., Rakotonanahary, A., Sousa, A. P., Kun-Rodrigues, C., Ralantoharijaona, T., Jan, F., Zaranaina, R., Rasolondraibe, E., Zaonarivelo, J. R., Andriaholinirina, N., & Chikhi, L. (2018). Genetic differentiation and demographic history of the northern rufous mouse Lemur (*Microcebus tavaratra*) across a fragmented landscape in northern Madagascar. *International Journal of Primatology*, *39*(1), 65–89.

Sgarlata, G. M., Salmona, J., Le Pors, B., Rasolondraibe, E., Jan, F., Ralantoharijaona, T., Rakotonanahary, A., Randriamaroson, J., Marques, A. J., Aleixo-Pais, I., de Zoeten, T., Ousseni, D. S. A., Knoop, S. B., Teixeira, H., Gabillaud, V., Miller, A., Ibouroi, M. T., Rasoloharijaona, S., Zaonarivelo, J. R., … Chikhi, L. (2019). Genetic and morphological diversity of mouse lemurs (*Microcebus* spp.) in northern Madagascar: The discovery of a putative new species? *American Journal of Primatology*, *81*(12), e23070.

bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Shirk, A. J., Landguth, E. L., & Cushman, S. A. (2017). A comparison of individual-based genetic distance metrics for landscape genetics. *Molecular Ecology Resources*, *17*(6), 1308–1317.

Shirk, A. J., Landguth, E. L., & Cushman, S. A. (2018). A comparison of regression methods for model selection in individual-based landscape genetic analysis. *Molecular Ecology Resources*, *18*(1), 55–67.

Skotte, L., Korneliussen, T. S., & Albrechtsen, A. (2013). Estimating individual admixture proportions from next generation sequencing data. *Genetics*, *195*(3), 693–702.

Slatkin, M. (1993). Isolation by distance in equilibrium and non-equilibrium populations. *Evolution*, *47*(1), 264–279.

Sokal, R. R. (1986). Spatial data analysis and historical processes. *Data Analysis and Informatics IV*, 29–43.

Solofondranohatra, C. L., Vorontsova, M. S., Hackel, J., Besnard, G., Cable, S., Williams, J., Jeannoda, V., & Lehmann, C. E. (2018). Grass functional traits differentiate forest and savanna in the Madagascar central highlands. *Frontiers in Ecology and Evolution*, *6*, 184.

Štorchová, H., & Olson, M. S. (2004). Comparison between mitochondrial and chloroplast DNA variation in the native range of *Silene vulgaris*. *Molecular Ecology*, *13*(10), 2909–2919.

Sutherland, W. J., Pullin, A. S., Dolman, P. M., & Knight, T. M. (2004). The need for evidence-based conservation. *Trends in Ecology & Evolution*, *19*(6), 305–308.

Tang, Q., Fung, T., & Rheindt, F. E. (2020). ResDisMapper: An r package for fine-scale mapping of resistance to dispersal. *Molecular Ecology Resources*, *20*(3), 819–831.

Van de Paer, C. (2017). Structural diversity and contrasted evolution of cytoplasmic genomes in flowering plants: A phylogenomic approach in Oleaceae. *PhD thesis, University of Toulouse III-Paul Sabatier* [PhD Thesis, Université Toulouse III – Paul Sabatier]. PhD thesis, University of Toulouse III-Paul Sabatier.

Van de Paer, C., Bouchez, O., & Besnard, G. (2018). Prospects on the evolutionary mitogenomics of plants: A case study on the olive family (Oleaceae). *Molecular Ecology Resources*, *18*(3), 407–423.

van Etten, J. (2012). R package gdistance: Distances and routes on geographical grids. *Journal of Statistical Software*, *76*(13), 1–21.

van Strien, M. J., Holderegger, R., & van Heck, H. J. (2015). Isolation-by-distance in landscapes: Considerations for landscape genetics. *Heredity*, *114*(1), 27–37.

van Strien, M. J., Keller, D., & Holderegger, R. (2012). A new analytical approach to landscape genetic modelling: Least-cost transect analysis and linear mixed models. *Molecular Ecology*, *21*(16), 4010–4023.

Vieilledent, G., Grinand, C., Rakotomalala, F. A., Ranaivosoa, R., Rakotoarijaona, J.-R., Allnutt, T. F., & Achard, F. (2018). Combining global tree cover loss data with historical national forest cover maps to look at six decades of deforestation and forest fragmentation in Madagascar. *Biological Conservation*, *222*, 189–197.

