Retracing the center of origin and evolutionary history of nutmeg Myristica fragrans, an emblematic spice tree species

Jakty Kusuma[1], Nora Scarcelli[1], Miguel Navascués[2], M. Couderc[1], Pierre Gerard[3], and Jérôme Duminil[4]

[1]IRD
[2]INRAE
[3]AgroParisTech
[4]Universite Libre de Bruxelles Departement de Biologie des Organismes

August 4, 2023

Abstract

The identification of cultivated plant centers of origin is of growing interest since the pioneering work of Nikolaï Vavilov. Cultivated species generally show a high level of genetic diversity in their center of origin, thus identifying these areas would have important implications for the conservation and sustainable management of their genetic resources. We aim to identify the center of origin and to infer recent evolutionary history of one of the most emblematic spice tree species ¬¬– nutmeg (Myristica fragrans; Myristicaceae). The species is thought to originate from the Banda islands (Southern Moluccas archipelago, Indonesia), but this has never been confirmed. We characterized the genetic diversity of this species across the Moluccas archipelago using recently developed nuclear microsatellites markers and whole plastid genome sequences. We found two main intra-specific genetic clusters: one in South Moluccas (Banda and Ambon islands) and another in North Moluccas (Ternate, Tidore and Bacan islands). The latter cluster showed higher genetic diversity than the South Moluccas cluster. We then inferred the demographic history of nutmeg and evaluated different evolutionary scenarios using Approximate Bayesian Computation (ABC) analyses. Populations from South Moluccas show patterns of a recent bottleneck, whereas North Molucca populations did not. Moreover, South Moluccas populations are inferred as ancestral, with subsequent population migration to the North Moluccas during the late Pleistocene to early Holocene period. These results have strong implications for conservation strategies of nutmeg.

INTRODUCTION

Inferring the center of domestication of a cultivated species is challenging. The development of agriculture occurred independently in different parts of the world from 11,000 to 5,000 years ago. Alphonse de Candolle was one of the first to question the geographical origin of cultivated plants using an integrated scientific approach . He synthesized data from different disciplines (botany, plant biogeography, archaeology, paleontology, history, linguistics, ethnology) to propose hypotheses about the geographic origin of 247 cultivated species. This work was later extended by Nikolaï Vavilov, who proposed main centers of domestication for cultivated plants . Vavilov's work has since been revisted and updated by others . Indeed, the centers of plant domestication *sensu* Vavilov was subsequently confirmed by a number of archaeological findings . Furthermore, additional progress has been made on this topic through the use of genomic data to study the origin of cultivated plants . However, despite these recent progress in research, the origins of many cultivated species still remain poorly understood.

Inferring the place of origin and biogeographic history of a cultivated species is important to develop robust conservation and sustainable management strategies. Generally, the genetic diversity within a species is not homogeneous across its distribution range. The center of origin of cultivated plants typically presents

higher levels of genetic diversity than their introduced populations as species propagation practices outside of its original distribution range often results in genetic bottlenecks . Identifying the center of origin of a target species is thus of utmost importance to develop conservation strategies that ensure the safeguarding of potentially valuable genetic resources. However, it is worth noting that this expectation can be obscured by gene flow with closely-related species, a relatively frequent phenomenon in plants . Indeed, during their propagation outside of their native range, cultivated species may enter into contact with closely-related species, creating new possibilities for interbreeding and increase of genetic diversity in the domesticated pool. This process is known to contribute to species adaptation to new environmental conditions during its propagation outside of its native range .

Genetic studies based on extensive sampling across a species' distribution can contribute to tracing back the biogeographical history of that species . For non-model organisms with a few genomic resources available, two types of genetic markers are commonly used: nuclear microsatellites (nSSR) and chloroplast DNA (cpDNA). These markers are used to infer past demographic events and to test different demographic scenarios with temporal calibrations .

Nutmeg, *Myristica fragrans* , is one of the most emblematic spice tree species. The exact location of the center of origin of the species is controversial, some authors hypothesize it to comprise a whole archipelago in Indonesia, the Moluccas (Maluku in Indonesian, also known as 'the Spice Islands'), whereas others mentioned the Banda islands only–a group of small islands in South Moluccas (Figure 1). The trading history of spice species has fascinated many, as evident in the extensive literature found on this topic. The Chinese were among the first involved in the trading of these species, with the earliest description of Banda islands and the nutmeg in the XIVth century . A short entry from Zhao Rugua's diary (c. 1225) seems to be the earliest written mention of the nutmeg. Unfortunately, the vocabulary used in this old text was insufficient for a definite identification of the species and region involved . The Portuguese were the first European explorers arriving in the region in search of the Spice Islands in the 16th century. They cite Banda islands as the center of production of nutmeg, suggesting that the cultivation of this species mainly took place on these islands during that time. Yet, Rumphius in his *Herbarii Amboinensis*explained nutmeg is distributed throughout the entire Moluccas: "*The Nutmeg Tree is much further distributed than the Clove tree, to wit, throughout almost the entire Moluccas, though its true see is the Province of Banda* , . . . " (though 'its true see' is in Banda archipelago (English translation by E.M. Beekman of the latin original text '*verus tamen ejus locas fixus est* Banda'). The words of Rumphius are not necessarily easy to interpret, and we might understand either that the native range of the species is the Banda archipelago (which has been repeatedly cited as such thereafter) or, that Banda islands represent the place where the species was cultivated at that time.

What is known with certainty is that Banda islands were the most important places of nutmeg cultivation during the Indonesian Dutch period (ca. 1600-1800). At this time, the Dutch East India Company (Vereenigde Oostindische Compagnie, VOC) applied a strong monopolistic policy to ensure its control over this important economic resource . Other European nations, such as France or England, did their best to circumvent the Dutch monopoly by getting planting material from Banda islands to propagate the tree in their tropical colonies . This part of history probably strengthened the general view that nutmeg's place of origin was Banda islands. Nowadays, the species is widely cultivated in different places in Indonesia and other tropical regions of the world.

Few archaeobotanical studies were conducted in Indonesia. To the best of our knowledge, only one study has identified historical traces of nutmeg consumption in Indonesia before European arrival . This study was held in one of the Banda islands, Pulau Ay (in English known as Ai Island), and traced the consumption to the Neolithic period (ca. 3500–2000 B.P.). Traces of consumption do not necessarily imply that nutmeg was cultivated at that time, but nevertheless this piece of evidence indicates the presence of the species in the Banda islands since more than 1000 years ago.

To address this gap in knowledge on the ancestral distribution of nutmeg, we applied a population genetic approach to improve our understanding of the biogeographical history of nutmeg in the Moluccas. More specifically, we tested whether the species originated from the Banda islands, or from the wider Moluccas.

We sampled the species across several Moluccas islands from North to South, and characterized its genetic diversity using a set of recently developed nSSR and whole plastid genomes.

MATERIALS AND METHODS

Study area: the Moluccas region

The Moluccas (Figure 1) constitute a natural biogeographical unit within the Wallacea region in the eastern part of Indonesia. They include two island groups: the North Moluccas (Maluku Utara province), which is part of a volcanic arc, and the South Moluccas (Maluku province) located on the Australasian continental plate.

