



HAL
open science

De l'histoire évolutive des noix de muscade indonésiennes

Jakty Kusuma

► **To cite this version:**

Jakty Kusuma. De l'histoire évolutive des noix de muscade indonésiennes. Biodiversité et Ecologie. AgroParisTech, 2023. Français. NNT : 2023AGPT0007 . tel-04842210

HAL Id: tel-04842210

<https://pastel.hal.science/tel-04842210v1>

Submitted on 17 Dec 2024

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

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

**THÈSE POUR OBTENIR LE GRADE DE DOCTEUR
DE L'INSTITUT DES SCIENCES ET INDUSTRIES DU VIVANT ET DE
L'ENVIRONNEMENT - AGROPARISTECH**

N° : 2023 AGPT 0007

En Sciences de l'évolution et de la Biodiversité

**École doctorale GAIA – Biodiversité, Agriculture, Alimentation, Environnement, Terre, Eau – n°584
Portée par l'Université de Montpellier**

Unité de recherche DIADE

ON THE EVOLUTIONARY HISTORY OF INDONESIAN NUTMEGS

Présentée par Jakty KUSUMA

Le 24 août 2023

**Sous la direction de Jérôme DUMINIL
et sous le co-encadrement de Pierre GÉRARD**

Devant le jury composé de

Amandine CORNILLE, Chargée de Recherche (HDR), CNRS

Myriam HEUERTZ, Directrice de Recherche, INRAE

Jacques DAVID, Professeur, L'Institut Agro Montpellier

Nicolas HUBERT, Directeur de Recherche, IRD

Nora SCARCELLI, Chargée de Recherche, IRD

Rapporteuse

Rapporteuse

Examineur

Président du Jury

Invitée

On the evolutionary history of Indonesian nutmegs



“The history of the nutmeg is one of conquest and exploitation – of both human life and the natural environment – and the origin of our contemporary climate crisis.”

The nutmeg’s curse: Parables for a planet in crisis.

by Amitav Ghosh

De l’histoire évolutive des noix de muscade indonésiennes

Par Jakty KUSUMA

Ecole Doctorale GAIA, Université de Montpellier – AgroParisTech

2019 – 2023

Résumé en français

Introduction

Les forêts de l'archipel Malaise représentent une opportunité intéressante pour la recherche sur l'histoire évolutive de la biodiversité tropicale en raison de leur richesse exceptionnelle en espèces végétales. Dans ces forêts, de nombreuses espèces végétales fournissent des ressources uniques très utiles pour l'homme (Middleton et al., 2019). L'interaction entre les humains et les forêts tropicales de l'archipel Malaise remonte à au moins quinze siècles (Reid, 1995), la forêt étant une source primaire de nourriture et de subsistances pour les communautés locales. Cependant, comme les activités humaines continuent de façonner le paysage et d'influencer l'utilisation des ressources forestières, il est inévitable que la composition de la biodiversité dans la région subisse des transformations.

Parmi les espèces d'arbres forestiers collectées ou récoltées dans les forêts naturelles plutôt que cultivées en milieu agricole, *Myristica fragrans* et *Myristica argentea* (regroupées dans cette thèse sous le vocable 'noix de muscade indonésiennes') se distinguent comme deux exemples notables utilisés à des fins diverses. La distribution naturelle de ces deux espèces est géographiquement limitée. *Myristica fragrans* se trouve principalement dans la région de Wallace (notamment l'archipel des Moluques), tandis que *M. argentea* est originaire de l'île de Papouasie (de Wilde, 1995, 2014).

L'aire de répartition étroite de ces espèces insulaires est caractérisée par un paysage fragmenté composé d'îles. En conséquence, l'isolement géographique et la petite taille des populations influencent la distribution de la diversité génétique de ces espèces, et l'attendu théorique serait une forte différenciation génétique entre

les îles et une faible diversité génétique au sein de chaque île (Barret, 1996; Frankham, 1996).

A l'origine, *M. fragrans* et *M. argentea* étaient précieuses pour les communautés locales en raison de leur importance en tant qu'arbres à épices (noix et macis). Prélevées dans leur milieu naturel, elles n'étaient pas soumises à des pratiques culturelles et commerciales à grande échelle. Cependant, en raison de la demande croissante pour les épices qu'elles fournissent, elles ont été introduites et cultivées en dehors de leur aire de répartition d'origine depuis le XVIIIe siècle (Brixius, 2015). Le déplacement délibéré de matériel végétal au-delà des frontières géographiques a facilité l'établissement de nouvelles populations. Cette dispersion par l'homme et l'expansion concomitante de l'aire de répartition de l'espèce posent la question du niveau de diversité génétique introduit au sein de ces populations.

L'évaluation de l'impact de la dispersion anthropique sur les flux génétique dans les aires de répartition naturelles pose des problèmes en raison de l'interaction complexe entre les processus de dispersion naturelle et anthropique. Il est difficile de démêler les contributions des deux processus lorsqu'ils se produisent simultanément (Einfeldt et al., 2020). Les caractéristiques géographiques et la gestion de la culture de ces deux espèces pourraient constituer des cas d'études particulièrement pertinents pour notre compréhension de la distribution de la diversité génétique et de l'histoire évolutive des arbres à épices. D'autre part, il a été suggéré que l'interaction humaine avec les espèces d'arbres cultivées affecte l'évolution des espèces (Cornille et al., 2015; Sánchez-Robles et al., 2022; Soares et al., 2019). Par conséquent, les caractéristiques géographiques uniques, l'histoire de leurs interactions avec l'Homme et les pratiques culturelles associées à *M.*

fragrans et *M. argentea* offrent des indications précieuses sur la répartition de la diversité génétique et l'histoire de l'évolution des espèces végétales dans la région.

Par ailleurs, la compréhension de la dynamique spatio-temporelle de la diversité génétique de ces espèces est cruciale pour la mise en place de stratégies de conservation, de gestion et d'utilisation durable face aux changements environnementaux et anthropiques en cours. Il est ainsi nécessaire d'étudier les mécanismes précis qui ont façonné l'histoire évolutive de la noix de muscade indonésienne, en tenant compte des impacts des fluctuations climatiques passées, des barrières géographiques, des pratiques culturelles et des interactions humaines. En utilisant les noix de muscade Indonésiennes comme espèces modèle, nous pouvons obtenir des informations précieuses sur la manière dont ces facteurs ont influencé leur diversité génétique, la structure et la dynamique de sa population tout au long de sa trajectoire évolutive.

Des facteurs environnementaux et anthropiques ont très probablement contribué à l'histoire de l'évolution des espèces de *Myristica*. L'évolution de la flore locale en relation avec l'évolution du climat, les modifications tectoniques et les pratiques de gestion humaine restent peu documentées. Les pratiques de gestion humaine des noix de muscade indonésiennes ont contribué à leur histoire évolutive. Par exemple, l'abattage des arbres de *M. fragrans* par la Compagnie Néerlandaise des Indes orientales (Vereenigde Oostindische Compagnie en Néerlandais ou VOC) pour maintenir un monopole sur la ressource, a probablement eu des conséquences sur les niveaux de diversité génétique de l'espèce. En outre, des événements fondateurs ont pu se produire en raison de la propagation des noix de muscade indonésiennes en dehors de leur distribution d'origine. Cependant, l'impact

des facteurs environnementaux et anthropiques sur la distribution de la diversité génétique des muscadiers indonésiens n'a pas été étudié à ce jour. Ce manque de connaissances est très préoccupant, car la diversité génétique est cruciale pour l'adaptation et la résilience aux environnements changeants. En outre, la compréhension de la structure de la population et de l'histoire démographique des muscadiers indonésiens peut fournir des informations précieuses sur leur histoire évolutive et éclairer les stratégies de conservation. Il est donc urgent de réaliser une étude complète sur la structure génétique et l'histoire démographique des muscadiers indonésiens, en tenant compte de l'impact potentiel des facteurs environnementaux et anthropiques. Une telle étude peut contribuer au développement de pratiques culturelles plus efficaces et durables, ainsi qu'à la conservation et à la gestion de cette culture importante.

Dans cette thèse, j'ai cherché à élucider l'histoire évolutive des noix de muscade indonésiennes, en me concentrant sur les deux espèces *M. fragrans* et *M. argentea*. J'ai divisé cette thèse en cinq chapitres différents présentés sous forme d'articles. La mise en page et le format de référence de ces chapitres ont été modifiés à partir des manuscrits originaux afin d'assurer une meilleure cohérence typographique au sein de cette thèse.

Le **chapitre 1** décrit les marqueurs microsatellites (SSR) développés dans le cadre de mes travaux chez *M. fragrans*. Nous fournissons les premières informations sur ses caractéristiques ainsi que sur sa transférabilité à d'autres espèces sœurs. Comme point de départ de ma thèse, ces marqueurs génétiques servent d'outil pour acquérir des informations sur la structure génétique et l'histoire

démographique des noix de muscade indonésiennes. Cet article a été publié dans *Molecular Biology Reports* (2020).

Le **chapitre 2** présente la séquence complète de l'ADN chloroplastique (cpDNA) de *M. argentea* et *M. fatua*. De l'assemblage de novo à l'annotation, nous décrivons pour la première fois le génome de l'ADN chloroplastique chez ces deux espèces. L'importance de ces ressources génomiques constitue une référence précieuse pour l'étude de la distribution phylogéographique des espèces focales. Ce manuscrit fait l'objet d'une révision mineure dans le journal *Mitochondrial DNA Part B*

Le **chapitre 3** traite de l'origine de *M. fragrans* dans les Moluques. Ce chapitre se concentre sur la génétique des populations et sur l'histoire biogéographique de la noix de muscade dans les Moluques. Plus précisément, nous avons cherché à déterminer si l'espèce était originaire des seules îles Banda ou de l'ensemble des Moluques. Ce manuscrit est en cours de soumission à *Molecular Ecology*.

Le **chapitre 4** présente l'histoire de l'introduction de *M. fragrans* en dehors de son aire de distribution naturelle. Il est bien documenté que l'espèce est maintenant répandue dans les zones tropicales du monde, et nous tentons de retracer ici l'histoire de son introduction ainsi que la distribution de la diversité génétique de l'espèce dans le monde.

Le **chapitre 5** décrit la distribution spatiale de la diversité génétique de *M. argentea*. Nous avons testé l'influence de différents facteurs (barrières historiques aux flux de gènes, pratiques de gestion récentes) sur la structure génétique et la dynamique historique des populations de l'espèce.

Résultats

Avant le début de l'étude, la compréhension globale de l'histoire évolutive de *M. fragrans* et de *M. argentea* était difficilement réalisable du fait de la rareté des marqueurs génétiques spécifiques chez ces espèces. Une recherche bibliographique approfondie menée en 2019 n'a révélé que deux ressources génétiques disponibles, à savoir des SSR chez l'espèce *M. malabarica* et la séquence incomplète du génome de l'ADN chloroplastique (voir (Hemmilä et al., 2010; Tallei & Kolondam, 2015)). En particulier, pour les SSR de *M. malabarica*, l'auteur n'a pas testé la transférabilité des marqueurs chez d'autres espèces de *Myristica*. Ces limitations ont mis en avant la nécessité d'approfondir les recherches et de développer des marqueurs génétiques plus fiables et mieux adaptés pour atteindre les objectifs de la thèse.

Par conséquent, il y avait une motivation pour développer de nouvelles ressources génomiques, comme détaillé dans les **chapitres 1 et 2** de cette étude. Les marqueurs SSR développés présentent un polymorphisme intéressant pour l'étude de la répartition de la diversité génétique chez *M. fragrans* et leur transférabilité à d'autres espèces de *Myristica* (*M. fatua*, *M. argentea* et *M. crassipes*) et à un autre genre de la famille des Myrsiticaceae (*Horsfieldia palauensis*) a été démontrée. En outre, cette étude a permis d'assembler et d'annoter de novo les génomes de l'ADN chloroplastique (ADNc) de deux espèces de *Myristica* (*M. argentea* et *M. fatua*) avec des couvertures de 41,2x et 31,8x, respectivement. Cette réalisation est importante car les séquences annotées du génome de l'ADN chloroplastique sont intéressants pour retracer l'histoire de l'évolution des espèces à l'aide de marqueurs transmis maternellement (c'est-à-dire

uniquement par voie de graines). De plus, la reconstruction phylogénétique qui en résulte, basée sur l'ensemble du génome de l'ADN chloroplastique, a confirmé une relation monophylétique des espèces de *Myristica*, comme indiqué précédemment (Sauquet et al., 2003), ce qui fournit des informations précieuses sur l'évolution et la classification du genre.

En utilisant ces SSR développés de novo et des génomes chloroplastiques entiers, l'histoire évolutive de *M. fragrans* dans son aire de répartition d'origine (l'archipel des Moluques) a été reconstituée (**chapitre 3**). Les populations de *M. fragrans* dans les Moluques se sont subdivisées en deux groupes génétiques intraspécifiques distincts, l'un situé dans le groupe des Moluques du Sud et l'autre dans celui des Moluques du Nord. Le groupe des Moluques du Nord présente une diversité génétique nettement plus grande que celui des Moluques du Sud. Il est probable que la diversité génétique plus faible au niveau des Moluques du Sud résulte des pratiques de gestion de l'espèce par les Néerlandais, ayant menées à une réduction drastique de la taille de population de l'espèce dans cette région. Au contraire, les populations des Moluques du Nord n'ont a priori subi aucune pratique de gestion. L'inférence démographique bayésienne a fourni des arguments convaincants en faveur d'une origine de l'espèce dans les Moluques du Sud, d'une colonisation des Moluques du Nord au cours du Pléistocène moyen à tardif, et d'un déclin démographique relativement récent (au cours du Pléistocène tardif, il y a environ 20 000 à 15 000 ans) dans les Moluques du Sud.

Dans le **chapitre 4**, nous avons élargi notre échantillonnage afin de couvrir au mieux la distribution mondiale de *M. fragrans* (populations naturelles et introduites) et de fournir une compréhension complète de l'histoire de la

propagation de l'espèce dans le monde. Outre l'utilisation de SSR et de génomes entiers cpDNA, nous avons également utilisé des séquences nucléaires ciblées pour renforcer nos déductions sur l'histoire évolutive de l'espèce. En analysant des données provenant de diverses sources, y compris des spécimens d'herbiers, l'étude présente un compte rendu détaillé de la répartition mondiale de *M. fragrans*. Nos résultats suggèrent que les populations de noix de muscade introduites proviennent principalement de matériel de propagation collecté dans les Moluques du Sud, mais aussi et dans une moindre mesure, des Moluques du Nord.

Enfin, dans le **chapitre 5**, nous avons démontré que la noix de muscade de Papouasie, *M. argentea*, présentait des discontinuités génétiques évidentes dans une aire de répartition aussi étroite que celle de la péninsule de Fakfak, sur l'île de Papouasie. Les analyses de regroupement bayésiennes ont révélé la présence de trois groupes génétiques distincts dans l'aire de répartition naturelle de l'espèce. Les populations introduites de l'espèce en dehors de son aire de distribution naturelle ont probablement dérivé principalement d'une population indigène (Raduria, au Sud de la région de Fakfak). Les populations de la péninsule de Fakfak partagent une distribution différente de la diversité génétique, tout en présentant généralement des niveaux de diversité génétique plus élevés que leurs homologues introduits.

Discussion

Cette étude visait à examiner les origines de *M. fragrans* et *M. argentea* à l'aide d'une approche génétique. Les hypothèses précédentes sur leurs origines étaient basées sur des données historiques mais manquaient de preuves génétiques.

Grâce à une analyse génétique utilisant des approches de ‘Approximate Bayesian Computation’ (ABC), l'étude a révélé que les populations de *M. fragrans* dans la région des Moluques du Sud représentent les populations ancestrales de l'espèce. La dispersion par l'eau est le mécanisme probable de colonisation de l'archipel des Moluques, comme le prouvent les preuves de la dispersion des graines par la mer chez la noix de muscade.

La dispersion des graines sur de longues distances joue un rôle crucial dans la colonisation des îles, les espèces végétales dispersées par l'eau ou les animaux atteignant des îles séparées par des centaines de kilomètres des populations d'origine. La divergence des pools génétiques chez *M. fragrans* suggère que la colonisation des Moluques du Sud vers les Moluques du Nord a probablement été facilitée par la dispersion par l'eau. Le moment de la colonisation correspond aux fluctuations marines passées, il y a environ 150 000 - 100 000 ans. La différenciation génétique et les différences de composition florale entre les Moluques du Sud et du Nord supportent le rôle d'une barrière géographique.

La présence de noix de muscade dans les Moluques du Nord soulève la question de la désignation des îles Banda comme source unique de l'épice à l'époque du commerce des épices. Les documents historiques indiquent la présence de muscadiers dans les Moluques du Nord, mais pour des raisons que nous ignorons, ils n'ont pas été cultivés de manière extensive. Les raisons pour lesquelles l'accent a été mis sur les îles Banda et le manque d'intérêt pour les sources alternatives dans les Moluques du Nord doivent être étudiées plus en détail.

Pour *M. argentea*, l'étude a révélé un schéma de divergence influencé par les fluctuations climatiques au cours du dernier maximum glaciaire en Papouasie

occidentale. Le dernier maximum glaciaire a déclenché des changements climatiques importants, entraînant l'émergence de nouvelles lignées chez des espèces confinées à des aires de répartition limitées. Les refuges forestiers ont servi de réservoirs génétiques et, lorsque le climat s'est réchauffé au cours de l'Holocène, les populations se sont étendues et ont recolonisé d'autres régions.