Vorontsova, M. S., Besnard, G., Forest, F., Malakasi, P., Moat, J., Clayton, W. D., Ficinski, P., Savva, G. M., Nanjarisoa, O. P., & Razanatsoa, J. (2016). Madagascar's grasses and grasslands: Anthropogenic or natural? *Proceedings of the Royal Society B: Biological Sciences*, *283*(1823), 20152262.

Wang, Y., Lu, J., Yu, J., Gibbs, R. A., & Yu, F. (2013). An integrative variant analysis pipeline for accurate genotype/haplotype inference in population NGS data. *Genome Research*, *23*(5), 833–842.

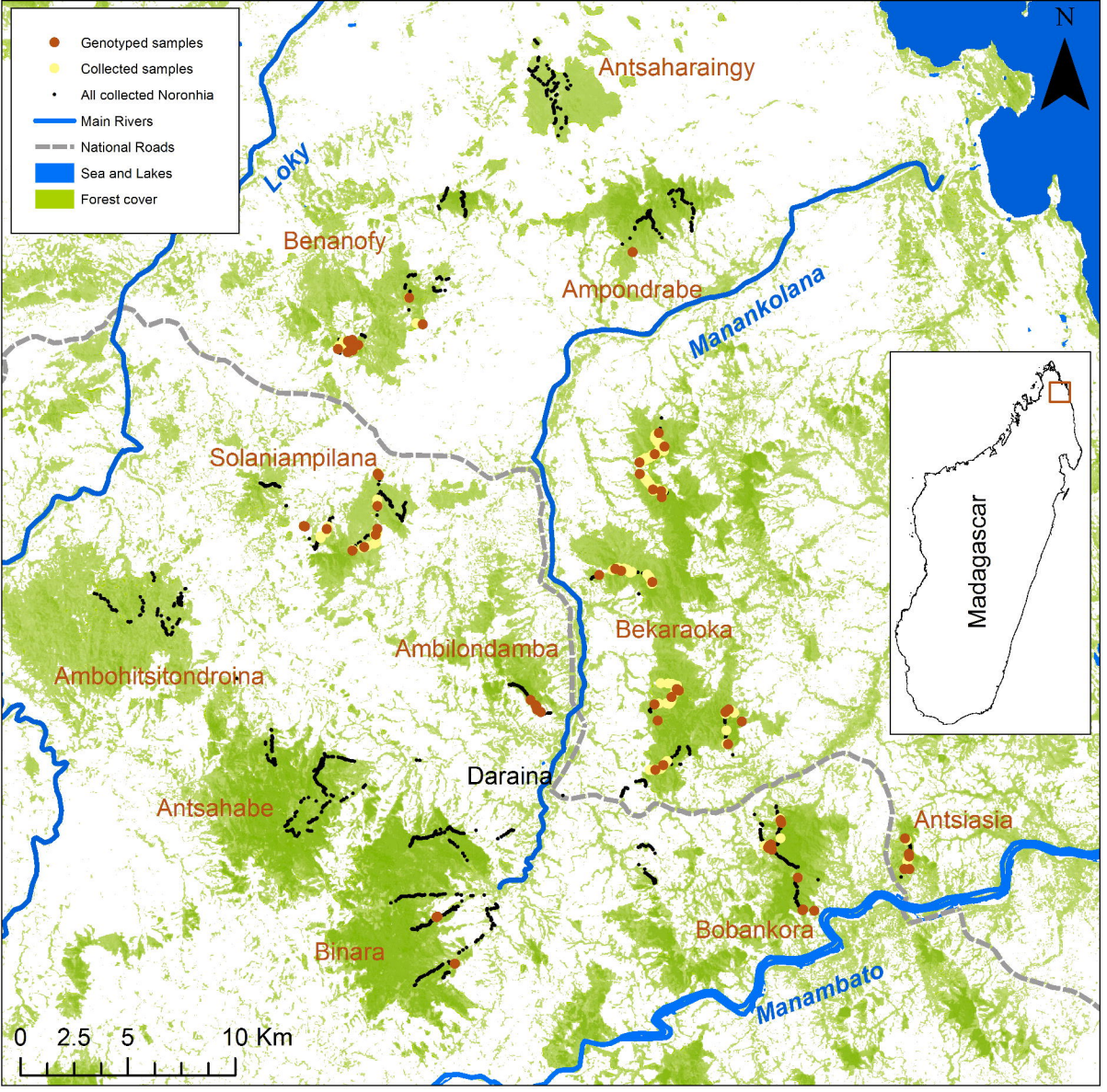
bioRxiv preprint doi: https://doi.org/10.1101/2020.11.25.394544; this version posted October 13, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