Plant material

A total of 393 individuals were collected at five different sites (Ambon, Banda islands, Ternate, Tidore, and Bacan) covering the geographic extent of the Moluccas archipelago from North to South (Figure 1). The Banda islands are a group of ten volcanic islands: we sampled four of them (hereafter "Banda" to refer to these sampling localities). Leaves of individuals were collected and stored in silica gel to dry rapidly for preservation until DNA extraction. Additional information were taken for the collecting sites: the name of the locality, geographic coordinates, and tree circumference. We refer to each island as a "population", and the different collecting sites in these islands are referred to as "subpopulations" (for details on the number of subpopulations per population and the number of individuals per population see Supplementary Table S1).

DNA extraction, nSSR genotyping and cpDNA sequencing

DNA extraction . Total genomic DNA was extracted using the MATAB protocol and chloroform separation with slight modifications following Mariac *et al* . . We checked DNA quality and quantity using a NanoDrop^{TM} One spectrophotometer (Thermo Scientific^{TM}). Extracted DNA was diluted in 50 μL Tris-EDTA (TE) buffer (1X) pH 8.0 and stored at -20°C.

nSSR genotyping . All individuals were genotyped using 14 polymorphic nuclear nSSR markers, with PCR conditions as described in Kusuma *et al.* . As one of the primer pair (*Myr* 42 locus) allowed for the amplification of two different genomic regions, corresponding information was split into two, namely *Myr* 42a and*Myr* 42b (see Kusuma *et al* ., 2020 for details). From this, we obtained our final dataset of 15 diploid genotypes for each individual. Genotyping was done on an ABI 3500 XL (Applied Biosystems) at CIRAD (Montpellier, France) using 1 μL of PCR product, 12 μL of Hi-Di Formamide (Applied Biosystems) and 0.3 μL GeneScan 500 LIZ dye Size Standard (Applied Biosystems). Genotypes were analyzed for each locus using the automatic procedure implemented in Geneious Prime$^{®}$ v.2022.2.2 and manually corrected.

cpDNA sequencing . We selected three to five individuals per subpopulation for genomic library construction ($N = 90$; see Supplementary File 1 for details) as well as four additional individuals of the sister species *Myristica argentea* (commonly known as the Papuan nutmeg, from Fakfak, Papua, Indonesia) that has been used as an outgroup in phylogenetic network analyses ($N = 94$ in total; see Supplementary File 1 for details). The protocols for library preparation, chloroplast enrichment, multiplexing, and sequencing are described in Mariac *et al* . . All sequenced individuals were analyzed as described in Scarcelli *et al.* . In short, data were demultiplexed and cleaned using the scripts DEMULTADAPT (*https://github.com/Maillol/demultadapt*) and CUTADAPT v.1.2.1 . Reads were then filtered using a PERL script (Supplementary File 2). After discarding low quality reads, mapping was performed with BWA v.0.7.12 using the *M. fragrans* cpDNA genome as reference. Unmapped reads were discarded and mapped reads were sorted using SAMTOOLS v.1.9 . Then VARSCAN v.2.3.7 was used to call Single Nucleotide Polymorphisms (SNPs). All SNPs were filtered using VCFtools , by deleting all the indels, and all non-biallelic SNPs. Then, we calculated the mean depth per individual, mean depth per SNP, percentage of missing data per individual, and percentage of missing data per SNP using VCFtools . We provide the command lines and scripts to prepare sequence data, mapping and calling of SNPs in Supplementary File 3.

Genetic diversity and population structure of M. fragrans

3

at nSSR loci

The distribution of genetic diversity in *M. fragrans* ($N = 393$ individuals) was characterized at the population level ($N_{POP} = 5$) by estimating the following indices using SPAGeDi v.1.5 : the number of alleles (N_A), observed and expected heterozygosity (H_O and H_E respectively), rarefied allelic richness (AR), and inbreeding coefficient (F_{IS}).

We then calculated the F_{ST} among populations and among sub-populations within each population using the *hierfstat* R package (Goudet, 2005) in R (R Core Team, 2013). Moreover, we characterized genetic structure using a Bayesian clustering approach as implemented in STRUCTURE v.x.x software with the following parameters: admixture model, sampling location as prior, independent allele frequencies. The number of clusters was set from K = 1 to 10, with 10 iterations for each K. We ran the analysis with 20,000 burn-in periods and 100,000 MCMC replicates. The value of K that best explained the data was estimated using STRUCTURE HARVESTER . We then summarized STRUCTURE results using CLUMPP and projected the barplot using CLUMPAK .

Genetic diversity and population structure of M. fragrans

at cpDNA loci

For cpDNA, we used the SNPs dataset obtained for the 94 individuals (90 *M. fragrans* , 4 *M. argentea*) to describe the geographic distribution of cpDNA genetic diversity. We estimated the haplotype number (H), number of segregating sites (S), and nucleotide diversity (π) using DnaSP v.6.12.03 . To analyze the genetic variation among groups and populations, we assigned the populations on each island to two different regions: Ambon and Banda were considered as a Southern Moluccas group and Ternate, Tidore and Bacan as a Northern Moluccas group. We used a molecular variance analysis (AMOVA) approach in ARLEQUIN v.3.5 assessing significant differences with 1,000 permutations.

Haplotype network reconstruction

We assigned cpDNA SNPs to different haplotypes using the *pegas* R package v.1.1 . We constructed the haplotype network using the maximum parsimonious tree method available in PopART v.1.7 , with a 95% statistical parsimony criterion. To test for phylogeographic signal, we calculated the N_{ST}, and the G_{ST} , using Permut v.1.2.1 applying 1,000 permutation tests .

Inference of population size change in M. fragrans

Past demographic events in *M. fragrans* populations from the Moluccas were inferred using Approximate Bayesian Computation (ABC) analyses. ABC approaches estimate posterior probability distributions substituting the calculation of likelihood by comparing observed data with simulated data . Using the nSSR dataset, we characterized the demographic evolution through time for each population using the program DIYABCskylineplot . Coalescent simulations and the computation of prior parameters and summary statistics were done using command line in DIYABC v.2.0 . We set the log-uniform priors for scaled effective population size (ϑ) ranging from 1 to 1,000,000 for each population. Time parameters were randomly drawn at the interval [1;2000], and uniform prior in the interval [0;1] for the generalized stepwise mutation model. We performed 100,000 simulations for each population.

To estimate the robustness of past demographic event inferences, we further explored the likelihood of demographic changes through time using an importance sampling algorithm under a maximum likelihood framework as implemented in MIGRAINE . First, we set the parameter of multistep mutation proportion (p GSM) from the range [0;1] on each population. We used the *OnePopVarSize* model since this model consider a single past change in population size. We determined high values of parameter points and trees to ensure smooth demographic projection with 20,000 trees, 500 points, and a total of 16 iterations. All demographic settings inferred by MIGRAINE were using the GSM model with a finite number of allelic states ($K = 50$).

Reconstructing population history of M. fragrans

4

We inferred the ancestry and divergence time of the genetic clusters defined by STRUCTURE (see Results section), using DIYABC Random Forest v.1.1.1 . Populations assigned to a given genetic cluster according to STRUCTURE results were pooled as a single genetic group to reduce demographic model complexity. For these analyses, we only kept individuals with a cluster assignment probability higher than 0.85.

We first performed the analysis on each dataset (nSSRs and cpDNA) separately to understand the behavior of confidence interval on the posterior distribution of demographic parameters. We then combined the two datasets and analyzed them together to assess the robustness of the analyses.