Les influences humaines, telles que la coupe massive des pieds de *M. fragrans* au cours du XVI^e siècle, ont eu un impact sur la diversité génétique dans les Moluques du Sud. En revanche, les populations natives de *M. argentea* dans la péninsule de Fakfak présentent des niveaux élevés de diversité génétique, ce qui indiquerait une pression par l'Homme sur cette ressource relativement réduite. L'identification de populations génétiquement diverses est précieuse pour les efforts de conservation et les programmes de sélection visant à développer des variétés résistantes.

Conclusion

En conclusion, cette étude fournit des preuves génétiques des origines géographiques de *M. fragrans* et de *M. argentea*, soulignant le rôle de la dispersion par l'Homme dans l'histoire de leur introduction et des fluctuations climatiques dans leur colonisation et leur divergence. Les différences dans les niveaux de diversité génétique et la dynamique des populations ont des implications importantes pour la compréhension de l'histoire évolutive et la conservation de ces espèces d'arbres.

Références

- Barret, S. C. H. (1996). The reproductive biology and genetics of island plants. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 351(1341), 725–733. <https://doi.org/10.1098/rstb.1996.0067>
- Brixius, D. (2015). A Pepper acquiring Nutmeg: Pierre Poivre, The French Spice Quest and the Role of Mediators in Southeast Asia, 1740s to 1770s. *Journal of the Western Society for French History*, 43(1).
- Cornille, A., Feurtey, A., Gélín, U., Ropars, J., Misvanderbrugge, K., Gladieux, P., & Giraud, T. (2015). Anthropogenic and natural drivers of gene flow in a temperate wild fruit tree: a basis for conservation and breeding programs in apples. *Evolutionary Applications*, 8(4), 373–384. <https://doi.org/10.1111/eva.12250>
- de Wilde, W. J. J. O. (1995). Census of *Myristica* (Myristicaceae) in New Guinea anno 1994. *Blumea: Biodiversity, Evolution and Biogeography of Plants*, 40(2), 237–344.
- de Wilde, W. J. J. O. (2014). *Flora Malesiana*. Pensoft Publishers. <https://doi.org/10.3897/ab.e1141>
- Einfeldt, A. L., Jesson, L. K., & Addison, J. A. (2020). Historical human activities reshape evolutionary trajectories across both native and introduced ranges. *Ecology and Evolution*, 10(13), 6579–6592. <https://doi.org/10.1002/ece3.6391>
- Frankham, R. (1996). Relationship of genetic variation to population size in wildlife. *Conservation Biology*, 10(6), 1500–1508. <https://doi.org/10.1046/j.1523-1739.1996.10061500.x>
- Hemmilä, S., Kumara, M., Ravikanth, G., Gustafsson, S., Vasudeva, R., Ganeshaiyah, K. N., Uma Shaanker, R., & Lascoux, M. (2010). Development of eleven microsatellite markers in the red-listed tree species *Myristica malabarica*. *Conservation Genetics Resources*, 2(S1), 305–307. <https://doi.org/10.1007/s12686-010-9212-7>
- Middleton, D. J., Armstrong, K., Baba, Y., Balslev, H., Chayamarit, K., Chung, R. C. K., Conn, B. J., Fernando, E. S., Fujikawa, K., Kiew, R., Luu, H. T., Aung, M. M., Newman, M. F., Tagane, S., Tanaka, N., Thomas, D. C., Tran, T. B., Utteridge, T. M. A., van Welzen, P. C., ... Wong, K. M. (2019). Progress on Southeast Asia's Flora projects. *Gardens' Bulletin Singapore*, 71(2), 267–319. [https://doi.org/10.26492/gbs71\(2\).2019-02](https://doi.org/10.26492/gbs71(2).2019-02)
- Reid, A. (1995). Humans and Forests in Pre-colonial Southeast Asia. *Environment and History*, 1(1), 93–110. <http://www.jstor.org/stable/20722966>

- Sánchez-Robles, J. M., García-Castaño, J. L., Balao, F., García, C., Terrab, A., & Talavera, S. (2022). Anthropogenic deforestation and climate dryness as drivers of demographic decline and genetic erosion in the southernmost European fir forests. *European Journal of Forest Research*, 141(4), 649–663. <https://doi.org/10.1007/s10342-022-01467-3>
- 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>
- Soares, L. A. S. S., Cazetta, E., Santos, L. R., França, D. de S., & Gaiotto, F. A. (2019). Anthropogenic disturbances eroding the genetic diversity of a threatened palm tree: A multiscale approach. *Frontiers in Genetics*, 10. <https://doi.org/10.3389/fgene.2019.01090>
- Tallei, T. E., & Kolondam, B. J. (2015). DNA Barcoding of Sangihe Nutmeg (*Myristica fragrans*) using matK Gene. *HAYATI Journal of Biosciences*, 22(1), 41–47. <https://doi.org/https://doi.org/10.4308/hjb.22.1.41>

Table of contents

Acknowledgements	i
Résumé en français	v
Table of contents	xiii
1. Introduction.....	1
1.1 Global context.....	1
1.2 Biogeography of the Malay Archipelago	6
1.2.1 Geological history.....	6
1.2.2 Floristic history	12
1.3 Spatio-temporal dynamics of nutmeg : the emblematic spice tree species ..	17
1.3.1 <i>Myristica fragrans</i> Houtt.....	18
1.3.2 <i>Myristica argentea</i> Warb.....	23
1.4 Evolutionary context.....	28
1.5 Genetic diversity in tree species	29
1.5.1 A microevolutionary perspectives	29
1.5.2 Evaluating the influence of evolutionary processes at intraspecific level... 38	
1.5.3 Retracing the demographic history of population	41
2. Contextual problem, thesis objectives and outline	44
References	47

Chapter 1: Microsatellites markers development for Indonesian nutmeg (*Myristica fragrans* Houtt.) and transferability to other Myristicaceae spp.

Abstract	71
Introduction	72
Methods and results.....	72
Conclusions	77
References	77
Tables	81
Supplementary online materials.....	84

Chapter 2: Complete chloroplast genome of two Nutmeg species *Myristica argentea* Warb 1891 and *Myristica fatua* Houtt 1774

Abstract	94
Introduction	95
Materials.....	95
Methods.....	96
Results	97
Discussions and conclusions	97

References	99
Figures	102
Supplementary online materials	105
Supplementary methods	107

Chapter 3: Retracing the center of origin and evolutionary history of nutmeg, *Myristica fragrans*, an emblematic spice tree species

Abstract.....	111
Introduction	112
Materials and methods.....	116
Results	122
Discussion.....	126
References	133
Tables and figures.....	142
Supplementary online materials	152
Supplementary files	165

Chapter 4: Propagation history and worldwide genetic diversity distribution of the Banda nutmeg (*Myristica fragrans* Houtt)

Abstract.....	171
Introduction	172
Results	174
Discussion.....	177
Conclusion.....	181
Materials and Methods	182
References	185
Tables and figures.....	190
Supplementary online materials	196
Supplementary files	198

Chapter 5: Influence of past climate change and human-mediated introduction on patterns of genetic diversity distribution of the Papuan nutmeg (*Myristica argentea* Warb)

Abstract.....	201
Introduction	203
Materials and methods.....	206
Results	211
Discussion.....	214
Conclusion.....	219
References	221
Tables and figures.....	230
Supplementary online materials	236

Supplementary files.....	242
7. Discussion.....	243
7.1 Chapter synthesis: main results	243
7.2 On the origin and genetic differentiation of Indonesian nutmegs: where and how?	246
7.3 On the genetic diversity of Indonesian nutmegs.....	250
7.4 On the population dynamics of nutmegs	253
7.5 Limitations on the regional understanding of the importance of genetic diversity.....	256
7.6 On the conservation and sustainable management: perspectives and the future	260
References	265
Appendixes.....	275

1. Introduction

1.1 Global context

There is nothing quite as captivating as the tapestry of biodiversity that envelops our planet. From its inception as “biological diversity” by Harris (1916), it has attracted our attention, spurring a quest to unravel the intricate ways in which biodiversity shapes living communities. As the world advances with unmaintained nature, concerns raised by Dasmann (1968), underscored the imperative necessity of safeguarding the natural riches of our world. In the realm of land plants, Myers (1988) introduced the concept of “hotspots”, paving the way for targeted conservation strategies. These “hotspots” highlighted regions of exceptional plant diversity that require special attention and conservation measures.

One of the most important hotspots regions is the Malay Archipelago. The Malay Archipelago contains three distinct biodiversity hotspots: Sundaland, Wallacea, and East Melanesian Islands (Hall, 2009a, 2009b, 2013; Figure 1). A staggering 15% of the world’s forest cover finds its home in this region (Stibig et al., 2014), underscoring its ecological importance. The forests of the Malay Archipelago are not only important in terms of their size but also in relation to the incredible diversity of plant species they are hosting (Cámara-Leret et al., 2020; Estoque et al., 2019), making them a crucial component of biodiversity.

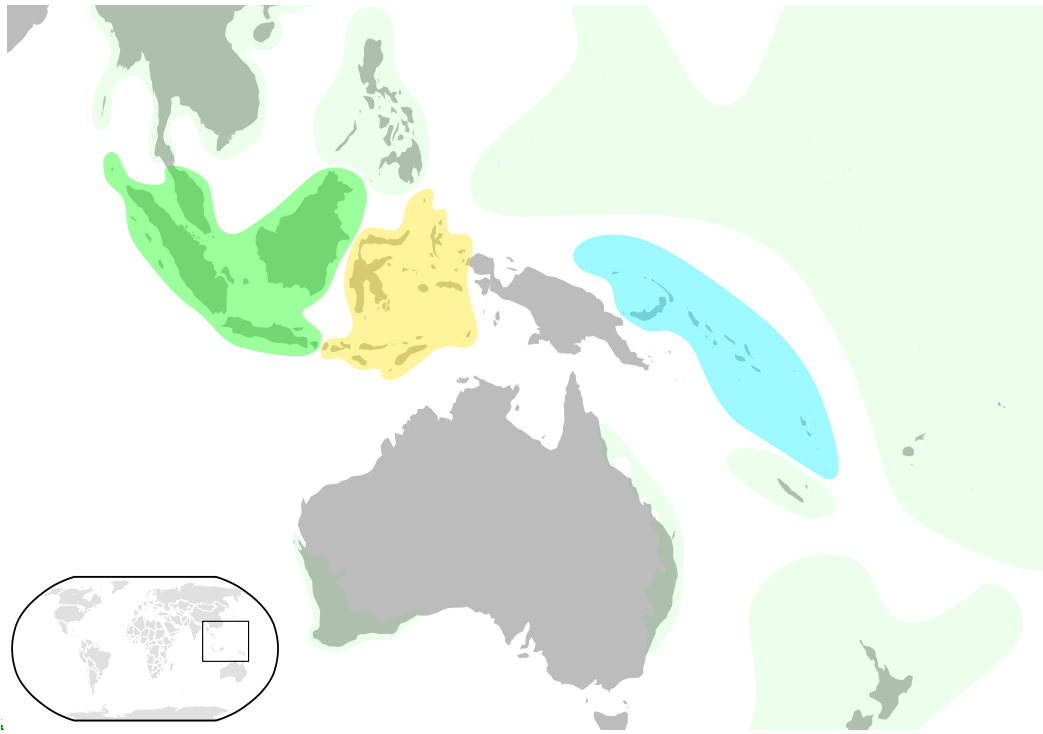


Figure 1. Malay Archipelago and its biodiversity hotspots. The Sundaland region is represented by the dark green color, the Wallacean region by yellow, and the East Melanesian Islands by blue. Light-green indicate other hotspots regions.

Despite its unique value, the biodiversity of the Malay Archipelago remains poorly documented. In a comparative analysis of research publications, various biodiversity hotspots, including the Congo Basin, African tropical forests, and the Amazonian Forest, were examined. The results were astounding, with the scarcity of biodiversity-related scientific literature dedicated to the Malay Archipelago is nothing short of startling (Figure 2). This research deficit calls for renewed efforts to delve into the region's unique ecosystems, unravel its hidden treasures, and shed light on its conservation needs.

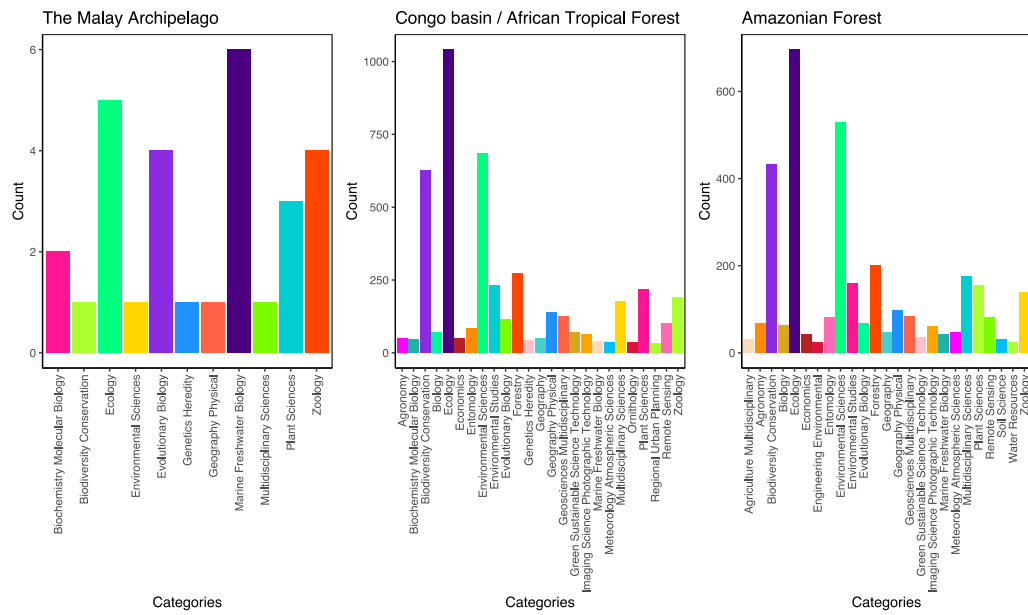


Figure 2. Comparative analysis of biodiversity studies and its derivatives in the Malay Archipelago, Congo Basin or African tropical forests, and the Amazon. The comparative analysis on biodiversity-related research was done in Web of Science and data plotted using R (R Core Team, 2022).

The forests of the Malay Archipelago present an exceptional opportunity for biodiversity research due to the rich array of plant species. Within these forests, numerous plant species contain unique substances that hold great significance for human use (Middleton et al., 2019). Historically, there was a huge interdependence between humans and the rainforests of the Malay Archipelago. It dates back at least fifteen centuries ago (Reid, 1995), with the forest acting as a primary source of food and sustenance for local communities. However, as human activities continue to shape the landscape and influence the utilization of forest tree species, it is inevitable that the composition of biodiversity in the region is undergoing transformative changes.

Among the forest tree species collected or harvested from natural forests, rather than being cultivated in agricultural settings, *Myristica fragrans* and *Myristica argentea* (referred as Indonesian nutmeg in this thesis) stand out as two

notable examples used for various purposes. The native distribution of the two species is geographically limited. *Myristica fragrans* is primarily found in the Wallacean region (notably the Moluccas archipelago), while *M. argentea* is native to Papua Island (de Wilde, 1995, 2014).

The narrow range of these insular species is characterized by a fragmented landscape consisting of discrete islands. As a consequence, the geographical isolation and small population sizes influence the distribution of diversity in both species, with potentially strong differentiation patterns among islands and low genetic diversity within each island (Barret, 1996; Frankham, 1996).

However, Frankham (1997) argued that certain island populations may nevertheless exhibit patterns of high genetic diversity thanks to high migration rates. There are several important considerations where migration rates are sufficiently high to counteract the effects of isolation and promote genetic diversity. In certain cases (see Tsumura & Ohba, 1993; Frankham, 1997), island populations have served as refuges during glacial periods, giving rise to smaller mainland populations that are derived from them. As a result, these refugia populations exhibit a higher genetic diversity compared to their continental counterpart. Thus, island species (such as Indonesian nutmegs, *M. fragrans* and *M. argentea*) offer unique insights to enhance our understanding in evolutionary processes.

Originally, *M. fragrans* and *M. argentea* were valuable to local communities, given their importance as spice species. They were initially not subject to large-scale commercial cultivation practices. Instead, they were typically harvested sustainably from their natural habitats. However, due to the growing demand for these species, they have been introduced and cultivated outside of their native range

since the 18th century (Brixius, 2015). The deliberate movement of plant material across geographic boundaries has facilitated the establishment of new populations in previously unoccupied areas. This human-mediated dispersal and the expansion of the species' range might have created new genetic lineages in these introduced populations.

Assessing the impact of human-mediated dispersal on gene flow in native ranges poses challenges due to the intricate interplay between natural and human-mediated dispersal processes. Disentangling the contributions of both processes becomes challenging when they are concurrently occurring (Einfeldt et al., 2020).

The geographical characteristic and cultivation management of these two species might provide paradigmatic examples on our understanding of the distribution of genetic diversity and evolutionary history. On the other hand, it has been suggested that human interaction with cultivated tree species will affect the species' evolution (see Cornille et al., 2015; Sánchez-Robles et al., 2022; Soares et al., 2019).

Therefore, the unique geographical characteristics, historical human interaction, and cultivation practices associated with *M. fragrans* and *M. argentea* offer valuable insights into the distribution of genetic diversity and evolutionary history of plant species in the region. Understanding the genetic dynamics of these species is crucial for their conservation and sustainable management in the face of ongoing environmental and anthropogenic changes. I contend that there is a need to investigate the precise mechanisms that have shaped the evolutionary history of the Indonesian nutmeg, taking into account the impacts of past climate fluctuations, geographical barriers, cultivation practices, and human interactions. By utilizing

the Indonesian nutmegs as a model species, we can gain valuable insights into how these factors have influenced their genetic diversity, population structure and dynamics throughout their evolutionary trajectory.