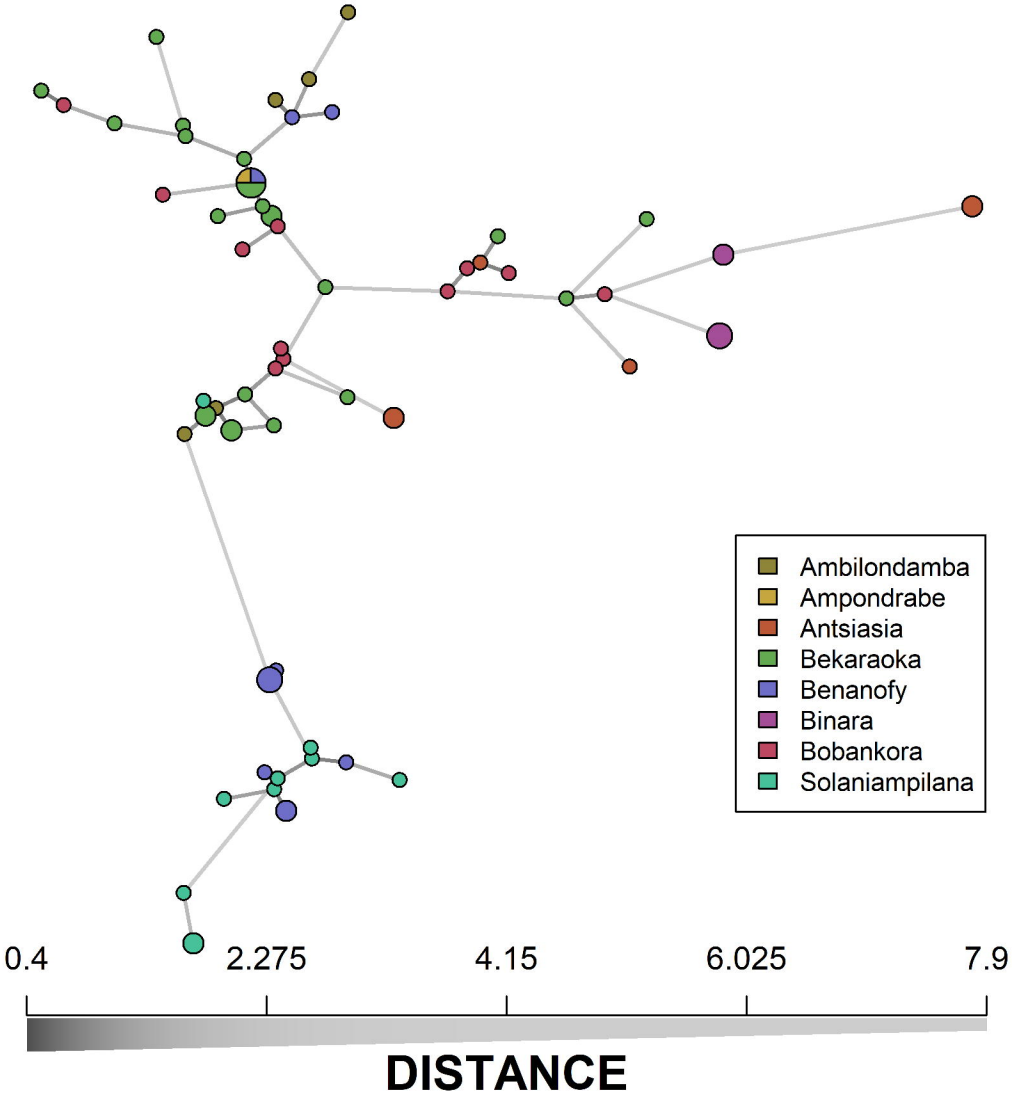
Warmuth, V. M., & Ellegren, H. (2019). Genotype-free estimation of allele frequencies reduces bias and improves demographic inference from RADSeq data. *Molecular Ecology Resources*, *19*(3), 586–596.