Two scenarios of species' population divergence were tested with historical demographic parameters drawn from the prior distribution (Supplementary Table S1), with no admixture event included in the analysis (Supplementary Figure S1). We also linked this demographic prior distribution according to historical data of *M. fragrans* distribution in the Moluccas as stated by Hanna and Lape to infer the ancestry and divergence time of the genetic groups.

For each scenario, 100,000 coalescent simulations were performed. We checked for the congruence of simulation results under the different scenario with the observed data using (i) Principal Component Analysis (PCA) on the summary statistics from simulated data sets, and (ii) projecting the observed data on the first two PC axes. We used ABC with random forest (RF) to perform scenario choice and parameter estimation, in DIYABCRF v.1.1.1 module. For scenario choice, we used the maximum number of simulated data sets (100,000 each scenario) and the best scenario were chosen according to a classification vote. For parameter estimation, RFs of 1000 trees were used to obtain median point estimates for each demographic parameter in the model, as well as a 95% posterior credible interval.

RESULTS

Genetic diversity and population structure of M. fragrans

at nSSR loci

We found a total number of 210 alleles using the 14 nSSR (corresponding to 15 diploid genotypes, see above), with the lowest number observed in Banda ($N_A = 68$) and the highest in Bacan ($N_A = 176$). The mean number of alleles per population (N_M) ranged from 5.00 to 11.73 across the five populations (Table 1). We found a lower genetic diversity in Ambon and Banda than in Ternate, Tidore and Bacan, as shown by the rarefied allelic richness statistics ($AR_{(k=46)}$, Table 1). The inbreeding coefficient (F_{IS}) values were found to be significantly greater than zero in Ambon, Ternate, and Bacan, indicating a signal of inbreeding.

The global genetic structure among populations as measured by the F_{ST} was 0.069. Meanwhile, the pairwise F_{ST} between populations ranged from 0.012 (between Ternate and Bacan) to 0.182 (between Ambon and Tidore) (Table 2). The genetic structure within populations (F_{ST} among sub-populations) was globally relatively low (F_{ST} values from 0.008 in Bacan to 0.040 in Tidore; see details in Supplementary Table S2).

Bayesian-clustering analyses suggest that the most likely number of intra-specific genetic clusters within our data set is 2 (Supplementary Figure S2). Individuals from Ambon and Banda are assigned to a first genetic cluster (Figure 2A / 2C), referred as the 'Southern Moluccas cluster' hereafter, whereas individuals from Ternate, Tidore and Bacan are generally assigned to a second genetic cluster, that we name the 'Northern Moluccas cluster' ($F_{ST} = 0.089$ between the two clusters). A part of these individuals show variable admixture proportions (probability of assignation to a cluster between 0.2 and 0.9) with the Southern Moluccas cluster (31%, 4%, and 26%, respectively for Ternate, Tidore and Bacan).

Our STRUCTURE results also indicate that the Northern Moluccas cluster is further subdivided into two groups at $K = 3$ (Figure 2), with the first cluster (2a) including Ternate, Tidore, and part of Bacan individuals (78%), and a second cluster (2b) including 22% of individuals from Bacan island ($F_{ST} = 0.055$ between cluster 2a and 2b). We further investigated this by running an additional analysis in STRUCTURE with individuals from the Northern Moluccas cluster ($N = 160$), and also Bacan island ($N = 75$; Supplementary Table S3). This additional analysis indicated $K = 2$ as the most plausible solution (Supplementary Figure

S3), supporting the subdivision of the North Moluccas cluster into 2a and 2b as observed in the previous analysis.

Genetic diversity and population structure of M. fragrans

at cpDNA loci

Of the 94 individuals analyzed for whole cpDNA genome data, 10 were discarded as they had more than 25% missing data. Of the remaining 84 individuals (with the four *M. argentea* individuals included), we obtained a total of 146 SNPs (109 SNPs if *M. argentea* individuals are excluded). We obtained a total of 30 haplotypes in *M. fragrans* and three additional haplotypes found exclusively in *M. argentea* . Within *M. fragrans* , we observed one common haplotype (*Hap* 3, ~60% of the individuals), that was present in all populations throughout the Moluccas archipelago (Figure 3; Supplementary Figure S4). Two other common haplotypes (*Hap* 4 and *Hap* 17) were present only in the North Moluccas. Unique haplotypes were found in all populations (three in Ternate, four in Tidore, seven in Bacan, seven in Ambon and six in Banda). The most common haplotype is placed in the center of the phylogenetic network (Figure 3). We also observe that *Hap* 3 differs from the *M. argentea* haplotypes by 48 mutations.

Diversity statistics obtained for the cpDNA dataset (number of segregating sites, nucleotide diversity Π) suggest a higher level of diversity in Bacan and Ternate (Table 3). At the cluster level, the nucleotide diversity is relatively lower in the Southern Moluccas cluster (Π =0.122) than in the Northern Moluccas cluster (Π =0.280). The AMOVA shows that the genetic variance is mainly explained by the cluster (Φ = 41.50% of the total variation, $p < 0.001$; Supplementary Table S4) and the sub-population (52%) levels. We found no signal of phylogeographic structure within our dataset (N_{ST}= 0.462 < G_{ST} = 0.532; $p < 0.05$).

Inference of population size changes in M. fragrans

The DIYABCskylineplot analyses show no evidence of significant population change over time in Ternate, Tidore, and Bacan populations, with a Bayes factor (*BF* , see Jeffreys, 1998) of 0.53, 0.45, and 0.34 respectively (Figure 4). MIGRAINE revealed a population expansion in Ternate, with $\vartheta_\mu > \vartheta^{anc}_\mu$. However, Ambon and Banda exhibit a strong relatively recent bottleneck (*BF* = 2.17 and *BF* = 141, respectively). Similar patterns of population size reduction in Ambon and Banda were also shown by our MIGRAINE analyses. The ratio of scaled population sizes ϑ_{ratio} = ϑ_0 / ϑ_A (see Supplementary Table 6) indicates a recent reduction in population size in Ambon and Banda populations.

Reconstructing population history of M. fragrans

The visual inspection of simulated and observed data in the PCA plots indicates that the models considered are able to reproduce the observed genetic diversity patterns across all our datasets (see Supplementary Figure S5, S6 and S7 for the nSSR, the cpDNA and both nSSR and cpDNA data respectively). From these results, we concluded that the proposed scenarios and the prior combinations were adequate to perform ABC analysis.

Model selection according to DIYABC-RF using all three data sets (nSSR, cpDNA, and both nSSR and cpDNA) favored scenario 1 as the best supported scenario, with a posterior probability of 0.73, 0.88 and 0.79 (Figure 5; Supplementary Figure S1. A; Supplementary Table S5). According to Jeffrey's scale (Jeffreys, 1998), the Bayes factor of the three posterior probabilities indicates significant support to scenario 1 (Supplementary Table S5). Divergence time (*t*) between the two genetic clusters of *M. fragrans* occurred around 2159 to 8231 generations ago, and the population bottleneck in the South Moluccas cluster occurred around 299 to 402 generations ago (Table 4).