1.2 Biogeography of the Malay Archipelago

1.2.1 Geological history

The Malay Archipelago, an archipelagic sequence located between the Australian continent and mainland Asia, comprises approximately 25,000 islands (Figure 3). The emergence of the Malay Archipelago is a complex process that has occurred over millions of years. While the precise timeline is still subject of scientific debate, it is hypothesized that the fragments that make up the Malay Archipelago were rifted from the Gondwana supercontinent over the past 300 million years, and then were accreted to the Eurasian subduction margin (Hall, 2009a). This means that the archipelago did not emerge at a specific point in time but that the process of its formation took place over a period of 300 million years.



Figure 3. The Malay Archipelago (sensu Wallace, 1886), indicated in green color, situated upon the equator between mainland Asia and Australia. Credit pictures, with color modification from the original: Wikimedia Commons licensed under the Creative Commons Attribution-Share Alike 4.0 International license.

The geological origin of the Malay Archipelago can be understood through the three main continental shelves that make up the archipelago: the Sunda Shelf, the Sahul Shelf, and the Wallacean region (Figure 4A; Atkins et al., 2020; Haddon et al., 1892; Raes & Van Welzen, 2009; Van Welzen et al., 2011; Wallace, 1862, 1886). The Sunda Shelf is a region that includes the western and central parts of the archipelago, including the islands of Sumatra, Java, and Borneo. In this thesis, I will specifically concentrate on two prominent regions within the Malay Archipelago that are respectively the centers of origin of *M. fragrans* and *M. argentea*: the Wallacean region and the Sahul Shelf. The Wallacean region (Figure 4B) is a transitional zone between the Sunda and Sahul shelves. It includes the islands of Sulawesi, the Moluccas (locally known as “*Kepulauan Maluku*”), and the Lesser Sunda Islands. The Wallacean region is believed to have formed through a

combination of tectonic and climatic processes, including the collision of the Sunda and Sahul shelves and changes in sea levels.



Figure 4. A. The approximate positions of the Sahul and Sunda shelves during the Last Glacial Maximum (in dark yellow) and current time (in light yellow). This region, at the South margin of the Southeast Asian region, have been subdivided in three biogeographical areas. These areas, namely Sahul, Sunda, and Wallacea (sensu Van Welzen et al., 2011), each exhibit a distinct flora and fauna due to their unique geological and biogeographical histories. B. The Wallacean region, a group of islands indicated in light blue area. Credit pictures, with color modification from the original: Wikimedia Commons licensed under the Creative Commons Attribution-Share Alike 4.0 International license.

Since the past 10 million years, the island groups in the Moluccas have undergone significant tectonic movements, although the nature of these movements has varied between the North and South Moluccas. Tectonic and volcanic activities are believed to have played significant roles in the shaping and emergence of these islands. Several islands such as Ternate and Tidore in the North Moluccas have remained submerged until 3 to 5 million years ago (Hall, 2013). The geological processes responsible for the formation and emergence of the islands in the Moluccas, have not yet been fully elucidated.

According to Wallace's concept of 'Island life' (Wallace, 1880), the Wallacean region is characterized as a group of islands without any continental connection. This region is distinct from the neighboring Sunda Shelf due to the absence of Sundaic groups and the presence of numerous endemic and Australasian species (Welzen & Raes, 2011).

Three hypothetical lines (demarcation boundaries) have been defined in the Wallacean region, namely the Wallace Line, the Weber Line, and the Lydekker Line (for a historical overview see Wallace, 1862; Lydekker, 1896; Weber, 1902). They were initially proposed as hypothetical boundaries to differentiate faunal populations. However, van Welzen et al. (2005) revealed that these lines also hold significant importance in delineating floral distributions and boundaries, supporting the independently unique tectonic history of the region.

The Sahul Shelf, encompassing the eastern portion of the archipelago, is an intriguing region with a rich geological history closely tied to climate dynamics. It forms part of the expansive Australian Plate, which, according to geological evidence, underwent a significant separation from the Gondwana supercontinent

approximately 180 million years ago (Ben-Avraham, 1973; Hall, 2013). Within this region lies the Papua Island (Figure 5), alternatively known as the New Guinea Island, which was once directly connected to the Australian continent.



Figure 5. Papua Island, situated in the eastern part of the Malay Archipelago, were once connected to Australia during the Last Glacial Maximum. Credit pictures, with color modification from the original: Wikimedia Commons licensed under the Creative Commons Attribution-Share Alike 4.0 International license.

According to various reports (Hill & Gleadow, 1989; Weiland & Cloos, 1996), it is believed that the separation of Papua Island with Australia occurred relatively recently, possibly within the past few million years of Earth's history. Approximately 50 million years ago, the Indo-Australian continental plate underwent significant tectonic activity (Baldwin et al., 2012; Hill & Gleadow, 1989). During the late Miocene to early Pliocene, around 5 to 7 million years ago, a significant portion of western Papua Island was submerged due to rising sea levels (Gold et al., 2017). This dynamic geological environment has contributed to the

formation of diverse terrains on the island. The interaction between tectonic plates has resulted large-scale faults, and has facilitated the formation of the mountain range (Holm et al., 2015, 2016).

As the Pleistocene epoch progressed, there was a gradual rise in sea levels from ~80,000 to ~10,000 years ago (Fairbairn et al., 2006). These conditions have conducted to the isolation of Papua Island from the Australian mainland (O'Connell et al., 2010). After the end of the Pleistocene period, the island underwent further geological and environmental changes. There was a continued rise in sea levels globally that affected the coastal areas leading to the inundation of low-lying regions and the formation of new coastlines, particularly in the Bird's Head area of West Papua (Oliver et al., 2017).

The West Papua mountainous range has undergone active tectonic movements, resulting in the creation of multiple ancient river networks (Fairbairn et al., 2006). However, the exact role of these tectonic activities in shaping the coastal landscape is still not fully understood (Hill & Gleadow, 1989; Hill & Hall, 2003). Similar to the Wallacean area mentioned earlier, it is crucial to investigate the impact of the island's active landmass (portion of land that is geologically active) on its floral biodiversity. Extensive research is necessary to gain a deeper understanding of how tectonic activities, environmental changes, and the interplay between these factors have shaped the development of distinctive plant life on the island.

1.2.2 Floristic history – The Moluccas and Papua Island

The Moluccas Archipelago

As the region has one of the most complex geological histories in the world (Ben-Avraham, 1973; Metcalfe, 2001), the Malay Archipelago exhibits a high degree of floral complexity. This region is characterized by its unique biogeographical status, being part of both the Malesia and Papuasias biogeographic region (Figure 6; Brummit, 2001). In the eastern boundary of Malesia, the Wallacean region acted as a biogeographical midpoint between Asian and Australian flora.



Figure 6. Biogeographic regions of the Malay Archipelago. The archipelago, situated between mainland Asia and Australia, exhibits a unique biogeographical condition. This archipelago is divided into two distinct biogeographic realms, Malesia (indicated in green) and Papuasias (indicated in orange).

Based on generic distributions, van Steenis (1950) did not consider the Wallacean region as a distinct phytogeographical area. More recent studies based on van Welzen et al. (2003) have suggested the importance of considering the Wallacean region as a distinct phytogeographical unit rather than combining it with the neighboring Sunda and Sahul regions. Van Welzen et al. (2011) have elucidated the plant families that exhibit a central distribution in the Wallacean region, highlighting the presence of herbaceous and woody plants within species-rich families. Furthermore, their study emphasizes the notable strong dispersal capacity exhibited by these plant groups.

One of the major archipelagos in the Wallacean area is the Moluccas archipelago (locally known as Maluku; Figure 7). The distribution of plants in the Moluccas is greatly influenced by past climatic events (Trethowan et al., 2022), as they were highly sensitive to hydroclimatic changes (Wicaksono et al., 2017). The region has experienced significant fluctuations in sea levels, which have had a considerable impact on the amount and location of exposed land surface. During glacial periods, sea levels were up to 120 m lower, resulting in the connection of different Moluccas islands (Hanebuth et al., 2000). However, sea passages remained between major island groups in the Moluccas, thus maintaining dispersal barriers (Morley & Flenley, 1987). The events of sea level changes occurred continuously during the middle Miocene, early Pliocene, and the Holocene (Woodruff, 2003).



Figure 7. Map of the Moluccas archipelago and Papua Island. North Moluccas indicated in yellow, South Moluccas indicated in pink, and Papua Island indicated in light blue. Picture By ASDFGHJ - Own work, Public Domain, <https://commons.wikimedia.org/w/index.php?curid=7350841>

Although the Moluccas is considered as a single biogeographical unit (Welzen et al., 2005), Rutgrink et al. (2018) suggested that the Moluccas should be divided into two different biogeographic groups (the North Moluccas and the South Moluccas). Actually, these authors show that the North and South Moluccas differ greatly in plant species composition, sharing only 50%–65% of their species. Continuous dispersal barriers and complex tectonic history of the islands might explain these differences.

Moluccas might be considered as ‘true oceanic islands’ because they probably have had no terrestrial connection to any surrounding land since their

emergence. Stuessy et al. (2014) proposed that factors influencing the genetic variation of plant species on oceanic islands are physical characteristics of the islands, biological features of the species, modes of speciation, hybridization and human-caused disturbances. The dispersal of plants by wind and bird/bat, coupled with the seasonal sea currents (Barmawidjaja et al., 1993), are not likely to get across sea barriers (Rutgrink et al., 2018). Thus, vicariance speciation is likely to have played a significant role in the formation of its floral diversity, since the presence of sea barriers has likely prevented long-term gene flow between populations. This isolation, combined with other factors such as founder events, may have led to the formation of new species through the process of allopatric speciation.

The origin of the floral diversity in the Moluccas is still unknown. The South Moluccas seems to be more diverse than the North Moluccas (1440 vs. 1133 species, Rutgrink et al., 2018). The first floral documentation in the Moluccas was done by Rumphius (1755a, 1755b), but he was mainly focusing on one particular island (Ambon) in the Southern Moluccas. Recent investigations (see Irmayanti et al., 2022) on an island in the Northern Moluccas (Bacan) reveals 40 tree species from 25 families, dominated by *Neolamarckia macrophylla*, *Celtis* spp., and *Aglaia* spp. In spite of an extensive search, no other literature on the floral composition of the Moluccas archipelago could be found.

Despite the importance of the Moluccas archipelago for the study of biogeographic patterns in the Malay Archipelago, its floral composition remains poorly documented. Compared to other regions in the world, the Moluccas have received limited attention from botanical researchers. Furthermore, the few

available data on the Moluccan flora are often fragmentary, making it difficult to draw definitive conclusions about the distribution and evolution of plants in this area. It is clear that further investigation is needed to fully understand the unique plant diversity and evolutionary history of this important region. Improved documentation of the Moluccan flora will not only provide insights into the evolution of plants in the region but will also inform conservation efforts for this unique and valuable ecosystem.

Papua Island

The geological origin of Papua Island is different from Moluccas archipelago. The gradual isolation of Papua Island during the Pleistocene fostered unique evolutionary processes, allowing for the development of diverse and endemic species adapted to local environmental conditions. This period of geographic and ecological separation played a crucial role in shaping the present-day biodiversity and ecological dynamics observed on Papua Island (Hartemink, 2010; Heads, 2001).

As the second largest island on Earth and the largest in the tropical region, Papua Island carries one of the most diverse and abundant floras on the planet (Cámara-Leret et al., 2020). According to this study, Papua Island would be ‘the world’s most floristically diverse island’, with an astonishing number of ~13,000 species. This is 19% larger than the number of species recorded in Madagascar island. This floral diversity results from various factors, including its past geological events. As a consequence of island’s notable tectonic activity, it gave rise to the emergence of mountain ranges that serve as both barriers and corridors

for the dispersal of plants. Heads (2001) proposed that there is a correlation between the tectonic history of Papua Island and its floral diversity, revealing the presence of numerous micro-endemic species within the mountainous ranges. These geological processes, coupled with changing sea levels during past glacial and interglacial periods, have influenced the migration and adaptation of plant species on the island.

The diverse climatic conditions in Papua Island encompass a range of factors, including variations in monsoons and the dynamic interplay between the Pacific and Indian Oceans (McAlpine, 1983). This has led to the development of diverse vegetation types, ranging from tropical rainforests in lowland areas to montane forests and alpine vegetation in the highlands. Throughout the Pleistocene period, the rise in temperatures gave rise to distinctive vegetation types characterized by a multitude of broadleaf species (Fairbairn et al., 2006).

The evolutionary history of Papua Island's flora represents a captivating enigma, with numerous unanswered questions. It remains unclear what are the forces behind their diversification and adaptation to the island's unique environment. Understanding the evolutionary trajectories of Papua Island's plant species is not only intellectually stimulating but also crucial for effective conservation strategies. This knowledge can guide conservation efforts aimed at preserving the island's rich biological heritage.

1.3 Spatio-temporal dynamics of nutmeg: the emblematic spice tree species

Myristicaceae are a family of flowering plants containing 21 genera native to Africa, Asia, Pacific islands, and the Americas. Among the genera found in the

Myristicaceae family, the genus *Myristica* contains the highest number of species, with around 170 species described to date in Indonesia and neighboring countries (de Wilde, 2014). *Myristica* spp. present a considerable ecological and ethnobotanical significance, as they are important food for many animals and are harvested by humans for timber (Soepadmo et al., 2004), spices, dart/arrow poison, medicine (Abourashed & El-Alfy, 2016) and a hallucinogenic snuff employed in medico-religious ceremonies (Steeves, 2011). As indicated by de Wilde (2014), the Malay Archipelago serves as the primary hub for *Myristica* species, with a significant concentration found within the Moluccas and Papua Island.

1.3.1 *Myristica fragrans* Houtt.

The most important species in the genus *Myristica* is nutmeg (*Myristica fragrans* Houtt.; Figure 7). The tree is usually 5–15 m tall, but can occasionally reach 20 or 30 m (de Wilde, 2014; Houttuyn et al., 1774; USDA, 2023). Nutmeg is a dioecious and entomophilous tree that thrives in warm and humid conditions with an annual rainfall of at least 150 cm. *Myristica fragrans* can grow up to 1300 meters above sea level and begin to yield fruit eight years after sowing. After reaching their optimal age of production around 25 years old, nutmeg trees continue to bear fruits for at least 60 years.

The fruit of *M. fragrans* is a strikingly beautiful pendulous drupe, resembling an apricot in appearance (Figure 8). When the fruit is fully mature, it opens from the base, revealing a brilliant crimson-colored aril, known as mace, which surrounds a single shiny brown seed, the nutmeg. The nutmeg seed is particularly

prized for its distinct aroma and flavor, making it a valued ingredient in numerous culinary dishes and spice blends around the world.

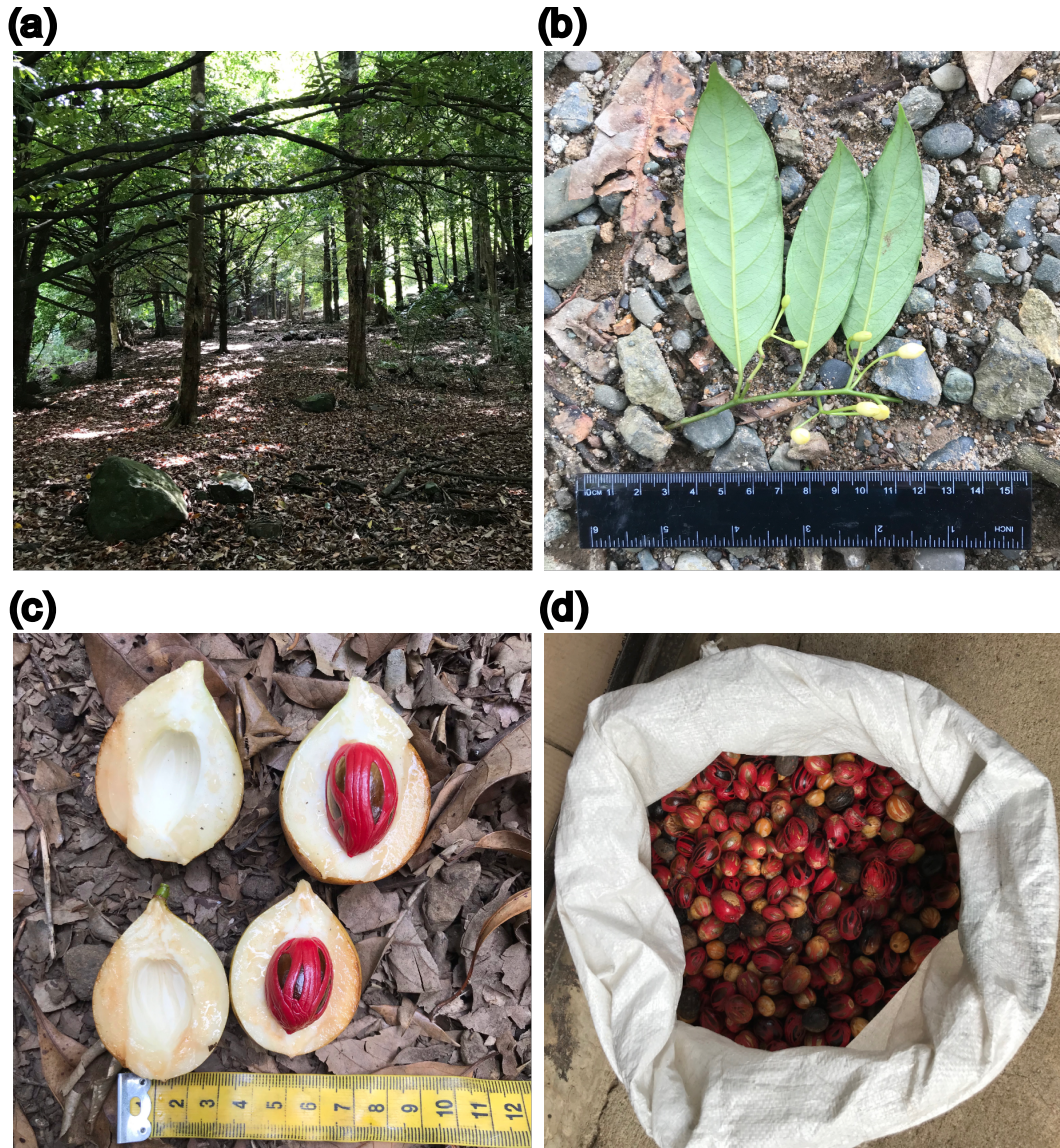


Figure 8. *Myristica fragrans* in Ternate, North Moluccas (a). Leaves and male flower features (b). Longitudinal cut of nutmeg fruit, exposing its seed covered with red aril (c). Seed and aril of nutmeg harvested by local farmers in Aceh, Sumatra Island (d). They will sort and grade the nutmeg before separating them into two different products, the nut (seed) and the mace (aril). © Jaky Kusuma, 2019.