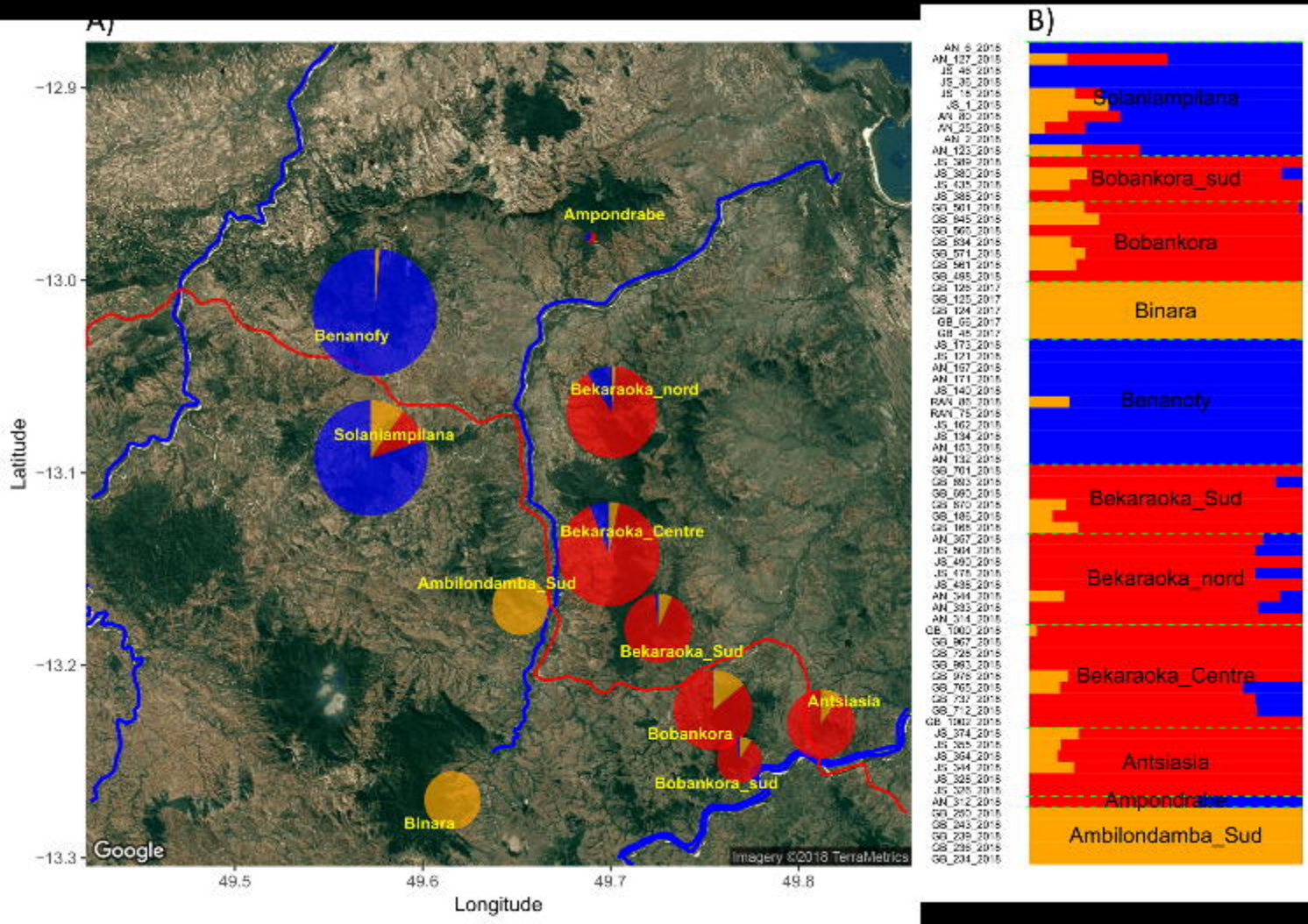
Weiß, C. L., Pais, M., Cano, L. M., Kamoun, S., & Burbano, H. A. (2018). nQuire: A statistical framework for ploidy estimation using next generation sequencing. *BMC Bioinformatics*, *19*(1), 1–8.