DISCUSSION

We aimed to identify the center of origin of the Indonesian nutmeg,*Myristica fragrans* . We applied a population genetic approach and used genetic data obtained from nuclear microsatellite genotyping and whole plastid genome sequencing. Our results suggest that *M. fragrans* originated from the South Moluccas,

and colonized North Moluccas islands well before the management of the species by humans (Middle, Late Pleistocene). The species was thus present across the entire Moluccas at the start of species trading history.

We found two different intra-specific genetic clusters in the Moluccas, one in South Moluccas (Banda and Ambon) and the other in North Moluccas (Ternate, Tidore and Bacan). The South Moluccas cluster appeared to be ancestral as indicated by our ABC analyses conducted on both nuclear and cpDNA data sets. In these analyses, the best supported demographic scenario corresponded to an ancestral population of the species in the South Moluccas and a colonization of the North Moluccas 2159 to 8231 generations ago, followed by a demographic bottleneck in the South Moluccas 299 to 402 generations ago. Assuming a generation time of 50 years in *M. fragrans* , species migration from South to North would have thus occurred c. 411-108 ka ago (corresponding to the Middle, Late Pleistocene), and the demographic bottleneck would have occurred c. 20-15 ka ago (Late Pleistocene).

The first individuals that colonized North Moluccas were probably carrying the cpDNA haplotype *Hap* 3. The *Hap* 3 haplotype is frequent and widespread across the entire Moluccas region and was also inferred to be ancestral, being directly connected to the outgroup (*M. argentea*) and had a central placement in the phylogenetic network. Most of the other haplotypes are endemic to a specific island or group of islands and appeared to have diverged from *Hap* 3. Also, South Moluccas and North Moluccas haplotypes tend to be in two different groups in the haplotype network, which is in line with the North/South genetic divide observed from nSSRs. Under this scenario, we would expect the presence of a phylogeographic signal, which is not the case. Recent human-driven dispersal of the species has potentially blurred this signal.

Species migration from South to North probably happened through animal or sea dispersal. Since the emergence of Moluccas islands in the Oligocene / Miocene and even during glacial maximas, no major land bridges have connected North and South Moluccas, even during the glacial maxima when the sea level was around 120 m lower than now (Faral et al., 2022; Hall, 1987; Hall & Wilson, 2000), which supports observed differences between the general flora of the North and South Moluccas (Rutgrink et al., 2018). The origin of the Asian genera *Myristica*has been dated around the Late Miocene based on plastid data (Doyle et al., 2004; Frost et al., 2022; Sauquet et al., 2003), at a time when both North and South Moluccas were already present. The geological history of North and South Moluccas islands and the timing of colonization obtained in our analyses thus suggest that the dispersal of*M. fragrans* was probably mediated by animals or ocean currents as opposed to rafting across tectonic terranes. Hornbills (Bucerotidae) are known to be involved in seed dispersal of Myristicaceae . These big birds potentially allow for long-dispersal events, although we cannot exclude that marine dispersal played a role in the demographic expansion of these species either. It has been demonstrated that seeds can be dispersed by marine currents over long geographic distances, and*M. fragrans* is not an exception (Bull, 2020). If some seeds kept their viability after marine dispersal, this might have also contributed to the dispersal of the species, as it has been shown for other tree species . It has also been suggested that the sea currents in the Moluccas during the Pleistocene flowed northwards (Barmawidjaja et al., 1993), allowing seed dispersal from South Moluccas to North Moluccas.

Since our demographic analyses suggest that *M. fragrans* colonized Northern Moluccas during the Pleistocene, this supports the presence of the species all across the Moluccas before the arrival of the Europeans. We found a number of historical accounts supporting this hypothesis

It seems that Europeans were in search of the mysterious historical place of cultivation of nutmeg, *i.e.* the place where nutmeg was cultivated before being imported to Europe along a complex network of different traders and nations. North Moluccas were known for the cultivation of cloves , and South Moluccas for the cultivation of nutmegs . Search for this 'golden place' and particular focus on the South Moluccas where the species was first grown in mass probably blurred their view that the species could be present elsewhere. Here, we can speculate that historical information on the distribution of the species was strongly controlled by the VOC. This idea is supported by the embargo set by the VOC on the publication of Rumphius 'Herbarium amboinense', a document that includes information about the trade and cultivation of the species (Friedrich, 2019; van der Werff & Raven, 2014).

Along the human-managed evolutionary history of *M. fragrans* , it is likely that the North Moluccas gene pool have remained unimpacted in contrast to the South Moluccas stock that have been strongly impacted. We observed higher levels of genetic diversity both for the SSR and cpDNA in the North Moluccas genetic cluster than for South Moluccas. The lowest level of genetic diversity observed in the South is probably due to the reduction in *M. fragrans* population size during the Dutch colonisation period in the 17th and 18th centuries. As a matter of fact, the Dutch secured a monopoly on the production (Robert & Trimen, 1880) and trade of nutmeg by adopting a strong control policy on the resource (extirpation of trees, destruction of a part of the production to keep prices high) (Wright, 1958). Such management practices have certainly negatively impacted the genetic diversity of the species. This reduction in population size in the South is actually supported by our demographic analyses, with a signal of a (relatively recent) bottleneck observed in Banda and a less-marked one in Ambon. In contrast, the demographic tests tend to support a constant population size in the North Moluccas islands. The dating of the population bottleneck is a bit older than the Holocene period, but dating recent bottleneck is almost impossible. Moreover, it is worth noting that it is extremely difficult to apply a time calibration in demographic analyses that use generation time as time unit for long-lived species such as trees. Here we have set an arbitrary generation time of 50 years which implies a detection of a bottleneck event before the Holocene period. Even though our analyses were based on only 15 nSSR, our results are roughly in agreement with the knowledge we have on the historical management of the species.

The observed pattern of higher genetic diversity in the North Moluccas cluster might also be the result of inter-specific gene flow. The potential hybridization of *M. fragrans* with other *Myristica* species present in these islands is not known. According to de Wilde (2014), 20 *Myristica* species would be present in the Moluccas, which suggest that a number of *Myristica* sister species can be found in sympatry with *M. fragrans* . The radiation of *Myristica* species is probably quite recent, as demonstrated by the high rate of transferability of nSSR markers from *M. fragrans* to sister species (Kusuma et al., 2020). A recent radiation implies that inter-specific gene flow may occur. It is worth noting that morphologically, *M. fragrans* individuals sampled in these islands are, to the best of our knowledge, similar to *M. fragrans* in South Moluccas. That would imply that if inter-specific gene flow occurred, it might have had a small effect on the phenotypic diversity of *M. fragrans* . In the absence of additional genetic data for other *Myristica* species present in the region, we cannot support or rule out this hypothesis that would deserve further investigation.

Perspectives about conservation and sustainable management and use .

Our results bring insights that will be important for conserving and managing the genetic diversity within *M. fragrans* . Genetic differentiation observed between North and South Moluccas in this study should be integrated into conservation strategies and management practices for this species. The genetic diversity of this species in Banda and Ambon has been particularly impacted by past human practices, with potential consequences on the adaptive potential of this species in these islands. To date, this lower level of genetic diversity does not seem to have caused any noticeable reduction in nutmeg production or susceptibility to pests and diseases, but such an impact may appear in the near future from indirect drivers of change, in particular climate change.