Myristica fragrans is widely cultivated across the tropics notably in China, Taiwan, Indonesia, Malaysia, Grenada, India, Sri Lanka and South America (Figure 9). The native range is believed to be in the Moluccas. Rumphius (1755b) believed

that the species has its origin in the Banda archipelago in the Moluccas, although he mentioned that nutmeg might also be present in the whole Moluccas. It appears that this hypothetical origin in the Banda archipelago has been adopted uncritically by all authors since this first description (see for example de Wilde, 2014). An archaeological survey recently demonstrated that nutmeg was already used by humans 3,000 years ago in Ay (or Ai) island, Banda archipelago (Lape et al., 2018).



Figure 9. The recent worldwide distribution of *Myristica fragrans* indicated with black dots. It is widely believed that the native distribution of *M. fragrans* is the Wallacean region in the Malay Archipelago (de Wilde, 2014). Note that distribution in the temperate region might correspond to national botanical gardens. Data retrieved from GBIF (2023).

The history of its native distribution is difficult to find. The earliest records dated back from Constantinople in the 6th century (Weiss, 2002). The Chinese, Arabs, and Indians contributed on trading the nutmeg to Europeans since 7th to early 13th century (Loewe, 1971). This uncertainty of the cradle place is supported by some 13th century Chinese writings, where “Banda” is mentioned, but where the word seemed to refer to the whole area of eastern Indonesia rather than to the Banda Island itself (Rockhill, 1914). Andaya (1991) also reported that the Moluccans took their spices (including nutmeg) from many islands during the local trade on spice

in 14th-15th century. Centuries later, it is known that Banda archipelago remain the ultimate source of nutmeg, spreading the products to Java (mainly Batavia; now called Jakarta, the capital of Indonesia) before being exported to other parts of the world through the Silk Road (Barjiyah, 2011; Mansyur, 2017; Taim, 2022).

Due to its very high commercial value during the medieval era, the Europeans started to obtain the direct source of nutmeg in the Moluccas, pioneered by the Portuguese in the 16th century (Ptak, 1992). Subsequently, with the arrival of Dutch traders in Banda Island, stringent measures were implemented to control the number of nutmeg trees, by frequent extirpation practices (Wright, 1958). These measures were taken to establish a monopoly on the resource and ensure that prices of nutmeg remained high. Thus, the nutmeg tree population from Banda archipelago was strongly managed by the Dutch East India Company (*Vereenigde Oostindische Compagnie*; VOC) for three centuries long (Bentley & Trimen, 1880). VOC was the public-trading company established by the Dutch government to manage spice cultivation and trade in Asia.

Pierre Poivre in the 18th century was one of the first to attempt to propagate *M. fragrans* outside of its native distribution. According to Brixius (2015), Poivre proposed to the *Compagnie des Indes Orientales* (CIO) to collect spices from different parts of the world in 1748. Poivre was determined to smuggle nutmeg from the Dutch-possessed Moluccas and introduce these spices to the Bourbon Island (île de la Réunion) with the aim of cultivating them for the French spice trade. However, the cultivation of nutmeg in Bourbon Island was not as successful as Poivre hoped to, leading to the importation of nutmeg from other regions such as Grenada in the Caribbean. It remains unclear how cultivators were instructed

during the first attempts of acclimatization. Yet, we do know that, prior to Poivre's departure in 1773, he printed and circulated '*Instruction sur la manière de planter et de cultiver avec succès les plantes et les graines de Géroflie et de Muscadier*' to local colonists.

Since the era of Spice Trade and through centuries after, this species has experienced various management practices. The cultivation of nutmeg in Indonesia is still using traditional methods. Particularly in the Moluccas, *M. fragrans* is treated as a wild tree in abandoned or neglected plantations, or was intentionally planted in wild forest (de Wilde, 2014). To upscale the production, breeding practices and vegetative propagation were applied during the middle of the 20th century (Flach, 1966). Being a dioecious species, farmers tend to avoid growing male trees as they do not produce fruits. To date, farmers or 'tree owners' often cut down male trees to reach a 1:10 male/female ratio in their plantations (Figure 10).

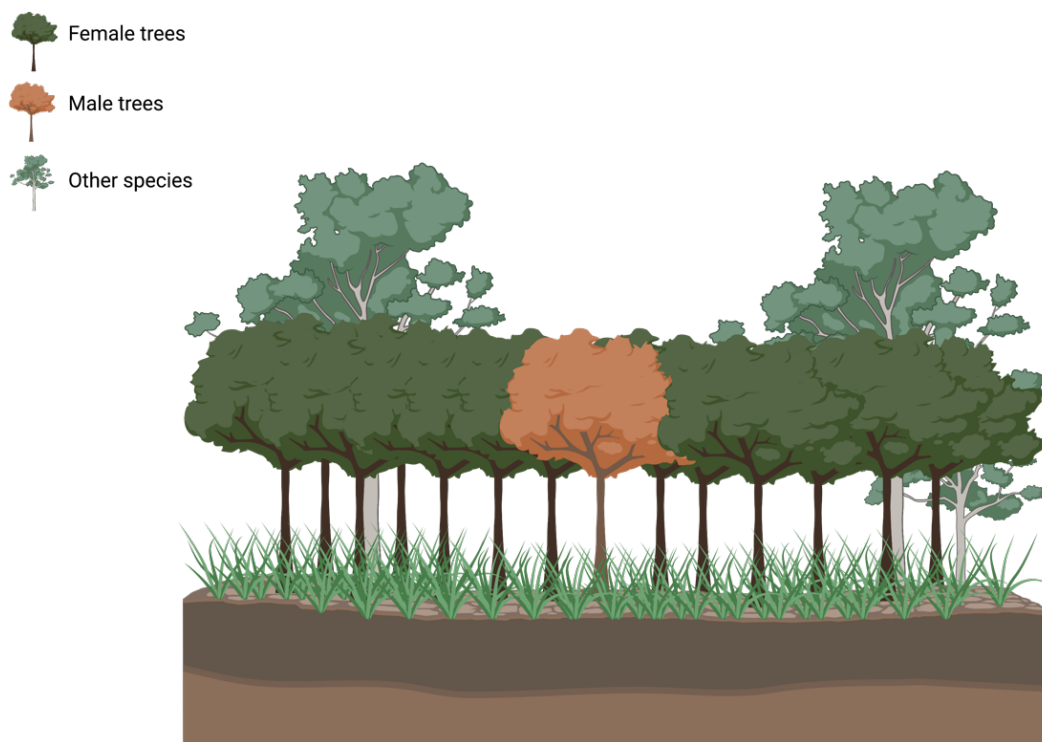


Figure 10. The figure illustrates the composition of a typical nutmeg forest plantation. It often involves intentional planting by farmers or the utilization of existing nutmeg trees within the forest.

Upon discovering a new forest area with nutmeg trees, farmers will typically assess the sex composition and actively limit the proportion of male trees. Every nutmeg tree often intercropped with other species (e.g., cloves, cinnamon, coconut, or *Ficus*) that are taller than the nutmeg trees and provide shade.

Some of nutmeg farmers, called ‘seedling seller’, collect seeds from a single highly-producing mother tree, grow seedlings in a nursery and sell them to other islands. In a specific administrative area such as a village, there is typically one nutmeg forest that serves as a seed source, and a particular mother tree can be identified as the main contributor (Rifany, 2016). Nowadays, as a result of their economic value, the species has been widely propagated, introduced, and mostly cultivated in wildcrafting system in Indonesia and across the tropical regions in the world.

The nutmeg is probably the most socio-economically important species from the *Myristica* genus. Its mace and nut are utilized as spices, medicine, and in the perfume industry. In Indonesia alone, production quantity is reaching around 40,000 tons/year (Ministry of Agriculture of the Republic of Indonesia, 2023). *Myristica fragrans* has contributed to more than US\$ 111,000,000 to smallholder farmers in 2021 (Central Bureau of Statistics of the Republic of Indonesia, 2023). Despite of its popularity, the nutmeg industry is still facing severe supply problems in both raw material and oil products (Juwita & Tsuchida, 2017), with a shortage of around 3,000 tons per year (Central Bureau of Statistics of the Republic of Indonesia, 2019).

1.3.2 *Myristica argentea* Warb.

To fulfill the nut and mace demand on a local and national scale, other species are utilized by farmers, even if they are recognized as a lower grade quality

compared to *M. fragrans*. *Myristica argentea*, widely known as Papuan nutmeg, is one of them. It has a slighter importance compared to *M. fragrans*. The species is probably native from Papua Island (de Wilde, 2014). Despite its low popularity compared to *M. fragrans*, this species became a leading commodity in West Papua (Musaad et al., 2016), and has started to be cultivated outside Papua. Similar to *M. fragrans*, *M. argentea*, or Papuan nutmeg, is a dioecious species that starts bearing fruits at the age of 7 years (Figure 11). The species' fruit stands out with its larger size compared to *M. fragrans*, showcasing an egg-shaped form instead of a round one. Traditionally, it served as a substitute when the yield of *M. fragrans* could not meet the demand. However, in present times, *M. argentea* has carved out its own distinct market segment. The taste of *M. argentea* fruit flesh is also sweeter and does not leave a bitter aftertaste. As a result, *M. argentea* yields two distinct products, specifically the fruit and its husk.

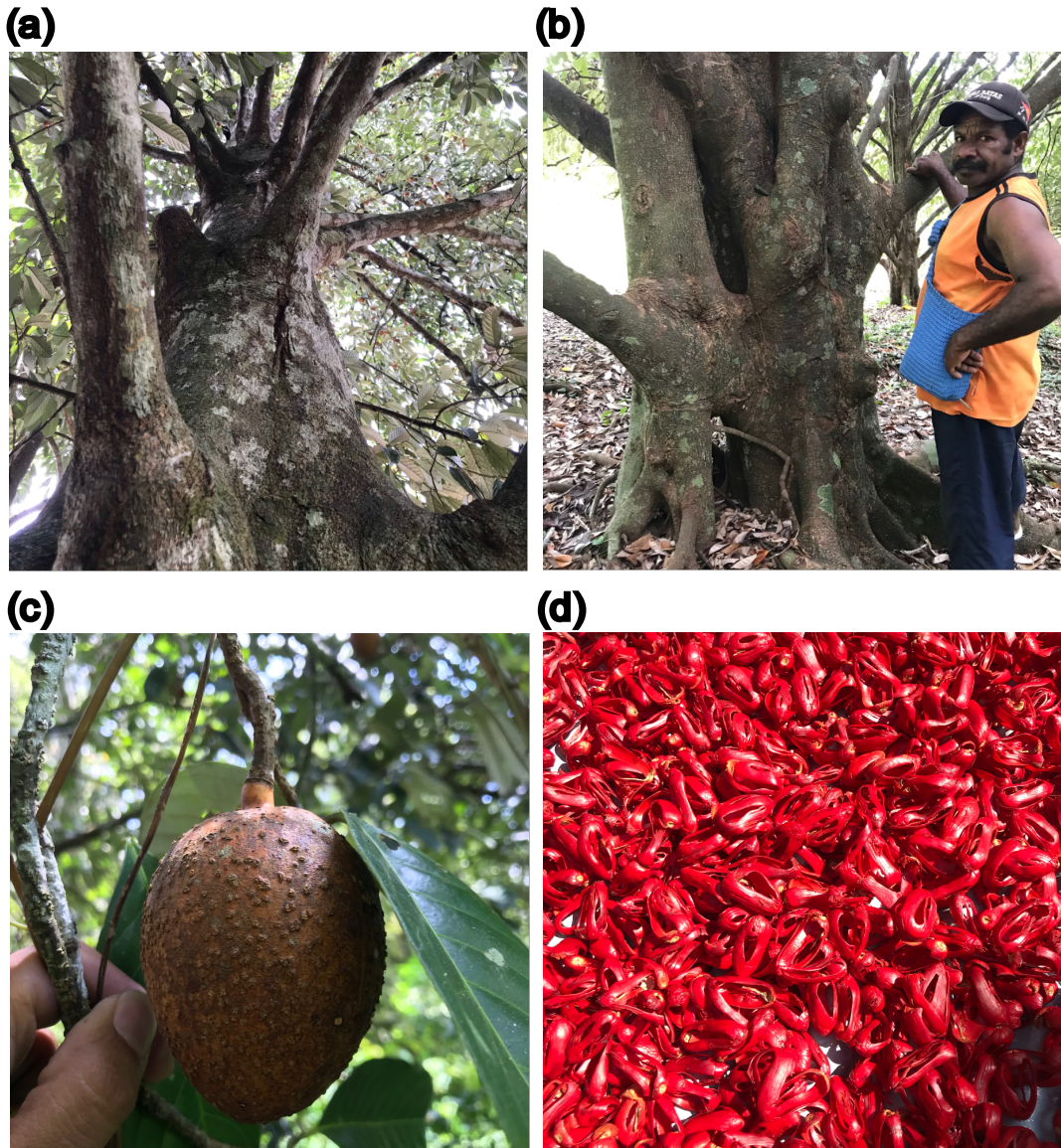


Figure 11. An old *Myristica argentea* tree in a Papuan forest (a). A Papuan nutmeg farmer with his old tree (b). Ripe fruit (c) and harvested mace (d) of Papuan nutmeg. © Jakty Kusuma, 2019.

The highest production is achieved around the age of 25 years, continuing until the plants are aged 60 - 70 years or more. Fruit harvest occurs several times a year, predominantly in March - April, June - July, and November - December, with a smaller harvest occurring outside these months. The range of fruit production is 450 – 4,750 fruits per tree, with an average production of 2,000 fruits per tree per

year. To maximize yield, trees with production rate of >2,000 are selected as seed sources (Wahyuni & Bermawie, 2020).

Despite its significance, botanical studies on this species remain elusive. There is very little information available about the biology, ecology, and ethnobotanical uses of *M. argentea*, making it challenging to establish sustainable management and use strategies. The majority of information found on the internet relies on interviews or local newspaper sources.

Myristica argentea is endemic to the Fakfak peninsula, in the Western Papua Island (or the Bird's Head region, Figure 12). Despite the scarcity of information, there is some available data on the dispersal mechanisms of this species. Correlated with the high number of bird species in the region (Mongabay, 2019), the natural dispersal of Papuan nutmeg seeds has been facilitated by the activities of birds over the course of centuries. The main dispersal agent for this species is *Rhyticeros plicatus* (Bucerotidae) or Blyth's Hornbill (BBKSDA, 2022) and the migration pattern of this bird is related with season in the West Papua (Kemp & Kirwan, 2020).

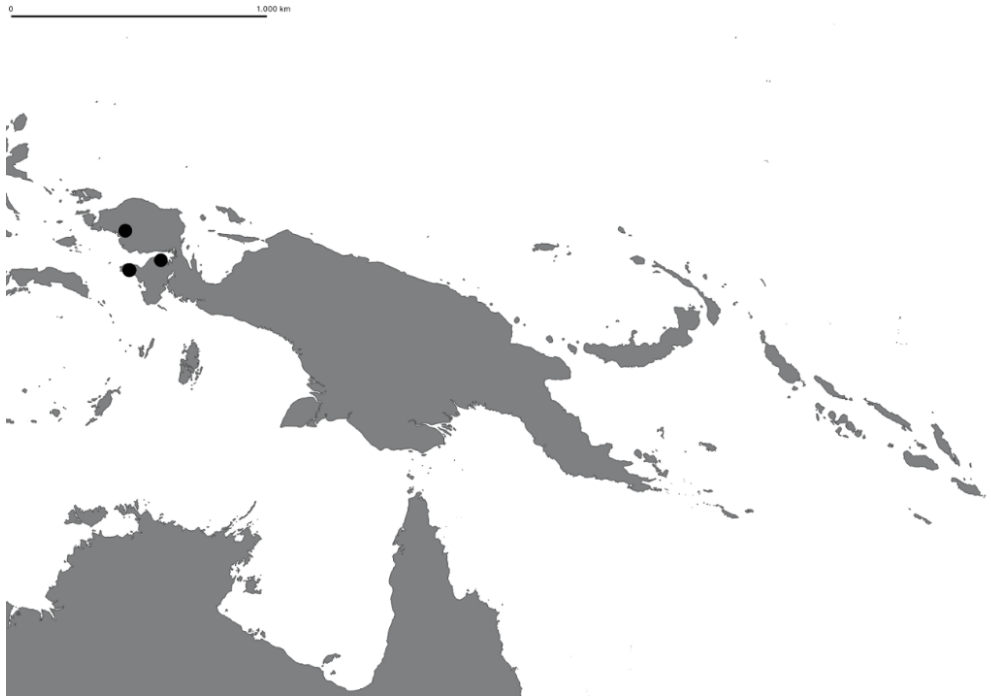


Figure 12. The distribution of *Myristica argentea* in the Western Papua Island (or the Bird's Head region) indicated with black dot. De Wilde (de Wilde, 1995, 2014) suggest that the native distribution of *M. argentea* is Fakfak Peninsula, West Papua. He also noted that the species is endemic in the Fakfak Peninsula. Data retrieved from GBIF (2023a).