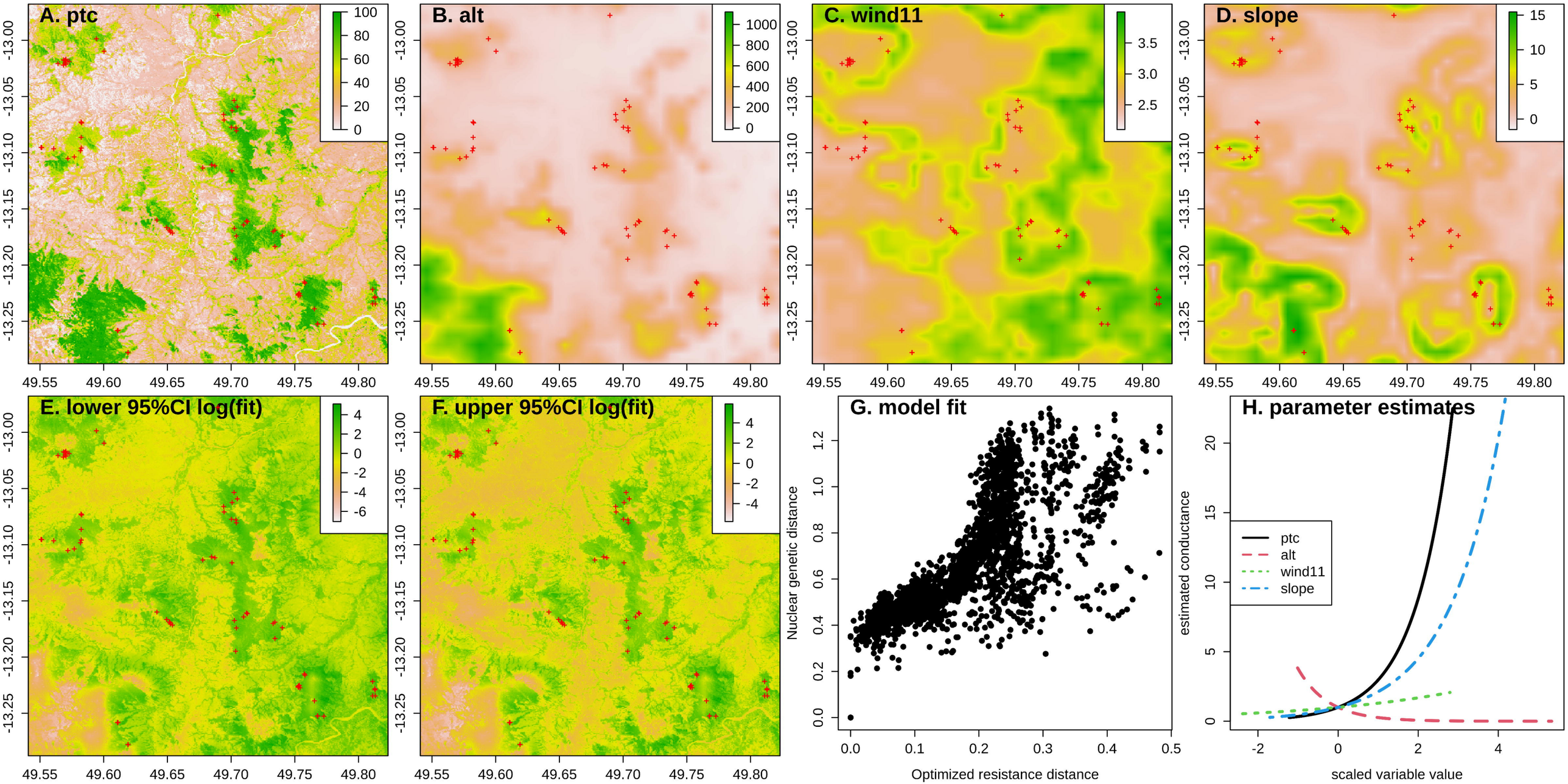
Winiarski, K. J., Peterman, W. E., & McGarigal, K. (2020). Evaluation of the R package 'resistancega': A promising approach towards the accurate optimization of landscape resistance surfaces. *Molecular Ecology Resources*, *20*(6), 1583–1596.

Wright, S. (1943). Isolation by distance. *Genetics*, *28*(2), 114–138.


Genotyped samples
Collected samples
All collected Noronhia
Main Rivers
National Roads
Sea and Lakes
Forest cover
N
Antsaharaingy
Loky
Benanofy
Ampondrabe
Manankolana
Solaniampilana
Madagascar
Bekaraoka
Ambilondamba
Ambohitsitondroina
Daraina
Antsahabe
Antsiasia
Bobankora
Binara
Manambato
0 2.5 5 10 Km


Ambilondamba
Ampondrabe
Antsiasia
Bekaraoka
Benanofy
Binara
Bobankora
Solaniampilana
0.4
2.275
4.15
6.025
7.9
DISTANCE


A)
Ampondrabe
Benanofy
Solaniampilana
Bekaraoka_nord
Bekaraoka_Centre
Ambilondamba_Sud
Bekaraoka_Sud
Antsiasia
Bobankora
Bobankora_sud
Binara
Google
Imagery ©2018 TerraMetrics
Latitude
Longitude
-12.9
-13.0
-13.1
-13.2
-13.3
49.5
49.6
49.7
49.8
B)
Solaniampilana
Bobankora_sud
Bobankora
Binara
Benanofy
Bekaraoka_Sud
Bekaraoka_nord
Bekaraoka_Centre
Antsiasia
Ampondrabe
Ambilondamba_Sud


A. ptc
B. alt
C. wind11
D. slope
E. lower 95%CI log(fit)
F. upper 95%CI log(fit)
G. model fit
H. parameter estimates
Nuclear genetic distance
Optimized resistance distance
estimated conductance
scaled variable value
ptc
alt
wind11
slope