Banda and Ambon islands were, and still are the main sources of planting material introduced in other parts of Indonesia and across all major tropical regions in the world . Given the relatively low level of genetic diversity in these islands, low number of male trees kept on farms, and the fact that seeds used to produce planting material are usually taken from one mother tree only, it is likely that the genetic diversity introduced into other regions of the world is also quite low. In addition, we have no information about potential differences in quality between North and South Moluccas. The Banda island nutmeg 'variety' is the most well-known, but it is not known whether if the popularity of this stock is the legacy of historical cultivation being centered around the Banda islands, or if it was intentionally selected based on resilience to environmental factors (e.g., soil properties), or genetic differences, or the combination of the above. Testing differences in the quality of the nuts by conducting common garden experiments would be needed to inform strategies of genetic resource management for this species . If no significant differences exist, it would be

interesting to consider the sourcing of additional material from North Moluccas in order to increase the level of genetic diversity and thus the potential resilience of introduced populations.

We cannot exclude that additional original genetic resources are present elsewhere. Here we have sampled only a representative part of the Moluccas islands. Indeed, this sampling regime was sufficient to infer the native range of the species, but it remains insufficient for conservation considerations. The presence of a third genetic cluster in Bacan island is interesting as it outlines an additional genetic structuring. We would need additional samples from the whole North Moluccas area (Halmahera, Sula, Obi, Misool in particular) to better interpret this result and get a more comprehensive view of the spatial genetic diversity of *M. fragrans* . Conducting an exhaustive sampling in a complex archipelago such as Indonesia is particularly complicated and time consuming (˜1000 islands for the Moluccas only). This probably explains why local flora remains largely undocumented. Concerning *M. fragrans* , we cannot rule out that some additional unique resources exist somewhere else in the Moluccas, such as in the more Southern islands (Kai, Aru, Tanimbar, etc), but a much more intensive sampling is needed to get the answer.

ACKNOWLEDGEMENTS

REFERENCES

Avise, J. C. (2009). Phylogeography: retrospect and prospect.*Journal of Biogeography* , *36* (1), 3–15. https://doi.org/10.1111/j.1365-2699.2008.02032.x

Barmawidjaja, B. M., Rohling, E. J., van der Kaars, W. A., Vergnaud Grazzini, C., & Zachariasse, W. J. (1993). Glacial conditions in the northern Molucca Sea region (Indonesia). *Palaeogeography, Palaeoclimatology, Palaeoecology* , *101* (1), 147–167. https://doi.org/https://doi.org/10.1016/0031-0182(93)90157-E

Baute, G. J., Kane, N. C., Grassa, C. J., Lai, Z., & Rieseberg, L. H. (2015). Genome scans reveal candidate domestication and improvement genes in cultivated sunflower, as well as post-domestication introgression with wild relatives. *The New Phytologist* ,*206* (2), 830–838. https://doi.org/10.1111/nph.13255

Beaumont, M. A., Zhang, W., & Balding, D. J. (2002). Approximate Bayesian Computation in Population Genetics. *Genetics* ,*162* (4), 2025–2035. https://doi.org/10.1093/genetics/162.4.2025

Brixius, D. (2018). A hard nut to crack: nutmeg cultivation and the application of natural history between the Maluku islands and Isle de France (1750s–1780s). *The British Journal for the History of Science* , *51* (4), 585–606. https://doi.org/10.1017/S0007087418000754

Burgarella, C., Barnaud, A., Kane, N. A., Jankowski, F., Scarcelli, N., Billot, C., Vigouroux, Y., & Berthouly-Salazar, C. (2019). Adaptive Introgression: An Untapped Evolutionary Mechanism for Crop Adaptation.*Frontiers in Plant Science* , *10* . https://doi.org/10.3389/fpls.2019.00004

Burkill, I. H. (1953). Habits of man and the origins of the cultivated plants of the Old World. *Proceedings of the Linnean Society of London* , *1* (64), 12–42.

Chybicki, I. J., & Burczyk, J. (2009). Simultaneous Estimation of Null Alleles and Inbreeding Coefficients. *Journal of Heredity* ,*100* (1), 106–113. https://doi.org/10.1093/jhered/esn088

Collin, F., Durif, G., Raynal, L., Lombaert, E., Gautier, M., Vitalis, R., Marin, J., & Estoup, A. (2021). Extending approximate Bayesian computation with supervised machine learning to infer demographic history from genetic polymorphisms using DIYABC Random Forest.*Molecular Ecology Resources* , *21* (8), 2598–2613. https://doi.org/10.1111/1755-0998.13413

Cornille, A., Gladieux, P., Smulders, M. J. M., Roldán-Ruiz, I., Laurens, F., Le Cam, B., Nersesyan, A., Clavel, J., Olonova, M., Feugey, L., Gabrielyan, I., Zhang, X.-G., Tenaillon, M. I., & Giraud, T. (2012). New Insight into the History of Domesticated Apple: Secondary Contribution of the European Wild Apple to the Genome of Cultivated Varieties. *PLoS Genetics* , *8* (5), e1002703. https://doi.org/10.1371/journal.pgen.1002703

Cornuet, J.-M., Pudlo, P., Veyssier, J., Dehne-Garcia, A., Gautier, M., Leblois, R., Marin, J.-M., & Estoup, A. (2014). DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics* , *30* (8), 1187–1189. https://doi.org/10.1093/bioinformatics/btt763

Crawford, K. M., & Whitney, K. D. (2010). Population genetic diversity influences colonization success. *Molecular Ecology* , *19* (6), 1253–1263. https://doi.org/10.1111/j.1365-294X.2010.04550.x

Csilléry, K., Blum, M. G. B., Gaggiotti, O. E., & François, O. (2010). Approximate Bayesian Computation (ABC) in practice. *Trends in Ecology & Evolution* , *25* (7), 410–418. https://doi.org/10.1016/j.tree.2010.04.001

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., & Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics* , *27* (15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools.*GigaScience* , *10* (2). https://doi.org/10.1093/gigascience/giab008

Darlington, C. D. (1973). *Chromosome Botany and the Origins of cultivated plants* (3rd editio). George Allen & Unwin.

Dayuan, W. (1339). *Summary Record of the Island Foreigners* . National Central Library, Taiwan.

de Candolle, A. (1855). *Géographie botanique raisonnée : ou exposition des faits principaux et des lois concernant la distribution geographique des plantes de l'epoque actuelle* . Librairie de Victor Masson.

de Candolle, A. (1882). *Origin of Cultivated Plants* . Kessinger Publishing, LLC, Whitefish, MT.

de Oliveira, S. M., Duijm, E., & Steege, H. ter. (2020). Chloroplast genome of the nutmeg tree: Myristica fragrans Houtt. (Myristicaceae).*BioRxiv* , 2020.02.25.964122. https://doi.org/10.1101/2020.02.25.964122

de Villemereuil, P., Gaggiotti, O. E., Mouterde, M., & Till-Bottraud, I. (2016). Common garden experiments in the genomic era: new perspectives and opportunities. *Heredity* , *116* (3), 249–254. https://doi.org/10.1038/hdy.2015.93

Dhont, F. (2022). Of Nutmeg and Forts: Indonesian Pride in the Banda Islands' Unique Natural and Cultural Landscape. *ETropic: Electronic Journal of Studies in the Tropics* , *21* (1), 83–98. https://doi.org/10.25120/etropic.21.1.2022.3864

Dick, C. W., Bermingham, E., Lemes, M. R., & Gribel, R. (2007). Extreme long-distance dispersal of the lowland tropical rainforest tree Ceiba pentandra L. (Malvaceae) in Africa and the Neotropics. *Molecular Ecology* , *16* (14), 3039–3049. https://doi.org/10.1111/j.1365-294X.2007.03341.x

10

Donkin, R. A. (2003). *Between east and west : the Moluccas and the traffic in spices up to the arrival of Europeans* . Philadelphia, PA : American Philosophical Society.