The Fakfak people have long recognized and made use of the root characteristics of *M. argentea* as a proactive measure to mitigate the risks of landslides and flooding in their hilly region. By firmly anchoring itself into the soil, this species helps maintain the integrity of the land, preventing potentially devastating slides and protecting vulnerable areas from the ravages of flooding. The profound understanding and utilization of the species' root system by the Fakfak people exemplify their deep-rooted knowledge of their environment and their resourceful approach to harmoniously coexist with the natural forces that shape their landscape.

Today, Papuan nutmeg continues to be a symbol of heritage and cultural significance, particularly in the Fakfak district. Its cultivation and trade remain important economic activities for local farmers and communities. The stories

shared by farmers and the long-standing tradition of nutmeg exports highlight the enduring legacy of this spice and its impact on the livelihoods of those involved in its production. In order to enhance the production of Papuan nutmeg and expand its value chain, the local Indonesian government has undertaken strategic measures. A key strategy involves increasing the planting area within the native range of this valuable species (Watts, 2021).

1.4 Evolutionary context

Charles Darwin discovered the variation within species in an island, the Galapagos archipelago (1859). This ignited an interest in understanding the mechanisms that generate and influence the evolution of this variation. With the discovery of DNA as the hereditary molecule (Watson & Crick, 1953), scientists were able to unravel the molecular underpinnings of evolutionary change. This allowed for a greater understanding of how genetic units within a species contribute to the emergence of new lineages and the diversification of life on earth, particularly in the island.

Among other living organisms, plants have a set of life-history traits that distinguish them from other life forms. Trees in particular, have a large stature and are long-lived species that are capable of producing abundant quantities of seeds throughout their lifetime (Moles et al., 2007; Qiu et al., 2022).

Trees are also capable to outcross with geographically distant individuals aided by wind, water, and animals (Elleouet & Aitken, 2019; Moreno et al., 2022; Nakanishi et al., 2021; Solís-Hernández & Fuchs, 2019; Sujji et al., 2021; Williams & Barnéoud, 2021). Their fruit and seed characteristics are noteworthy, as they can

attract frugivorous animals. This can lead to a possible autochory and hydrochory dispersal (Hampe, 2004; Nilsson et al., 2010).

Trees, by providing timber and Non-Timber Forest Products (NTFPs), significantly contribute to human wellbeing, notably for health, nutrition, cultural values, income, energy, and shelter (Shackleton et al., 2020; Timko et al., 2010). They also play a vital role in broader ecological functioning, indirectly contributing to human wellbeing (Shackleton et al., 2018).

Those above-mentioned characteristics can have significant impacts on the distribution of genetic diversity of tree populations. As a consequence, trees are one of the most interesting groups for studying the mechanisms driving micro-evolutionary processes and population genetics (Laikre et al., 2020; Razgour et al., 2019).

1.5 Genetic diversity in tree species

1.5.1 A micro-evolutionary perspective

Mutation

Genetic diversity originates from mutation, which is the process by which changes occur in the DNA sequence of a genome. Mutations can occur spontaneously as a result of natural processes, and they can be induced by external factors such as radiation, chemicals, or errors during DNA replication. They are the primary source of genetic variation within a population (Lee et al., 2012; Y. O. Zhu et al., 2014).

The assessment of mutation rates, their types, and molecular spectra plays a crucial role in advancing our understanding of evolutionary questions. This

includes determining divergence times (Kumar, 2005), exploring population genetic diversity and effective population size (Johnson & Barton, 2005; Lynch, 2010; Sung et al., 2012), studying intra-organism mutation rate variations (Wang et al., 2019), and investigating adaptation to specific environments (Kondrashov & Kondrashov, 2010).

Differences in mutation rates among flowering plants can be attributed to various factors, including variations in life forms, plant height, generation times, genome size, and environmental conditions (Gaut et al., 1992; Lanfear et al., 2013; Smith & Donoghue, 2008). These variations concern both nuclear and organellar genomes (Bromham et al., 2015). In nuclear genes, the mutation rate differs between coding and non-coding regions. Moreover, studies conducted by de la Torre et al. (2017) suggest a mutation rate of $5.35e^{-9}$ in Angiosperms, while a different mutation rate of $2.5e^{-8}$ has been observed in Palmaceae (Gros-Balthazard et al., 2017), indicating variations in mutation rates among different plant groups.

Gene flow

The distribution of genetic diversity is also influenced by gene flow. This gene exchange contributes to the genetic diversity within populations, and can notably result in the transfer of certain alleles from one population to another. Gene flow can occur through dispersal of pollen and seed with agent like wind, insects, birds, or other animals (Bezemer et al., 2019; Grant et al., 2019; Wessinger, 2021).

Trees are often capable of long-distance gene flow. Several studies suggests that various dispersal agent are responsible for long-distance gene flow in trees (Kremer et al., 2012; Latouche-Hallé et al., 2004; Williams, 2010a, 2010b).

According to Nathan et al. (2008), there are six primary mechanisms through which long-distance seed-mediated gene flow occurs in plants: (1) open terrestrial landscapes, (2) large animals, (3) migratory animals, (4) meteorological events, (5) ocean currents, and (6) human transportation.

Human-mediated propagation of plant species, such as introduction to a new habitat, habitat restoration, translocations, or assisted migration, can influence gene flow patterns. Cultivated species can be introduced to new geographical areas/habitats outside of their original distribution. These human activities can artificially facilitate gene flow as individuals from different populations are intentionally or unintentionally mixed through human intervention (see Aitken & Whitlock, 2013; Krutovsky et al., 2012). This mixing can occur through various means such as the intentional transfer of seeds or the transplantation of individuals to new locations. Theoretically, the pattern in human-mediated gene flow can be observed. However, this pattern is still poorly documented.

In an example of a Lauraceae tree species (Kameyama et al., 2017), the pattern of human-mediated gene flow is observed in this cultivated tree population. According to Kameyama et al. (2017), distinguishing the impacts of human activities from natural processes is often challenging due to several reasons. Firstly, human activities and natural processes are interconnected and influence each other, making it difficult to separate their effects. Secondly, human-mediated translocation of species often lacks historical documentation, making it challenging to trace their origins accurately. Lastly, extensive gene flow within species further complicates the assignment of geographical origin (see Bagnoli et al., 2009; L. Baldwin et al., 2006; Petit, 2004).

Genetic drift

In natural conditions, habitat fragmentation could result from environmental characteristics such as climate changes (e.g. last glaciation). It can occur through a punctual event, such as a volcanic eruption (Bemmels et al., 2022; Moreno-Gonzalez, 2022) and colonization through long-distance dispersal (Alsos et al., 2015) in the context of oceanic islands. These mechanisms create isolated pockets of surviving individuals with a decrease of genetic diversity. One key evolutionary force influenced by these activities is genetic drift, which refers to random changes in gene frequencies within a population due to chance events. Genetic drift can lead to substantial loss of genetic variation, especially in small populations experiencing population bottlenecks, thereby limiting their genetic capacity to adapt to environmental changes.

In natural populations, a founder effect, i.e., the limited genetic diversity of founding individuals compared to the source population, is observed where a small group of individuals establishes a new population in a previously unoccupied area. In the context of oceanic islands, the intensity of the founder effect is influenced by factors such as distance to the source region, frequency of dispersal events, and species traits related to dispersal (Alsos et al., 2015). In tree species, Austerlitz et al. (2000) discovered that the impact of the founder effect is significantly diminished due to the trees' long-life cycle, even in cases of limited pollen flow. This phenomenon can be attributed to the following explanation: upon the establishment of a tree population, initial growth during the early years primarily relies on the influx of new juvenile migrants (seed flow) rather than reproduction (Austerlitz et al., 2000).

In emerging islands formed through volcanic activities, such as Hawaii (Wilding et al., 2023) or Ternate and Tidore in the Wallacean region (Faral et al., 2022), genetic drift may also occur. The founder effect can be observed in a population that experiences a volcanic eruption. During a volcanic eruption, the intense heat, ash, and lava can cause widespread destruction of the existing habitat, leading to the displacement or even extinction of the original population. In such cases, the surviving individuals that manage to establish new populations in the aftermath of the eruption are considered the founders of these newly established populations (Figure 12b; Gallardo et al., 1995; Vandergast et al., 2004).

On the other hand, anthropogenic activities, such as agriculture, can have a significant impact on the distribution of genetic diversity in plant species. For instance, the introduction of forest species outside their native range has been shown to illustrate these effects (Figure 13a; Brundu & Richardson, 2016; Dickie et al., 2014).

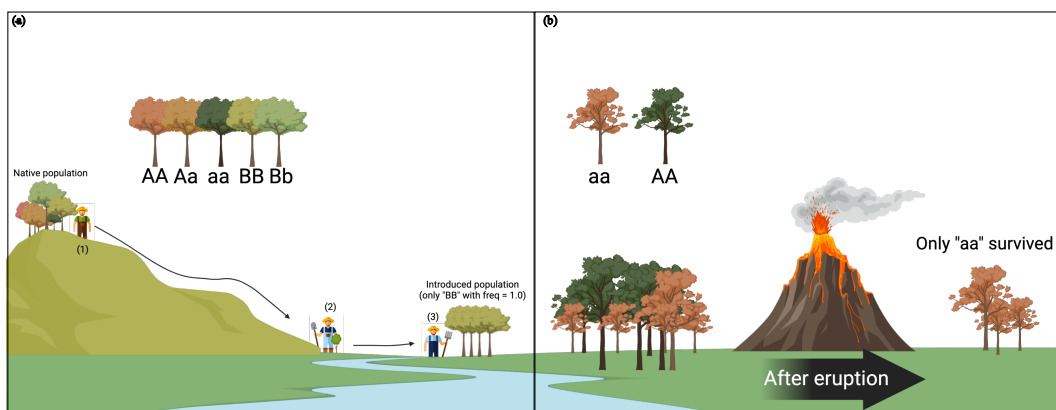


Figure 13. Founding event process that generates genetic drift, illustrated with human activities (a) and natural events (b). Individual colored trees represent the proportion of their allele frequencies. Human intentional selection of preferred traits in a group of tree individuals (1), followed by planting them outside the native range (2), can result in the loss of alleles in the introduced population after multiple generations (3). Meanwhile, natural events such as volcanic eruption could eradicate large proportion of allele frequencies and leaving small percentage of alleles in the following generations. Created with <https://BioRender.com>.

Newly introduced populations might exhibit a decrease in genetic diversity due to a founder effect. But in some cases (e.g., Thompson et al., 2016), genetic diversity could increase at the later stages of population establishment, if multiple introductions occur (i.e. subsequent gene flow). As an example, the introduced populations of canarywood survived the founding event and gradually increased its genetic diversity after several generations (Cordeiro et al. 2019).

Furthermore, extensive human activities in the forest (e.g., deforestation, land clearing, etc.) can lead to breaking down continuous forest into smaller and isolated fragments. Forest fragmentation can cause genetic drift in tree populations in several ways. First, when a forest is fragmented, the size and spatial arrangement of tree populations change. Second, forest fragmentation reduces connectivity between populations (Fountain et al., 2016). As a result, smaller/geographically-isolated populations will be more sensitive to genetic drift. It is widely reported that loss of genetic diversity through genetic drift is common in a fragmented population (Cueva-Agila et al., 2019; Fountain et al., 2016; Sánchez-Robles et al., 2022). The evolution of genetic diversity due to forest fragmentation is often linked to human activities, (e.g., logging, see Soares et al., 2019), but it is worth noting that forest fragmentation might also occur naturally. The impact of these environmental events and human activities will thus have a strong impact on species' diversity and evolutionary history.

Selection

Another significant factor that influences species' genetic diversity is natural and artificial selection. Natural selection acts on genetic variation that exists within

populations and can drive the evolution of adaptive traits. Selection can be driven by various abiotic factors (climate, soil, etc.), and biotic factors (Parker et al., 2003). For example, in temperate forests, natural selection can play a significant role in shaping the distribution of genetic diversity in plant species due to the strong selection pressures imposed by seasonal changes in temperature and precipitation (Hu et al., 2022). Similarly, in tropical forests, natural selection can be influenced by factors such as frugivorous birds (Assis et al., 2022; Lavabre et al., 2016). For example, the preference of frugivorous birds for fruits with certain traits (e.g. color, size, etc.) creates a selection pressure favoring trees that produce fruits with those desired characteristics. Thus, individuals that produce fruits with desirable traits are more likely to survive and reproduce, leading to the evolution of those traits over time.

In contrast to natural selection, artificial selection is a deliberate process where humans choose which populations or individuals are allowed to reproduce, resulting in cultivated populations that diverge morphologically and genetically from their wild progenitors. This phenomenon is also known as the domestication syndrome (Zohary et al., 2012). In trees, domestication processes mainly aimed to target particular traits such as vegetative characteristics (Anjarwalla et al., 2017; Cornille et al., 2019; Ferrer et al., 2020; Julca et al., 2020; Mañourová et al., 2019).

When human intentionally select and cultivate plants, it leads to genetic and phenotypic changes that distinguish them from their wild ancestors. Over time, domesticated plants undergo genetic divergence from their wild counterparts due to human-driven selective pressures. Stetter et al. (2017) suggests that while a significant portion of the early selection exerted by humans on wild species may

have been unintentional, as the domestication process progresses, intentional selection by humans becomes a prominent factor.

During the process of domestication, selection has primarily targeted polygenic variation rather than focusing on a single gene or locus with a large effect (Meyer & Purugganan, 2013; Purugganan, 2022). Instead of favoring specific alleles at individual genes, selection during domestication has acted on the combined effects of multiple genes, resulting in shifts in the average phenotype of the domesticated population. This means that the overall expression of traits has been altered, leading to observable differences between domesticated plants and their wild ancestors (Stetter et al., 2017). While selection has caused changes in the average phenotype, it has not completely eliminated genetic variation within domesticated populations. Despite the selective pressures exerted during domestication, some degree of genetic variation is likely to persist within domesticated populations.

The result of the domestication process can be seen in the cultivated plants that exhibit notable distinction from their wild ancestor. Examples of such divergence can be observed in species of the grasses family such as rice (Sang & Ge, 2007), maize (van Heerwaarden et al., 2011), and wheat (Charmet, 2011). These cultivated species exhibit distinct traits that set them apart from their wild ancestors. These differences encompass a range of characteristics including grain size, shape, plant architecture, and texture in rice (Choi et al., 2017; Londo et al., 2006), seed-bearing ears and kernel properties in maize (Andersson & Purugganan, 2022), and seed size and spike morphology in wheat (Leigh et al., 2022; Zeibig et al., 2022). Through deliberate improvement efforts, domesticated plants can

develop a reliance on humans for their survival. This is due to the absence of adaptations required to compete effectively with wild plants or overcome physiological limitations that impede their reproductive capabilities (Hufford et al., 2012).

The focus on grass plants in studies of domestication is largely due to their life history traits and the relatively short generation times associated with annual grasses. Annual plants complete their life cycle within a single year, allowing for rapid accumulation of genetic changes over a few generations. This accelerated pace of evolution makes it easier to observe and study the effects of artificial selection and detect clear patterns of domestication.

Trees have much longer lifespans and slower generation times compared to annual plants. They often take several years or even decades to reach reproductive maturity, making it challenging to study their domestication processes. The prolonged time required for trees to reach reproductive age, as well as their complex mating systems and larger genome sizes, contribute to slower rates of evolutionary change and less pronounced genetic divergence between cultivated and wild tree populations. Consequently, the concept of domestication as applied to grass plants may not be directly applicable to trees, and researchers studying tree domestication often face difficulties in defining and identifying clear patterns of domestication.

The four evolutionary forces mentioned above are not always operating with the same proportion in the evolutionary history of species. But gene flow, followed by genetic drift, are likely to have a larger proportion as reported on some studies. However, we also need to consider if the relative influence of these evolutionary forces is equivalent. When examining the evolutionary history of cultivated

Myristica species through the use of neutral genetic markers, two primary evolutionary forces come into play: gene flow and drift. Migration, whether facilitated by natural means or human intervention, can introduce new genetic material into populations, leading to a potential diversification. On the other hand, drift, particularly during founding events or events of population size reduction, can result in random changes in allele frequencies due to sampling effects. By considering these two forces in the study of cultivated *Myristica* species, a deeper understanding of their evolutionary dynamics can be gained.

1.5.2 Evaluating the influence of evolutionary processes at an intraspecific level

If we want to understand how tree populations respond to changing environmental conditions (e.g., climate change) and human activities (e.g., human cultivation), their genetic variation is a key factor that we need to assess (Alberto et al., 2013; Caetano-Andrade et al., 2020; Milesi et al., 2023; Rajora & Zinck, 2021). By analyzing the genetic variation within and among populations, we can infer the processes that have shaped the genetic makeup of cultivated species over time. Since their first introduction (S. Wright, 1949), F -statistics (or fixation indices) have been widely used to quantify the extent and distribution of genetic variation within and among populations. The different F -statistics look at different levels of population structure. F_{IT} is the inbreeding coefficient of individuals (I) relative to the total (T); F_{IS} is the inbreeding coefficient of individuals (I) relative to the subpopulation (S) and F_{ST} is the effect of subpopulations (S) compared to the total population (T).

Inbreeding coefficient (F_{IS})

The inbreeding coefficient (F_{IS}) compares expected and observed heterozygosity (H_O and H_E) within a population and allows to make hypotheses on the evolutionary processes affecting this population (Nei & Roychoudhury, 1974; S. Wright, 1949). Several studies pointed out that heterozygosity levels in tree population are highly correlated with their climatic fluctuation, their distribution range (Péros et al., 2021; Vanden Abeele et al., 2021) and their phenotypical plasticity (Burkardt et al., 2021). For example, if a tree population experiencing selective logging, genetic drift will have a strong impact on allele frequencies and, overtime, we will expect $H_O < H_E$, i.e., a deficit in heterozygotes highlighted by a positive F_{IS} (also see Kitamura, 2011; Lefèvre et al., 2004; Waqar et al., 2021).

In addition to the factors mentioned earlier, it is important to consider the potential effects of crossing between inbred individuals on the observed heterozygosity within a population. Inbreeding can occur through selfing or when species have limited dispersal abilities, leading to a local genetic structure and increased cross between individuals. Such cross-breeding between inbred individuals can further influence the inbreeding coefficient and impact the genetic diversity of a population (Slate et al., 2004). It is worth noting that a high F_{IS} may not solely be attributed to inbreeding. Other factors, such as the presence of null alleles, can also contribute to the observed patterns (De Meeûs, 2018). Distinguishing between these two effects can be challenging, as both inbreeding and the presence of null alleles can lead to an increase in homozygosity. Further investigations, including the use of molecular markers and statistical methods, are

necessary to disentangle the relative contributions of inbreeding and null alleles to the observed F_{IS} values. But in general, the molecular markers will be affected in a same way in the population experienced inbreeding.

Moreover, stable regions are likely to exhibit long-term population persistence, facilitating the accumulation of genetic diversity and promoting high levels of gene flow across the region (Fjeldså & Lovett, 1997; Leffler et al., 2012). Therefore, in such a population, we will expect $H_O > H_E$ and a low level of inbreeding, highlighted by a negative F_{IS} .

Fixation index (F_{ST})

Another complementary way to assess genetic variation is using the fixation index, F_{ST} . Indeed, the repartition of genetic variation within and between populations gives clues on intensity of gene flow between populations. Thus, species with limited gene flow, will exhibit significant genetic differentiation among populations, corresponding to a high F_{ST} value (close to 1). Factors such as the composition and configuration of the landscape (e.g. a geographical barrier such as a river, the sea in the case of archipelago, a mountain, etc.) and the absence (or limited travelling capacities) of pollinator or seed dispersal agents can prevent genetic exchange among populations which leads to genetic differentiation. The degree of genetic differentiation and the effective population size are key factors influencing the strength of genetic drift. Stronger effects are observed in highly differentiated populations and a gradual reduction of genetic diversity over time is observed in populations with small effective sizes (Knowles & Richards, 2005; Toczydlowski & Waller, 2019).

At the opposite, if we consider several patches of forest with persistent connectivity between them, we will observe low levels of genetic differentiation among patches, highlighted by low F_{ST} values (close to 0). In such a species, patterns of genetic isolation by distance can be observed. Low genetic differentiation thus indicates the absence of geographical barriers to gene flow and/or efficient gene flow (Bors et al., 2019; Branch et al., 2017).

It is important to note that F -statistics (both F_{IS} and F_{ST}) have some limitations, such as sensitivity to sample size, assumptions of certain population genetic models (like Hardy-Weinberg equilibrium), and potential confounding effects of demographic processes (Meirmans & Hedrick, 2011).

1.5.3 Retracing the demographic history of populations

Demographic processes collectively determine the changes in population size, structure, and composition, which are essential for understanding the patterns and mechanisms of population dynamics. The population size is expected to undergo evolutionary changes over time, including decline, expansion and stability. These changes are influenced by a variety of factors, such as climate change and human activities (Nieto-Blázquez et al., 2022; Ochoa-Zavala et al., 2019; Wahlsteen et al., 2023). Therefore, the characterization of population size evolution through time can be used to understand the effects of past events, such as glaciation and climate change events, but also of recent/historical ecological processes and anthropogenic activities on species evolutionary history (Aoki et al., 2019; Guerrina et al., 2022; Lu et al., 2020). The study of the demographic history of a species is based on the analysis of different summary statistics, like allele frequency-based

statistics (e.g., mean or variance of allele frequencies), genetic diversity-based statistics (e.g., H_0), and population differentiation statistics (F -statistics) (Beaumont et al. 2002).

Colonization events are good examples to explore how summary statistics allow to reconstruct a demographic history. When tree species colonize new habitats, the demographic history of the population is shaped by factors such as founder effects, genetic drift and limited gene flow. In this case, colonizing individuals might present a lower heterozygosity and potentially different allele frequencies compared to source populations.

Fady & Bozzano (2021) present an insightful argument highlighting the limitations of solely relying on effective population size (N_e) to understand the genetic dynamics of tree populations during colonization events. They emphasize that a reduction in N_e does not necessarily indicate a detrimental loss of genetic diversity. To comprehensively grasp the complexities of tree population dynamics, it is crucial to consider additional factors beyond N_e alone. Factors such as genetic drift, gene flow, and selection pressure play significant roles in shaping the genetic diversity and structure of tree populations, even if N_e appears to be stable or recovering (Goncalves et al., 2019; Hematzadeh et al., 2023; Navascués et al., 2006; Rico et al., 2021; Yoichi et al., 2021; S. Zhu et al., 2020). Determining the timing of population size changes remains a challenging endeavor despite its feasibility.

Time calibration of demographic analysis is complex due to several factors. They include mutation rate of genetic markers, generation time (in the context of the species' reproductive lifespan), and selection of analytical approaches (e.g. model specification, genealogy reconstruction; see Beaumont et al., 2002;

Drummond, 2005). In many cases, very little information is available on the mutation rate and generation time. However, using widely known mutation rate and generation time might provide useful insights, but need to be treated carefully (Catalan et al., 2022; Kamm et al., 2020).

The estimation of historical population size changes can be inferred through observed and simulated summary statistics through approximate Bayesian computation (ABC) framework. Posterior probability distributions are estimated from the collection of parameter values used in the selected simulations (see Beaumont, 2010). Then we can project this population dynamics using the Skyline plot (Pybus et al., 2000; Figure 14). This graphical representation consists of several periods of constant size, with instantaneous changes of sizes between consecutive periods. Hence, the posterior probability regarding the number of periods can be employed to distinguish between stable and changing demography. This is achieved by estimating the Bayes factor (Jeffreys, 1998), comparing the hypothesis of a single period (constant population size) to multiple demographic periods (variable population size).

In some special situations, where a small founding (e.g. island) population splits off from the main population, the estimated bottleneck time (t) of the derived population can be used as a proxy for its divergence time (see also Dussert et al., 2015; Foote et al., 2016; Tamaki et al., 2016). At a temporal scale, it is widely acknowledged that biological diversification in tropical forests has been predominantly influenced by climatic fluctuations during the Late Pleistocene (Haffer, 1969). However, the accuracy of this reconstruction will depend on an adequate statistical method for demographic inferences.

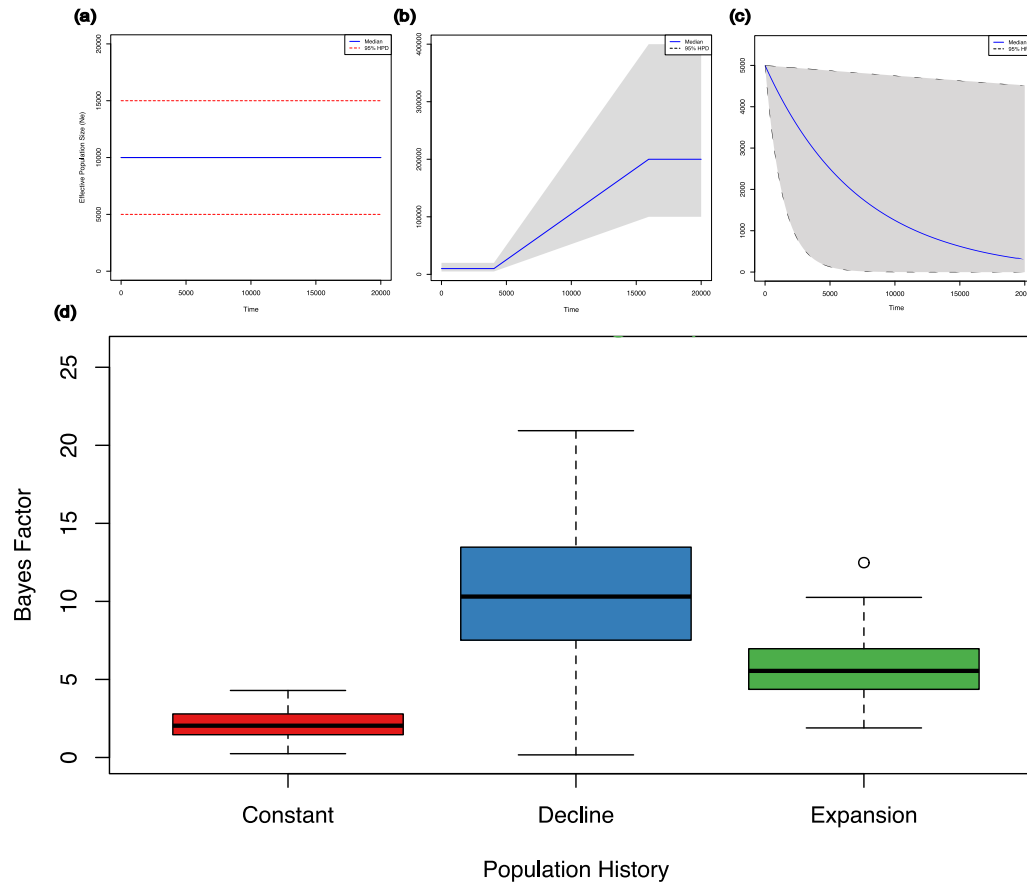


Figure 14. Skyline plots illustrating the (a) constant size, (b) population decline, and (c) population expansion. Median value is in blue color and 95% of highest posterior density (HPD) interval in grey of the posterior probability distribution. Distribution of Bayes factor of each population history illustrated with boxplot (d). We use Jeffrey’s scale (1998) to easily interpret the Bayes factor values. We can draw the conclusion that in “constant” there is no evidence of population size change, in “decline” there is very strong evidence of population size change, and in “expansion” there is strong evidence of population size change (refer to Jeffreys, 1998, for complete interpretation). The value described in this plot are hypothetical.

2. Contextual problem, thesis objectives and outlines

Both environmental and anthropogenic factors have most probably contributed to the evolutionary history of *Myristica* species. The species originate from oceanic islands from the Malay Archipelago. The evolutionary history of the local flora in relation to the evolution of climate, tectonic modifications and human management practices remains poorly documented. Human management practices on Indonesian nutmegs have contributed to their evolutionary history (e.g. cutting

down of *M. fragrans* trees by the VOC to maintain a monopoly on the resource). Additionally, founder events might have occurred due to the propagation of Indonesian nutmegs outside of their native distribution. However, the impact of environmental and anthropogenic factors on the distribution of genetic diversity in Indonesian nutmegs has not been explored to date. This lack of knowledge is of major concern, as genetic diversity is crucial for adaptation and resilience to changing environments. Moreover, understanding the population structure and demographic history of Indonesian nutmegs can provide valuable insights into their evolutionary history and inform conservation strategies. Therefore, there is an urgent need to develop a comprehensive study that investigate the genetic structure and demographic history of Indonesian nutmegs, taking into account the potential impact of environmental and anthropogenic factors. Such a study can contribute to the development of more efficient and sustainable cultivation practices, as well as to the conservation and management of this important crop.

In this thesis dissertation, I aimed to unravel the evolutionary history of Indonesian nutmegs, focusing on the two species *Myristica fragrans* and *Myristica argentea*. I divided this thesis into five different chapters presented in an article format. The layout and reference format of these chapters were edited from the original manuscripts to provide a typographic consistency within the manuscript.

Chapter 1 describes newly developed microsatellite markers from *M. fragrans*. We provide the initial information on their characteristics as well as their transferability to other sister species. As a starting point of my thesis, these genetic markers act as a tool to acquire information on the genetic structure and

demographic history of the focal species. This paper was published in *Molecular Biology Reports* (2020).

Chapter 2 presents the reconstruction of the whole chloroplast (cp) DNA genome for *M. argentea* and *M. fatua*. From *de novo* assembly to annotation, we describe the cpDNA genome for the first time in these two species. The importance of these genomic resources provide a valuable reference for studying phylogeographic distribution of the focal species. This manuscript is under minor revision in *Mitochondrial DNA part B*.

Chapter 3 discusses the origin of *M. fragrans* in the Moluccas. This chapter focuses on the population genetic and on the biogeographical history of the nutmeg in the Moluccas. More specifically, we were testing whether the species originated from Banda islands only or from the whole Moluccas. This manuscript has been submitted to *Molecular Ecology*.

Chapter 4 presents the worldwide history of introduction in *M. fragrans*. It is well documented that the species is now widespread in the tropical areas of the world, and we here describe the genetic diversity distribution of the species worldwide.

Chapter 5 describes the spatial distribution of the genetic diversity of the narrowly distributed and micro-endemic *M. argentea*. We tested the influence of different factors (historical barriers to gene flow, recent management practices) on the genetic structure and historical population dynamics of the species.

Finally, the results of these chapters are summarized, combined, and discussed in a *General Discussion* section.

References

- Abourashed, E. A., & El-Alfy, A. T. (2016). Chemical diversity and pharmacological significance of the secondary metabolites of nutmeg (*Myristica fragrans* Houtt.). *Phytochemistry Reviews*, *15*(6), 1035–1056. <https://doi.org/10.1007/s11101-016-9469-x>
- Aitken, S. N., & Whitlock, M. C. (2013). Assisted gene flow to facilitate local adaptation to climate change. *Annual Review of Ecology, Evolution, and Systematics*, *44*(1), 367–388. <https://doi.org/10.1146/annurev-ecolsys-110512-135747>
- Alberto, F. J., Aitken, S. N., Alía, R., González-Martínez, S. C., Hänninen, H., Kremer, A., Lefèvre, F., Lenormand, T., Yeaman, S., Whetten, R., & Savolainen, O. (2013). Potential for evolutionary responses to climate change – evidence from tree populations. *Global Change Biology*, *19*(6), 1645–1661. <https://doi.org/10.1111/gcb.12181>
- Alsos, I. G., Ehrich, D., Eidesen, P. B., Solstad, H., Westergaard, K. B., Schönswetter, P., Tribsch, A., Birkeland, S., Elven, R., & Brochmann, C. (2015). Long-distance plant dispersal to North Atlantic islands: colonization routes and founder effect. *AoB PLANTS*, *7*. <https://doi.org/10.1093/aobpla/plv036>
- Andaya, L. (1991). Local Trade Networks in Maluku in the 16th, 17th and 18th Centuries. *Cakalele*, *2*(2), 71–96.
- Andersson, L., & Purugganan, M. (2022). Molecular genetic variation of animals and plants under domestication. *Proceedings of the National Academy of Sciences*, *119*(30). <https://doi.org/10.1073/pnas.2122150119>
- Anjarwalla, P., Ofori, D., Owino, A., Matuku, D., Adika, W., Njogu, K., & Kehlenbeck, K. (2017). Testing different grafting methods for vegetative propagation of baobab (*Adansonia digitata* L.) in Kenya to assist its domestication and promote cultivation. *Forests, Trees and Livelihoods*, *26*(2), 85–95. <https://doi.org/10.1080/14728028.2016.1202791>
- Aoki, K., Tamaki, I., Nakao, K., Ueno, S., Kamijo, T., Setoguchi, H., Murakami, N., Kato, M., & Tsumura, Y. (2019). Approximate Bayesian computation analysis of EST-associated microsatellites indicates that the broadleaved evergreen tree *Castanopsis sieboldii* survived the Last Glacial Maximum in multiple refugia in Japan. *Heredity*, *122*(3), 326–340. <https://doi.org/10.1038/s41437-018-0123-9>
- Assis, A. P. A., Galetti, M., Maia, K. P., & Guimarães, P. R. (2022). Reduced evolutionary potential of a frugivorous bird species in fragmented forests. *Frontiers in Ecology and Evolution*, *10*. <https://doi.org/10.3389/fevo.2022.804138>

- Atkins, H. J., Bramley, G. L. C., Johnson, M. A., Kartonegoro, A., Nishii, K., Kokubugata, G., Moeller, M., & Hughes, M. (2020). A molecular phylogeny of Southeast Asian *Cyrtandra* (Gesneriaceae) supports an emerging paradigm for Malesian plant biogeography. *Frontiers of Biogeography*, 12(1). <https://doi.org/10.21425/F5FBG44184>
- Austerlitz, F., Mariette, S., Machon, N., Gouyon, P.-H., & Godelle, B. (2000). Effects of Colonization Processes on Genetic Diversity: Differences Between Annual Plants and Tree Species. *Genetics*, 154(3), 1309–1321. <https://doi.org/10.1093/genetics/154.3.1309>
- Bagnoli, F., Vendramin, G. G., Buonamici, A., Doulis, A. G., Gonzalez-Martinez, S. C., La Porta, N., Magri, D., Raddi, P., Sebastiani, F., & Fineschi, S. (2009). Is *Cupressus sempervirens* native in Italy? An answer from genetic and palaeobotanical data. *Molecular Ecology*, 18(10), 2276–2286. <https://doi.org/10.1111/j.1365-294X.2009.04182.x>
- Baldwin, L., Tosti, N., Ricciolini, C., Belaj, A., Arcioni, S., Pannelli, G., Germana, M. A., Mulas, M., & Porceddu, A. (2006). Genetic Structure of Wild and Cultivated Olives in the Central Mediterranean Basin. *Annals of Botany*, 98(5), 935–942. <https://doi.org/10.1093/aob/mcl178>
- Baldwin, S. L., Fitzgerald, P. G., & Webb, L. E. (2012). Tectonics of the New Guinea Region. *Annual Review of Earth and Planetary Sciences*, 40(1), 495–520. <https://doi.org/10.1146/annurev-earth-040809-152540>
- Barjiyah, U. (2011). Orang Buangan (Bannelingen) pada perkebunan pala di Kepulauan Banda. *Paramita: Historical Studies Journal*, 20(1). <https://doi.org/10.15294/paramita.v20i1.1054>
- 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](https://doi.org/https://doi.org/10.1016/0031-0182(93)90157-E)
- Barret, S. C. H. (1996). The reproductive biology and genetics of island plants. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 351(1341), 725–733. <https://doi.org/10.1098/rstb.1996.0067>
- BBKSDA, P. B. (2022). *Globally recognized nutmeg from the east (Myristica argentea Warb). in Indonesian: Rempah dari timur yang telah mendunia: Pala Papua (Myristica argentea Warb).* Accessed in 20 May 2023 in <https://Bbksda-Papuabarat.Com/Rempah-Dari-Timur-Indonesia-Yang-Telah-Mendunia-Pala-Papua-Myristica-Argentea-Warb/>.