Doyle, J. A., Sauquet, H., Scharaschkin, T., & Le Thomas, A. (2004). Phylogeny, Molecular and Fossil Dating, and Biogeographic History of Annonaceae and Myristicaceae (Magnoliales). *International Journal of Plant Sciences* , *165* (S4), S55–S67. https://doi.org/10.1086/421068

Earl, D. A., & VonHoldt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* , *4* (2), 359–361. https://doi.org/10.1007/s12686-011-9548-7

Ellen, R. (2003). Introduction. In *On the Edge of the Banda Zone*(pp. 1–20). University of Hawai'i Press.

Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* , *10* (3), 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

Faral, A., Lavigne, F., Mutaqin, B. W., Mokadem, F., Achmad, R., Ningrum, R. W., Lahitte, P., Hadmoko, D. S., & Mei, E. T. W. (2022). A 22,000-year tephrostratigraphy record of unidentified volcanic eruptions from Ternate and Tidore islands (North Maluku, Indonesia). *Journal of Volcanology and Geothermal Research* , *423* , 107474. https://doi.org/https://doi.org/10.1016/j.jvolgeores.2022.107474

Fish, S. (2012). *The manila-acapulco galleons : the treasure ships of the pacific : with an annotated list of the transpacific galleons 1565-1815* . AuthorHouse.

Friedrich, S. (2019). The Importance of Being a Good Employee: Georg Everhard Rumphius, the Dutch East India Company, and Knowledge in the Late Seventeenth Century. *Early Modern Low Countries* ,*3* (2), 183. https://doi.org/10.18352/emlc.110

Frost, L., Santamaria-Aguilar, D. A., Singletary, D., & Lagomarsino, L. P. (2022). Neotropical niche evolution of Otoba trees in the context of global biogeography of the nutmeg family. *Journal of Biogeography* , *49* (1), 156–170. https://doi.org/10.1111/jbi.14290

Gandour, M., Hessini, K., & Abdelly, C. (2008). Understanding the population genetic structure of coastal species (Cakile maritima): seed dispersal and the role of sea currents in determining population structure. *Genetics Research* , *90* (2), 167–178. https://doi.org/DOI: 10.1017/S0016672308009269

Goudet, J. (2005). hierfstat, a package for r to compute and test hierarchical F-statistics. *Molecular Ecology Notes* , *5* (1), 184–186. https://doi.org/10.1111/j.1471-8286.2004.00828.x

Hall, R. (1987). Plate boundary evolution in the Halmahera region, Indonesia. *Tectonophysics* , *144* (4), 337–352. https://doi.org/https://doi.org/10.1016/0040-1951(87)90301-5

Hall, R., & Wilson, M. E. J. (2000). Neogene sutures in eastern Indonesia. *Journal of Asian Earth Sciences* , *18* (6), 781–808. https://doi.org/https://doi.org/10.1016/S1367-9120(00)00040-7

Hanaoka, S., Chien, C.-T., Chen, S.-Y., Watanabe, A., Setsuko, S., & Kato, K. (2014). Genetic structures of Calophyllum inophyllum L., a tree employing sea-drift seed dispersal in the northern extreme of its distribution. *Annals of Forest Science* , *71* (5), 575–584. https://doi.org/10.1007/s13595-014-0365-5

Hanna, W. (1991). *Indonesian Banda: Colonialism and its aftermath in the Nutmeg Islands* (U. S. Inst.for the Study of Human Issues, Ed.; 1st ed.). Yayasan Warisan dan Budaya Banda Naira.

Hanna, W. A. (1990). *Turbulent times past in Ternate and Tidore / by Willard A. Hanna & Des Alwi* (D. Alwi, Ed.). Yayasan Warisan dan Budaya Banda Naira.

Hanson, A. A. (1952). The Origin, Variation, Immunity, and Breeding of Cultivated Plants. *Agronomy Journal* , *44* (2), 102. https://doi.org/https://doi.org/10.2134/agronj1952.00021962004400020016x

Hardy, O. J., & Vekemans, X. (2002). spagedi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, *2* (4), 618–620. https://doi.org/10.1046/j.1471-8286.2002.00305.x

Harlan, J. R. (1986). Plant Domestication: Diffuse Origins and Diffusions. In C. Barigozzi (Ed.), *The origin and domestication of cultivated plants* (Vol. 16, pp. 21–34). Elsevier. https://doi.org/https://doi.org/10.1016/B978-0-444-42703-8.50007-5

Harris, D. R. (1990). Vavilov's concept of centres of origin of cultivated plants: its genesis and its influence on the study of agricultural origins. *Biological Journal of the Linnean Society*, *39* (1), 7–16. https://doi.org/10.1111/j.1095-8312.1990.tb01608.x

Hirota, S. K., Yasumoto, A. A., Nitta, K., Tagane, M., Miki, N., Suyama, Y., & Yahara, T. (2021). Evolutionary history of Hemerocallis in Japan inferred from chloroplast and nuclear phylogenies and levels of interspecific gene flow. *Molecular Phylogenetics and Evolution*, *164*, 107264. https://doi.org/10.1016/j.ympev.2021.107264

Hostetler, L. (2005). *Qing colonial enterprise : ethnography and cartography in early modern china* . University of Chicago Press.

Hutsemekers, V., Szovenyi, P., Shaw, A. J., Gonzalez-Mancebo, J.-M., Munoz, J., & Vanderpoorten, A. (2011). Oceanic islands are not sinks of biodiversity in spore-producing plants. *Proceedings of the National Academy of Sciences*, *108* (47), 18989–18994. https://doi.org/10.1073/pnas.1109119108

Iberl, K., Poschlod, P., & Reisch, C. (2023). Restoration of calcareous grasslands by natural recolonization after forest clearing and its impact on the genetic variation of three common herb species. *Biodiversity and Conservation*, *32* (2), 671–690. https://doi.org/10.1007/s10531-022-02518-2

Ibrahim, M. A., Cantrell, C. L., Jeliazkova, E. A., Astatkie, T., & Zheljazkov, V. D. (2020). Utilization of Nutmeg (Myristica fragrans Houtt.) Seed Hydrodistillation Time to Produce Essential Oil Fractions with Varied Compositions and Pharmacological Effects. *Molecules*, *25* (3), 565. https://doi.org/10.3390/molecules25030565

Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, *23* (14), 1801–1806. https://doi.org/10.1093/bioinformatics/btm233

Janzen, G. M., Wang, L., & Hufford, M. B. (2019). The extent of adaptive wild introgression in crops. *New Phytologist*, *221* (3), 1279–1288. https://doi.org/10.1111/nph.15457

Jordan, A. (2015). SPICE ISLAND STEW: CREOLIZATION OF FOODWAYS ON COLONIAL ERA NUTMEG PLANTATIONS, MALUKU PROVINCE, INDONESIA. *Journal of Indo-Pacific Archaeology* , *37* , 33. https://doi.org/10.7152/jipa.v37i0.14746

Kannan, L., & Wheeler, W. C. (2012). Maximum Parsimony on Phylogenetic networks. *Algorithms for Molecular Biology* , *7* (1), 9. https://doi.org/10.1186/1748-7188-7-9

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., & Drummond, A. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* , *28* (12), 1647–1649. https://doi.org/10.1093/bioinformatics/bts199

Keay, J. (2006). *The Spice Route: A History* (Reprint). University of California Press.