- Beaumont, M. A. (2010). Approximate Bayesian Computation in Evolution and Ecology. *Annual Review of Ecology, Evolution, and Systematics*, 41(1), 379–406. <https://doi.org/10.1146/annurev-ecolsys-102209-144621>
- 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>
- Bemmels, J. B., Haddrath, O., Colbourne, R. M., Robertson, H. A., & Weir, J. T. (2022). Legacy of supervolcanic eruptions on population genetic structure of brown kiwi. *Current Biology*, 32(15), 3389–3397.e8. <https://doi.org/10.1016/j.cub.2022.05.064>
- Ben-Avraham, Z. (1973). *Structural framework of the Sunda Shelf and vicinity*. Massachusetts Institute of Technology and Woods Hole Oceanographic Institution. <https://doi.org/10.1575/1912/1265>
- Bentley, R., & 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. 1). J. & A. Churchill. <https://www.biodiversitylibrary.org/item/78161>
- Bezemer, N., Hopper, S. D., Krauss, S. L., Phillips, R. D., & Roberts, D. G. (2019). Primary pollinator exclusion has divergent consequences for pollen dispersal and mating in different populations of a bird-pollinated tree. *Molecular Ecology*, 28(22), 4883–4898. <https://doi.org/10.1111/mec.15264>
- Bors, E. K., Herrera, S., Morris, J. A., & Shank, T. M. (2019). Population genomics of rapidly invading lionfish in the Caribbean reveals signals of range expansion in the absence of spatial population structure. *Ecology and Evolution*, 9(6), 3306–3320. <https://doi.org/10.1002/ece3.4952>
- Branch, C. L., Jahner, J. P., Kozlovsky, D. Y., Parchman, T. L., & Pravosudov, V. V. (2017). Absence of population structure across elevational gradients despite large phenotypic variation in mountain chickadees (*Poecile gambeli*). *Royal Society Open Science*, 4(3), 170057. <https://doi.org/10.1098/rsos.170057>
- Brixius, D. (2015). A Pepper acquiring Nutmeg: Pierre Poivre, The French Spice Quest and the Role of Mediators in Southeast Asia, 1740s to 1770s. *Journal of the Western Society for French History*, 43(1).
- Bromham, L., Hua, X., Lanfear, R., & Cowman, P. F. (2015). Exploring the Relationships between Mutation Rates, Life History, Genome Size, Environment, and Species Richness in Flowering Plants. *The American Naturalist*, 185(4), 507–524. <https://doi.org/10.1086/680052>

- Brummit, R. K. (2001). *World Geographic Scheme for Recording Plant Distributions* (R. K. Brummit, Ed.; Edition 2, Vol. 2). Hunt Institute for Botanical Documentation, Carnegie Mellon University (Pittsburgh). <http://rs.tdwg.org/wgsrpd/doc/data/>
- Brundu, G., & Richardson, D. M. (2016). Planted forests and invasive alien trees in Europe: A Code for managing existing and future plantings to mitigate the risk of negative impacts from invasions. *NeoBiota*, 30, 5–47. <https://doi.org/10.3897/neobiota.30.7015>
- Burkardt, K., Pettenkofer, T., Ammer, C., Gailing, O., Leinemann, L., Seidel, D., & Vor, T. (2021). Influence of heterozygosity and competition on morphological tree characteristics of *Quercus rubra* L.: a new single-tree based approach. *New Forests*, 52(4), 679–695. <https://doi.org/10.1007/s11056-020-09814-1>
- Caetano-Andrade, V. L., Clement, C. R., Weigel, D., Trumbore, S., Boivin, N., Schöngart, J., & Roberts, P. (2020). Tropical Trees as Time Capsules of Anthropogenic Activity. *Trends in Plant Science*, 25(4), 369–380. <https://doi.org/10.1016/j.tplants.2019.12.010>
- Cámara-Leret, R., Frodin, D. G., Adema, F., Anderson, C., Appelhans, M. S., Argent, G., Arias Guerrero, S., Ashton, P., Baker, W. J., Barfod, A. S., Barrington, D., Borosova, R., Bramley, G. L. C., Briggs, M., Buerki, S., Cahen, D., Callmander, M. W., Cheek, M., Chen, C.-W., ... van Welzen, P. C. (2020). New Guinea has the world's richest island flora. *Nature*, 584(7822), 579–583. <https://doi.org/10.1038/s41586-020-2549-5>
- Catalan, A., Höhna, S., Lower, S. E., & Duchon, P. (2022). Inferring the demographic history of the North American firefly *Photinus pyralis*. *Journal of Evolutionary Biology*, 35(11), 1488–1499. <https://doi.org/10.1111/jeb.14094>
- Charvet, G. (2011). Wheat domestication: Lessons for the future. *Comptes Rendus Biologies*, 334(3), 212–220. <https://doi.org/10.1016/j.crvbi.2010.12.013>
- Choi, J. Y., Platts, A. E., Fuller, D. Q., Hsing, Y.-I., Wing, R. A., & Purugganan, M. D. (2017). The rice paradox: Multiple origins but single domestication in Asian rice. *Molecular Biology and Evolution*, msx049. <https://doi.org/10.1093/molbev/msx049>
- Cordeiro, E. M. G., Macrini, C. M., Sujii, P. S., Schwarcz, K. D., Pinheiro, J. B., Rodrigues, R. R., Brancalion, P. H. S., & Zucchi, M. I. (2019). Diversity, genetic structure, and population genomics of the tropical tree *Centropogon tomentosum* in remnant and restored Atlantic forests. *Conservation Genetics*, 20(5), 1073–1085. <https://doi.org/10.1007/s10592-019-01195-z>

- Cornille, A., Antolín, F., Garcia, E., Vernesi, C., Fietta, A., Brinkkemper, O., Kirleis, W., Schlumbaum, A., & Roldán-Ruiz, I. (2019). A Multifaceted Overview of Apple Tree Domestication. *Trends in Plant Science*, 24(8), 770–782. <https://doi.org/10.1016/j.tplants.2019.05.007>
- Cornille, A., Feurtey, A., Gélin, U., Ropars, J., Misvanderbrugge, K., Gladieux, P., & Giraud, T. (2015). Anthropogenic and natural drivers of gene flow in a temperate wild fruit tree: a basis for conservation and breeding programs in apples. *Evolutionary Applications*, 8(4), 373–384. <https://doi.org/10.1111/eva.12250>
- Cueva-Agila, A., Vélez-Mora, D., Arias, D., Curto, M., Meimberg, H., & Brinegar, C. (2019). Genetic characterization of fragmented populations of *Cinchona officinalis* L. (Rubiaceae), a threatened tree of the northern Andean cloud forests. *Tree Genetics & Genomes*, 15(6), 81. <https://doi.org/10.1007/s11295-019-1393-y>
- Darwin, C. (1859). *On the origin of species by means of natural selection, or preservation of favoured races in the struggle for life*. London : John Murray, 1859. <https://search.library.wisc.edu/catalog/9934839413602122>
- Dasmann, R. (1968). A Different Kind of Country. In *Oryx* (2009/04/24, Vol. 9, Issue 6). Cambridge University Press. <https://doi.org/DOI:10.1017/S0030605300007389>
- De La Torre, A. R., Li, Z., Van de Peer, Y., & Ingvarsson, P. K. (2017). Contrasting Rates of Molecular Evolution and Patterns of Selection among Gymnosperms and Flowering Plants. *Molecular Biology and Evolution*, 34(6), 1363–1377. <https://doi.org/10.1093/molbev/msx069>
- De Meeûs, T. (2018). Revisiting FIS, FST, Wahlund Effects, and Null Alleles. *Journal of Heredity*, 109(4), 446–456. <https://doi.org/10.1093/jhered/esx106>
- de Wilde, W. J. J. O. (1995). Census of *Myristica* (Myristicaceae) in New Guinea anno 1994. *Blumea: Biodiversity, Evolution and Biogeography of Plants*, 40(2), 237–344.
- de Wilde, W. J. J. O. (2014). *Flora Malesiana*. Pensoft Publishers. <https://doi.org/10.3897/ab.e1141>
- Dickie, I. A., Bennett, B. M., Burrows, L. E., Nuñez, M. A., Peltzer, D. A., Porté, A., Richardson, D. M., Rejmánek, M., Rundel, P. W., & van Wilgen, B. W. (2014). Conflicting values: ecosystem services and invasive tree management. *Biological Invasions*, 16(3), 705–719. <https://doi.org/10.1007/s10530-013-0609-6>

- Drummond, A. J. (2005). Bayesian Coalescent Inference of Past Population Dynamics from Molecular Sequences. *Molecular Biology and Evolution*, 22(5), 1185–1192. <https://doi.org/10.1093/molbev/msi103>
- Dussert, Y., Snirc, A., & Robert, T. (2015). Inference of domestication history and differentiation between early- and late-flowering varieties in pearl millet. *Molecular Ecology*, 24(7), 1387–1402. <https://doi.org/10.1111/mec.13119>
- Einfeldt, A. L., Jesson, L. K., & Addison, J. A. (2020). Historical human activities reshape evolutionary trajectories across both native and introduced ranges. *Ecology and Evolution*, 10(13), 6579–6592. <https://doi.org/10.1002/ece3.6391>
- Elleouet, J. S., & Aitken, S. N. (2019). Long-distance pollen dispersal during recent colonization favors a rapid but partial recovery of genetic diversity in *Picea sitchensis*. *New Phytologist*, 222(2), 1088–1100. <https://doi.org/10.1111/nph.15615>
- Estoque, R. C., Ooba, M., Avitabile, V., Hijioka, Y., DasGupta, R., Togawa, T., & Murayama, Y. (2019). The future of Southeast Asia's forests. *Nature Communications*, 10(1), 1829. <https://doi.org/10.1038/s41467-019-09646-4>
- Fady, B., & Bozzano, M. (2021). Effective population size does not make a practical indicator of genetic diversity in forest trees. *Biological Conservation*, 253, 108904. <https://doi.org/10.1016/j.biocon.2020.108904>
- Fairbairn, A. S., Hope, G. S., & Summerhayes, G. R. (2006). Pleistocene Occupation of New Guinea's Highland and Subalpine Environments. *World Archaeology*, 38(3), 371–386. <http://www.jstor.org/stable/40026639>
- 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>
- Ferrer, M. M., Montañez-Escalante, P. I., Ruenes-Morales, M. del R., Estrada-Medina, H., & Jiménez-Osornio, J. (2020). Growing out of the tropical forests: domestication syndrome of native Mesoamerican trees in Mayan homegardens. *Genetic Resources and Crop Evolution*, 67(3), 587–604. <https://doi.org/10.1007/s10722-019-00833-2>
- Fjeldså, J., & Lovett, J. (1997). Biodiversity and environmental stability. *Biodiversity and Conservation*, 6(3), 315–323. <https://doi.org/10.1023/A:1018304522320>

- Flach, M. (1966). *Nutmeg cultivation and its sex-problem : an agronomical and cytogenetical study of the dioecy in Myristica fragrans Houtt. and Myristica argentea Warb.* <https://edepot.wur.nl/187949>
- Foote, A. D., Vijay, N., Ávila-Arcos, M. C., Baird, R. W., Durban, J. W., Fumagalli, M., Gibbs, R. A., Hanson, M. B., Korneliussen, T. S., Martin, M. D., Robertson, K. M., Sousa, V. C., Vieira, F. G., Vinař, T., Wade, P., Worley, K. C., Excoffier, L., Morin, P. A., Gilbert, M. T. P., & Wolf, J. B. W. (2016). Genome-culture coevolution promotes rapid divergence of killer whale ecotypes. *Nature Communications*, 7(1), 11693. <https://doi.org/10.1038/ncomms11693>
- Fountain, T., Nieminen, M., Sirén, J., Wong, S. C., Lehtonen, R., & Hanski, I. (2016). Predictable allele frequency changes due to habitat fragmentation in the Glanville fritillary butterfly. *Proceedings of the National Academy of Sciences*, 113(10), 2678–2683. <https://doi.org/10.1073/pnas.1600951113>
- Frankham, R. (1996). Relationship of Genetic Variation to Population Size in Wildlife. *Conservation Biology*, 10(6), 1500–1508. <https://doi.org/10.1046/j.1523-1739.1996.10061500.x>
- Frankham, R. (1997). Do island populations have less genetic variation than mainland populations? *Heredity*, 78(3), 311–327. <https://doi.org/10.1038/hdy.1997.46>
- Gallardo, M. H., Köhler, N., & Araneda, C. (1995). Bottleneck effects in local populations of fossorial *Ctenomys* (Rodentia, Ctenomyidae) affected by vulcanism. *Heredity*, 74(6), 638–646. <https://doi.org/10.1038/hdy.1995.87>
- Gaut, B. S., Muse, S. V., Clark, W. D., & Clegg, M. T. (1992). Relative rates of nucleotide substitution at the *rbcL* locus of monocotyledonous plants. *Journal of Molecular Evolution*, 35(4), 292–303. <https://doi.org/10.1007/BF00161167>
- GBIF.org. (2023a, May 15). *Myristica argentea* occurrence. GBIF Occurrence Download <https://doi.org/10.15468/DI.Uue8w9>.
- GBIF.org. (2023b, May 15). *Myristica fragrans* occurrence. GBIF Occurrence Download <https://doi.org/10.15468/DI.Uue8w9>.
- Gold, D. P., White, L. T., Gunawan, I., & BouDagher-Fadel, M. K. (2017). Relative sea-level change in western New Guinea recorded by regional biostratigraphic data. *Marine and Petroleum Geology*, 86, 1133–1158. <https://doi.org/10.1016/j.marpetgeo.2017.07.016>
- Goncalves, A. L., García, M. V., Heuertz, M., & González-Martínez, S. C. (2019). Demographic history and spatial genetic structure in a remnant population of the subtropical tree *Anadenanthera colubrina* var. *cebil* (Griseb.) Altschul

- (Fabaceae). *Annals of Forest Science*, 76(1), 18.
<https://doi.org/10.1007/s13595-019-0797-z>
- Grant, E. L., Conroy, G. C., Lamont, R. W., Reddell, P. W., Wallace, H. M., & Ogbourne, S. M. (2019). Short distance pollen dispersal and low genetic diversity in a subcanopy tropical rainforest tree, *Fontainea picrosperma* (Euphorbiaceae). *Heredity*, 123(4), 503–516. <https://doi.org/10.1038/s41437-019-0231-1>
- Gros-Balthazard, M., Galimberti, M., Kousathanas, A., Newton, C., Ivorra, S., Paradis, L., Vigouroux, Y., Carter, R., Tengberg, M., Battesti, V., Santoni, S., Falquet, L., Pintaud, J.-C., Terral, J.-F., & Wegmann, D. (2017). The Discovery of Wild Date Palms in Oman Reveals a Complex Domestication History Involving Centers in the Middle East and Africa. *Current Biology*, 27(14), 2211–2218.e8. <https://doi.org/https://doi.org/10.1016/j.cub.2017.06.045>
- Guerrina, M., Theodoridis, S., Minuto, L., Conti, E., & Casazza, G. (2022). First evidence of post-glacial contraction of Alpine endemics: Insights from *Berardia subacaulis* in the European Alps. *Journal of Biogeography*, 49(1), 79–93. <https://doi.org/10.1111/jbi.14282>
- Haddon, A. C., Sollas, W. J., & Cole, G. A. J. (1892). On the Geology of Torres Straits. *The Transactions of the Royal Irish Academy*, 30, 419–476. <http://www.jstor.org/stable/30079786>
- Haffer, J. (1969). Speciation in Amazonian Forest Birds. *Science*, 165(3889), 131–137. <https://doi.org/10.1126/science.165.3889.131>
- Hall, R. (2009a). *Indonesia, Geology* (R. Gillespie & D. Clague, Eds.; pp. 454–460). University of California Press. <https://doi.org/doi:10.1525/9780520943728-104>
- Hall, R. (2009b). Southeast Asia’s changing palaeogeography. *Blumea-Biodiversity, Evolution and Biogeography of Plants*, 54(1–2), 148–161.
- Hall, R. (2013). The palaeogeography of Sundaland and Wallacea since the Late Jurassic. *Journal of Limnology*, 72, e1.
- Hampe, A. (2004). Extensive hydrochory uncouples spatiotemporal patterns of seedfall and seedling recruitment in a “bird-dispersed” riparian tree. *Journal of Ecology*, 92(5), 797–807. <https://doi.org/10.1111/j.0022-0477.2004.00918.x>
- Hanebuth, T., Statterger, K., & Grootes, P. M. (2000). Rapid Flooding of the Sunda Shelf: A Late-Glacial Sea-Level Record. *Science*, 288(5468), 1033–1035. <https://doi.org/10.1126/science.288.5468.1033>