Kitamura, S. (2011). Frugivory and seed dispersal by hornbills (Bucerotidae) in tropical forests. *Acta Oecologica* , *37* (6), 531–541. https://doi.org/10.1016/j.actao.2011.01.015

Koboldt, D. C., Zhang, Q., Larson, D. E., Shen, D., McLellan, M. D., Lin, L., Miller, C. A., Mardis, E. R., Ding, L., & Wilson, R. K. (2012). VarScan 2: Somatic mutation and copy number alteration discovery in

cancer by exome sequencing. *Genome Research* , *22* (3), 568–576. https://doi.org/10.1101/gr.129684.111

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). <scp>Clumpak</scp> : a program for identifying clustering modes and packaging population structure inferences across *K* . *Molecular Ecology Resources* , *15* (5), 1179–1191. https://doi.org/10.1111/1755-0998.12387

Kusuma, J., Scarcelli, N., Couderc, M., Mariac, C., Zekraoui, L., & Duminil, J. (2020). Microsatellite markers development for Indonesian nutmeg (Myristica fragrans Houtt.) and transferability to other Myristicaceae spp. *Molecular Biology Reports* , *47* (6), 4835–4840. https://doi.org/10.1007/s11033-020-05535-y

Lape, P., Peterson, E., Tanudirjo, D., Shiung, C.-C., Lee, G.-A., Field, J., & Coster, A. (2018). New Data from an Open Neolithic Site in Eastern Indonesia. *Asian Perspectives* , *57* (2), 222–243. https://doi.org/10.1353/asi.2018.0015

Lape, P. V. (2000). Political dynamics and religious change in the late pre-colonial Banda Islands, Eastern Indonesia. *World Archaeology* ,*32* (1), 138–155. https://doi.org/10.1080/004382400409934

Leblois, R., Pudlo, P., Neron, J., Bertaux, F., Reddy Beeravolu, C., Vitalis, R., & Rousset, F. (2014). Maximum-Likelihood Inference of Population Size Contractions from Microsatellite Data. *Molecular Biology and Evolution* , *31* (10), 2805–2823. https://doi.org/10.1093/molbev/msu212

Leigh, J. W., & Bryant, D. (2015). <scp>popart</scp> : full-feature software for haplotype network construction. *Methods in Ecology and Evolution* , *6* (9), 1110–1116. https://doi.org/10.1111/2041-210X.12410

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* , *25* (14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, X., Wei, G., El-Kassaby, Y. A., & Fang, Y. (2021). Hybridization and introgression in sympatric and allopatric populations of four oak species. *BMC Plant Biology* , *21* (1), 266. https://doi.org/10.1186/s12870-021-03007-4

Liao, Y. (2015). Medicine. In *A History of Chinese Science and Technology* (pp. 1–159). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-44166-4_1

Librado, P., & Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* , *25* (11), 1451–1452. https://doi.org/10.1093/bioinformatics/btp187

Liu, J., Lindstrom, A. J., Chen, Y., Nathan, R., & Gong, X. (2021). Congruence between ocean-dispersal modelling and phylogeography explains recent evolutionary history of Cycas species with buoyant seeds.*New Phytologist* , *232* (4), 1863–1875. https://doi.org/10.1111/nph.17663

Mariac, C., Luong, V., Kapran, I., Mamadou, A., Sagnard, F., Deu, M., Chantereau, J., Gerard, B., Ndjeunga, J., Bezancon, G., Pham, J.-L., & Vigouroux, Y. (2006). Diversity of wild and cultivated pearl millet accessions (Pennisetum glaucum [L.] R. Br.) in Niger assessed by microsatellite markers. *Theoretical and Applied Genetics* ,*114* (1), 49–58. https://doi.org/10.1007/s00122-006-0409-9

Mariac, C., Scarcelli, N., Pouzadou, J., Barnaud, A., Billot, C., Faye, A., Kougbeadjo, A., Maillol, V., Martin, G., Sabot, F., Santoni, S., Vigouroux, Y., & Couvreur, T. L. P. (2014). Cost-effective enrichment hybridization capture of chloroplast genomes at deep multiplexing levels for population genetics and phylogeography studies. *Molecular Ecology Resources* , *14* (6), 1103–1113. https://doi.org/10.1111/1755-0998.12258

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal* , *17* (1), 10. https://doi.org/10.14806/ej.17.1.200

Meegahakumbura, M. K., Wambulwa, M. C., Li, M.-M., Thapa, K. K., Sun, Y.-S., Moller, M., Xu, J.-C., Yang, J.-B., Liu, J., Liu, B.-Y., Li, D.-Z., & Gao, L.-M. (2018). Domestication Origin and Breeding History of the Tea Plant (Camellia sinensis) in China and India Based on Nuclear Microsatellites and cpDNA Sequence Data. *Frontiers in Plant Science* , *8* . https://doi.org/10.3389/fpls.2017.02270

Navascues, M., Leblois, R., & Burgarella, C. (2017). Demographic inference through approximate-Bayesian-computation skyline plots. *PeerJ*, *5*, e3530. https://doi.org/10.7717/peerj.3530

Neophytou, C., Potzelsberger, E., Curto, M., Meimberg, H., & Hasenauer, H. (2020). Population bottlenecks have shaped the genetic variation of Ailanthus altissima (Mill.) Swingle in an area of early introduction. *Forestry: An International Journal of Forest Research*, *93*(4), 495–504. https://doi.org/10.1093/forestry/cpz019

Paradis, E. (2010). pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics*, *26*(3), 419–420. https://doi.org/10.1093/bioinformatics/btp696

Petit, R. (2012). *Logiciels Permut et cpSSR* (1.2.1). INRA.

Petit, R. J., Brewer, S., Bordacs, S., Burg, K., Cheddadi, R., Coart, E., Cottrell, J., Csaikl, U. M., van Dam, B., Deans, J. D., Espinel, S., Fineschi, S., Finkeldey, R., Glaz, I., Goicoechea, P. G., Jensen, J. S., Konig, A. O., Lowe, A. J., Madsen, S. F., . . . Kremer, A. (2002). Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. *Forest Ecology and Management*, *156*(1–3), 49–74. https://doi.org/10.1016/S0378-1127(01)00634-X

Pigafetta, V. (1522). *Premier voyage autour du monde sur l'escadre de Magellan*. Librairie Ch. DeLagrave.

Pons, O., & Petit, R. J. (1996). Measwring and Testing Genetic Differentiation With Ordered Versus Unordered Alleles. *Genetics*, *144*(3), 1237–1245. https://doi.org/10.1093/genetics/144.3.1237

Price, M. R., & Hadfield, M. G. (2014). Population Genetics and the Effects of a Severe Bottleneck in an Ex Situ Population of Critically Endangered Hawaiian Tree Snails. *PLoS ONE*, *9*(12), e114377. https://doi.org/10.1371/journal.pone.0114377

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, *155*(2), 945 LP – 959.

Ray, N., & Excoffier, L. (2009). Inferring Past Demography Using Spatially Explicit Population Genetic Models. *Human Biology*, *81*(2–3), 141–157. https://doi.org/10.3378/027.081.0303

Ricklefs, M. C. (1991). A History of Modern Indonesia since c. 1300. In *A History of Modern Indonesia since c. 1300* (2nd Editio, pp. 24–36). Macmillan.

Robert, B., & Trimen, H. (1880). *Medicinal plants: Being descriptions with original figures of the principal plants employed in medicine and an account of the characters, properties, and uses of their parts and products of medicinal value. Vol. 2*. London: J. & A. Churchill.

Rousset, F., Reddy Beeravolu, C., & Leblois, R. (2018). Likelihood computation and inference of demographic and mutational parameters from population genetic data under coalescent approximations. *Journal de La Societe Francaise de Statistique*, *159*(3), 142–166.

Rumpf, G. E., Changuion, Francis., Uytwerf., H., & Catuffe., J. (1741). *Herbarium amboinense : plurimas conplectens arbores, frutices, herbas, plantas terrestres & aquaticas, quae in Amboina et adjacentibus reperiuntur insulis adcuratissime descriptas iuxta earum formas, cum diuersis denominationibus cultura, usu, ac virtutib: Vol. v.2*. Apud Fransicum Changuion, Joannem Catuffe, Hermannum Uytwerf.

Rumphius, G. E. (1755a). *Herbarii Amboinensis auctuarium*. Amstelaedami : Apud Mynardum Uytwerf, & Viduam ac Filium S. Schouten, 1755.

Rumphius, G. E. (1755b). The Ambonese Herbal, Volume 2. In E. M. Beekman (Ed.), *Book XII: Concerning the Little Sea Trees, and Stony Sea Growths, Which Resemble Plants; Auctuarium, or Augmentation of The Ambonese Herbal* (1st ed.). Yale University Press.

Rutgrink, A. L. J., Visser, M., & van Welzen, P. C. (2018). Differences between the floras of the North and South Moluccas (Indonesia). *Journal of Systematics and Evolution*, *56*(6), 652–662.

https://doi.org/10.1111/jse.12301

Sauer, C. O. (1952). *Agricultural origins and dispersals* (1st editio). American Geographical Society.

Sauquet, H., Doyle, J. A., Scharaschkin, T., Borsch, T., Hilu, K. W., Chatrou, L. W., & Le Thomas, A. (2003). Phylogenetic analysis of Magnoliales and Myristicaceae based on multiple data sets: implications for character evolution. *Botanical Journal of the Linnean Society* , *142* (2), 125–186. https://doi.org/10.1046/j.1095-8339.2003.00171.x

Scarcelli, N., Mariac, C., Couvreur, T. L. P., Faye, A., Richard, D., Sabot, F., Berthouly-Salazar, C., & Vigouroux, Y. (2016). Intra-individual polymorphism in chloroplasts from <scp>NGS</scp> data: where does it come from and how to handle it? *Molecular Ecology Resources* , *16* (2), 434–445. https://doi.org/10.1111/1755-0998.12462

Schery, R. W. (1952). *Plants for man* (1st editio). Prentice-Hall, Inc., Englewood Cliffs.

Schreiber, M., Stein, N., & Mascher, M. (2018). Genomic approaches for studying crop evolution. *Genome Biology* , *19* (1), 140. https://doi.org/10.1186/s13059-018-1528-8

Smith, C. E. (1968). The New World Centers of Origin of Cultivated Plants and the Archaeological Evidence. *Economic Botany* ,*22* (3), 253–266.

Smith, C. E. (1969). From Vavilov to the Present: A Review.*Economic Botany* , *23* (1), 2–19.

Sonnerat, P. (1776). *Voyage a la Nouvelle Guinee* . Chez Ruault, Libraire, rue de la Harpe (De l'imprimerie de Demonville, Imp. Lib. rue S.

van der Werff, H., & Raven, P. H. (2014). The Herbarium Amboinense, from Creator to Translator. *Allertonia* , *13* , 47–55.

van Ruth, S. M., Silvis, I. C. J., Alewijn, M., Liu, N., Jansen, M., & Luning, P. A. (2019). No more nutmegging with nutmeg: Analytical fingerprints for distinction of quality from low-grade nutmeg products.*Food Control* , *98* , 439–448. https://doi.org/10.1016/j.foodcont.2018.12.005

Vavilov, N. (1926). *Studies on the Origin of Cultivated Plants*(1st ed.).

Wang, X.-H., Li, J., Zhang, L.-M., He, Z.-W., Mei, Q.-M., Gong, X., & Jian, S.-G. (2019). Population Differentiation and Demographic History of the Cycas taiwaniana Complex (Cycadaceae) Endemic to South China as Indicated by DNA Sequences and Microsatellite Markers. *Frontiers in Genetics* , *10* . https://doi.org/10.3389/fgene.2019.01238

Wright, H. R. C. (1958). The Moluccan Spice Monopoly, 1770-1824.*Journal of the Malayan Branch of the Royal Asiatic Society* ,*31* (4 (184)), i–127.

Yao, Y.-X., Shang, X.-P., Yang, J., Lin, R.-Z., Huai, W.-X., & Zhao, W.-X. (2020). Genetic Variation May Have Promoted the Successful Colonization of the Invasive Gall Midge, Obolodiplosis robiniae, in China. *Frontiers in Genetics* , *11* . https://doi.org/10.3389/fgene.2020.00387

Zhang, Q., Zhang, D., Yu, K., Ji, J., Liu, N., Zhang, Y., Xu, M., Zhang, Y.-J., Ma, X., Liu, S., Sun, W.-H., Yu, X., Hu, W., Lan, S.-R., Liu, Z.-J., & Liu, W. (2021). Frequent germplasm exchanges drive the high genetic diversity of Chinese-cultivated common apricot germplasm.*Horticulture Research* , *8* . https://doi.org/10.1038/s41438-021-00650-8

DATA ACCESSIBILITY AND BENEFIT-SHARING

Data has been submitted to DRYAD and should be accessible at https://10.5061/dryad.qjq2bvqm0.

COMPETING INTERESTS

The authors declare no competing financial interests.

AUTHOR CONTRIBUTIONS

JK and JD conducted the field work in Ambon; JK conducted remaining field work in Banda, Ternate, Tidore, and Bacan; JK, NS, and JD conceptualized and designed research; JK and MC performed lab experiment; JK, NS, and MdN analyzed the data; JK and JD wrote original draft of the manuscript; PG contributed to data analysis and draft of the manuscript. All authors read and approved the final manuscript. JK is a PhD student registered at AgroParisTech, France.

Hosted file

`Figures_MolEcol_Kusuma_et_al.docx` available at https://authorea.com/users/646443/articles/658359-retracing-the-center-of-origin-and-evolutionary-history-of-nutmeg-myristica-fragrans-an-emblematic-spice-tree-species

Hosted file

`Tables_MolEcol_Kusuma_et_al.docx` available at https://authorea.com/users/646443/articles/658359-retracing-the-center-of-origin-and-evolutionary-history-of-nutmeg-myristica-fragrans-an-emblematic-spice-tree-species