- Harris, J. A. (1916). The Variable Desert. *The Scientific Monthly*, 3(1), 41–50. <http://www.jstor.org/stable/6182>
- Hartemink, A. E. (2010). The invasive shrub *Piper aduncum* in Papua New Guinea: a review. *Journal of Tropical Forest Science*, 22(2), 202–213. <http://www.jstor.org/stable/23616728>
- Heads, M. (2001). Regional patterns of biodiversity in New Guinea plants. *Botanical Journal of the Linnean Society*, 136(1), 67–73. <https://doi.org/10.1111/j.1095-8339.2001.tb00556.x>
- Hematzadeh, A., Esmailzadeh, O., Jalali, S. G., Mirjalili, M. H., Walas, Ł., & Yousefzadeh, H. (2023). Genetic diversity and structure of English yew (*Taxus baccata* L.) as a tertiary relict and endangered tree in the Hyrcanian forests. *Biodiversity and Conservation*, 32(5), 1733–1753. <https://doi.org/10.1007/s10531-023-02573-3>
- Hill, K. C., & Gleadow, A. J. W. (1989). Uplift and thermal history of the Papuan Fold Belt, Papua New Guinea: Apatite fission track analysis. *Australian Journal of Earth Sciences*, 36(4), 515–539. <https://doi.org/10.1080/08120098908729507>
- Hill, K. C., & Hall, R. (2003). Mesozoic-Cenozoic evolution of Australia's New Guinea margin in a west Pacific context. In R. Hillis & R. Müller (Eds.), *Evolution and Dynamics of the Australian Plate* (1st ed., Vol. 372). Geological Society of America. <https://doi.org/10.1130/0-8137-2372-8.265>
- Holm, R. J., Rosenbaum, G., & Richards, S. W. (2016). Post 8 Ma reconstruction of Papua New Guinea and Solomon Islands: Microplate tectonics in a convergent plate boundary setting. *Earth-Science Reviews*, 156, 66–81. <https://doi.org/10.1016/j.earscirev.2016.03.005>
- Holm, R. J., Spandler, C., & Richards, S. W. (2015). Continental collision, orogenesis and arc magmatism of the Miocene Maramuni arc, Papua New Guinea. *Gondwana Research*, 28(3), 1117–1136. <https://doi.org/10.1016/j.gr.2014.09.011>
- Houttuyn, M. (Martinus), Linné, C. von, & Philips, J. C. (1774). *Natuurlyke historie : of, Uitvoerige beschryving der dieren, planten, en mineraalen: Vol. d.2:stuk 3 (1774)*. By F. Houttuyn [and others]. <https://www.biodiversitylibrary.org/item/135748>
- Hu, D., Xu, Y., Chai, Y., Tian, T., Wang, K., Liu, P., Wang, M., Zhu, J., Hou, D., & Yue, M. (2022). Spatial Distribution Pattern and Genetic Diversity of *Quercus wutaishanica* Mayr Population in Loess Plateau of China. *Forests*, 13(9), 1375. <https://doi.org/10.3390/f13091375>

- Hufford, M. B., Xu, X., van Heerwaarden, J., Pyhäjärvi, T., Chia, J.-M., Cartwright, R. A., Elshire, R. J., Glaubitz, J. C., Guill, K. E., Kaeppler, S. M., Lai, J., Morrell, P. L., Shannon, L. M., Song, C., Springer, N. M., Swanson-Wagner, R. A., Tiffin, P., Wang, J., Zhang, G., ... Ross-Ibarra, J. (2012). Comparative population genomics of maize domestication and improvement. *Nature Genetics*, 44(7), 808–811. <https://doi.org/10.1038/ng.2309>
- Irmayanti, L., Ashari, R., Peniwidiyanti, Umanailo, F., Rangkuti, A. B., Fatrawana, A., Nurhikmah, Nurdin, A. S., & Nur, M. (2022). Flora composition and diversity in Mount Sibela Educational Forest, South Halmahera, North Maluku. *IOP Conference Series: Earth and Environmental Science*, 959(1), 012015. <https://doi.org/10.1088/1755-1315/959/1/012015>
- Jeffreys, H. (1998). *The Theory of Probability* (3rd Editio). Oxford University Press.
- Johnson, T., & Barton, N. (2005). Theoretical models of selection and mutation on quantitative traits. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1459), 1411–1425. <https://doi.org/10.1098/rstb.2005.1667>
- Julca, I., Marcet-Houben, M., Cruz, F., Gómez-Garrido, J., Gaut, B. S., Díez, C. M., Gut, I. G., Alioto, T. S., Vargas, P., & Gabaldón, T. (2020). Genomic evidence for recurrent genetic admixture during the domestication of Mediterranean olive trees (*Olea europaea* L.). *BMC Biology*, 18(1), 148. <https://doi.org/10.1186/s12915-020-00881-6>
- Juwita, R., & Tsuchida, S. (2017). Current conditions and profitability of the nutmeg industry in Bogor regency, Indonesia. *Journal of the International Society for Southeast Asian Agricultural Sciences*, 23, 33–44.
- Kameyama, Y., Furumichi, J., Li, J., & Tseng, Y.-H. (2017). Natural genetic differentiation and human-mediated gene flow: the spatiotemporal tendency observed in a long-lived *Cinnamomum camphora* (Lauraceae) tree. *Tree Genetics & Genomes*, 13(2), 38. <https://doi.org/10.1007/s11295-017-1119-y>
- Kamm, J., Terhorst, J., Durbin, R., & Song, Y. S. (2020). Efficiently Inferring the Demographic History of Many Populations With Allele Count Data. *Journal of the American Statistical Association*, 115(531), 1472–1487. <https://doi.org/10.1080/01621459.2019.1635482>
- Kemp, A. C., & Kirwan, G. M. (2020). Blyth's Hornbill (*Rhyticeros plicatus*). In J. del Hoyo, A. Elliott, J. Sargatal, D. Christie, & E. de Juana (Eds.), *Birds of the World*. Cornell Lab of Ornithology. <https://doi.org/10.2173/bow.blyhor1.01>
- 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>

- Knowles, L. L., & Richards, C. L. (2005). Importance of genetic drift during Pleistocene divergence as revealed by analyses of genomic variation. *Molecular Ecology*, *14*(13), 4023–4032. <https://doi.org/10.1111/j.1365-294X.2005.02711.x>
- Kondrashov, F. A., & Kondrashov, A. S. (2010). Measurements of spontaneous rates of mutations in the recent past and the near future. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1544), 1169–1176. <https://doi.org/10.1098/rstb.2009.0286>
- Kremer, A., Ronce, O., Robledo-Arnuncio, J. J., Guillaume, F., Bohrer, G., Nathan, R., Bridle, J. R., Gomulkiewicz, R., Klein, E. K., Ritland, K., Kuperinen, A., Gerber, S., & Schueler, S. (2012). Long-distance gene flow and adaptation of forest trees to rapid climate change. *Ecology Letters*, *15*(4), 378–392. <https://doi.org/10.1111/j.1461-0248.2012.01746.x>
- Krutovsky, K. V., Burczyk, J., Chybicki, I., Finkeldey, R., Pyhäjärvi, T., & Robledo-Arnuncio, J. J. (2012). Gene Flow, Spatial Structure, Local Adaptation, and Assisted Migration in Trees. In *Genomics of Tree Crops* (pp. 71–116). Springer New York. https://doi.org/10.1007/978-1-4614-0920-5_4
- Kumar, S. (2005). Molecular clocks: four decades of evolution. *Nature Reviews Genetics*, *6*(8), 654–662. <https://doi.org/10.1038/nrg1659>
- Laikre, L., Hoban, S., Bruford, M. W., Segelbacher, G., Allendorf, F. W., Gajardo, G., Rodríguez, A. G., Hedrick, P. W., Heuertz, M., Hohenlohe, P. A., Jaffé, R., Johannesson, K., Liggins, L., MacDonald, A. J., Orozco-Wengel, P., Reusch, T. B. H., Rodríguez-Correa, H., Russo, I.-R. M., Ryman, N., & Vernesi, C. (2020). Post-2020 goals overlook genetic diversity. *Science*, *367*(6482), 1083–1085. <https://doi.org/10.1126/science.abb2748>
- Lanfear, R., Ho, S. Y. W., Jonathan Davies, T., Moles, A. T., Aarssen, L., Swenson, N. G., Warman, L., Zanne, A. E., & Allen, A. P. (2013). Taller plants have lower rates of molecular evolution. *Nature Communications*, *4*(1), 1879. <https://doi.org/10.1038/ncomms2836>
- 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>
- Latouche-Hallé, C., Ramboer, A., Bandou, E., Caron, H., & Kremer, A. (2004). Long-distance pollen flow and tolerance to selfing in a neotropical tree species. *Molecular Ecology*, *13*(5), 1055–1064. <https://doi.org/10.1111/j.1365-294X.2004.02127.x>
- Lavabre, J. E., Gilarranz, L. J., Fortuna, M. A., & Bascompte, J. (2016). How does the functional diversity of frugivorous birds shape the spatial pattern of seed dispersal? A case study in a relict plant species. *Philosophical Transactions of*

- the Royal Society B: Biological Sciences*, 371(1694), 20150280. <https://doi.org/10.1098/rstb.2015.0280>
- Lee, H., Popodi, E., Tang, H., & Foster, P. L. (2012). Rate and molecular spectrum of spontaneous mutations in the bacterium *Escherichia coli* as determined by whole-genome sequencing. *Proceedings of the National Academy of Sciences*, 109(41). <https://doi.org/10.1073/pnas.1210309109>
- Lefèvre, F., Fady, B., Fallour-Rubio, D., Ghosn, D., & Bariteau, M. (2004). Impact of founder population, drift and selection on the genetic diversity of a recently translocated tree population. *Heredity*, 93(6), 542–550. <https://doi.org/10.1038/sj.hdy.6800549>
- Leffler, E. M., Bullaughey, K., Matute, D. R., Meyer, W. K., Ségurel, L., Venkat, A., Andolfatto, P., & Przeworski, M. (2012). Revisiting an Old Riddle: What Determines Genetic Diversity Levels within Species? *PLoS Biology*, 10(9), e1001388. <https://doi.org/10.1371/journal.pbio.1001388>
- Leigh, F. J., Wright, T. I. C., Horsnell, R. A., Dyer, S., & Bentley, A. R. (2022). Progenitor species hold untapped diversity for potential climate-responsive traits for use in wheat breeding and crop improvement. *Heredity*, 128(5), 291–303. <https://doi.org/10.1038/s41437-022-00527-z>
- Loewe, M. (1971). Spices and Silk: Aspects of World Trade in the First Seven Centuries of the Christian Era. *Journal of the Royal Asiatic Society of Great Britain and Ireland*, 2, 166–179. <http://www.jstor.org/stable/25203290>
- Londo, J. P., Chiang, Y.-C., Hung, K.-H., Chiang, T.-Y., & Schaal, B. A. (2006). Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. *Proceedings of the National Academy of Sciences*, 103(25), 9578–9583. <https://doi.org/10.1073/pnas.0603152103>
- Lu, R.-S., Chen, Y., Tamaki, I., Sakaguchi, S., Ding, Y.-Q., Takahashi, D., Li, P., Isaji, Y., Chen, J., & Qiu, Y.-X. (2020). Pre-quaternary diversification and glacial demographic expansions of *Cardiocrinum* (Liliaceae) in temperate forest biomes of Sino-Japanese Floristic Region. *Molecular Phylogenetics and Evolution*, 143, 106693. <https://doi.org/10.1016/j.ympev.2019.106693>
- Lydekker, R. (1896). *A geographical history of mammals*. The University press. <https://www.biodiversitylibrary.org/item/75002>
- Lynch, M. (2010). Evolution of the mutation rate. *Trends in Genetics*, 26(8), 345–352. <https://doi.org/10.1016/j.tig.2010.05.003>
- Mañourová, A., Leuner, O., Tchoundjeu, Z., Van Damme, P., Verner, V., Přibyl, O., & Lojka, B. (2019). Medicinal Potential, Utilization and Domestication

- Status of Bitter Kola (*Garcinia kola* Heckel) in West and Central Africa. *Forests*, 10(2), 124. <https://doi.org/10.3390/f10020124>
- Mansyur, S. (2017). Periode kolonial di pesisir Pulau Seram (Maluku): Kontak awal hingga terbentuknya Kota Bula. *Jurnal Walennae*, 15(1), 59. <https://doi.org/10.24832/wln.v15i1.1>
- McAlpine, J. R. (John R. (1983). *Climate of Papua New Guinea* (G. Keig 1945-, R. Falls, & CSIRO, Eds.). Commonwealth Scientific and Industrial Research Organization in association with Australian National University Press.
- Meirmans, P. G., & Hedrick, P. W. (2011). Assessing population structure: F_{st} and related measures. *Molecular Ecology Resources*, 11(1), 5–18. <https://doi.org/10.1111/j.1755-0998.2010.02927.x>
- Metcalf, I. (2001). Faunal and Floral Migration and Evolution in SE Asia-Australasia . In I. Metcalfe, Jeremy. M. B. Smith, M. Morwood, & I. Davidson (Eds.), *Palaeozoic and Mesozoic tectonic evolution and biogeography of SE Asia-Australasia* (1st ed., Vol. 1, pp. 21–23). A. A. Balkema Publishers.
- Meyer, R. S., & Purugganan, M. D. (2013). Evolution of crop species: genetics of domestication and diversification. *Nature Reviews Genetics*, 14(12), 840–852. <https://doi.org/10.1038/nrg3605>
- Middleton, D. J., Armstrong, K., Baba, Y., Balslev, H., Chayamarit, K., Chung, R. C. K., Conn, B. J., Fernando, E. S., Fujikawa, K., Kiew, R., Luu, H. T., Aung, M. M., Newman, M. F., Tagane, S., Tanaka, N., Thomas, D. C., Tran, T. B., Utteridge, T. M. A., van Welzen, P. C., ... Wong, K. M. (2019). Progress on Southeast Asia's Flora projects. *Gardens' Bulletin Singapore*, 71(2), 267–319. [https://doi.org/10.26492/gbs71\(2\).2019-02](https://doi.org/10.26492/gbs71(2).2019-02)
- Milesi, P., Kastally, C., Dauphin, B., Cervantes, S., Bagnoli, F., Budde, K. B., Cavers, S., Ojeda, D. I., Fady, B., Faivre-Rampant, P., González-Martínez, S. C., Grivet, D., Gugerli, F., Jorge, V., Lesur-Kupin, I., Olsson, S., Opgenoorth, L., Pinosio, S., Plomion, C., ... Pyhäjärvi, T. (2023). Synchronous effective population size changes and genetic stability of forest trees through glacial cycles. *BioRxiv*, 2023.01.05.522822. <https://doi.org/10.1101/2023.01.05.522822>
- Moles, A. T., Ackerly, D. D., Tweddle, J. C., Dickie, J. B., Smith, R., Leishman, M. R., Mayfield, M. M., Pitman, A., Wood, J. T., & Westoby, M. (2007). Global patterns in seed size. *Global Ecology and Biogeography*, 16(1), 109–116. <https://doi.org/10.1111/j.1466-8238.2006.00259.x>
- Mongabay. (2019). *Total number of bird species by country*. Accessed in 23 May 2023 in <https://Rainforests.Mongabay.Com/03birds.Htm>.

- Moreno, M. A., Boshier, D., Tarazi, R., Tambarussi, E. V., Aguiar, B. I., Ferraz, E. M., Kageyama, P. Y., & Sebbenn, A. M. (2022). Long distance pollen dispersal and intrapopulation genetic structure determined by clonal diversity in *Hymenaea stigonocarpa* populations of the Brazilian savanna. *Trees*, 36(1), 471–484. <https://doi.org/10.1007/s00468-021-02222-8>
- Moreno-Gonzalez, R. (2022). Vegetation responses to past volcanic disturbances at the *Araucaria araucana* forest-steppe ecotone in northern Patagonia. *Ecology and Evolution*, 12(10). <https://doi.org/10.1002/ece3.9362>
- Morley, R. J., & Flenley, J. R. (1987). Late Cainozoic Vegetational and Environmental Changes in the Malay Archipelago. In T. C. Whitmore (Ed.), *Biogeographical Evolution in the Malay Archipelago* (1st ed., Vol. 4, pp. 50–59). Oxford Monograph.
- Myers, N. (1988). Threatened biotas: “Hot spots” in tropical forests. *The Environmentalist*, 8(3), 187–208. <https://doi.org/10.1007/BF02240252>
- Nakanishi, A., Goto, S., Sumiyoshi, C., & Isagi, Y. (2021). Directional seed and pollen dispersal and their separate effects on anisotropy of fine-scale spatial genetic structure among seedlings in a dioecious, wind-pollinated, and wind-dispersed tree species, *Cercidiphyllum japonicum*. *Ecology and Evolution*, 11(12), 7754–7767. <https://doi.org/10.1002/ece3.7609>
- Nathan, R., Schurr, F. M., Spiegel, O., Steinitz, O., Trakhtenbrot, A., & Tsoar, A. (2008). Mechanisms of long-distance seed dispersal. *Trends in Ecology & Evolution*, 23(11), 638–647. <https://doi.org/10.1016/j.tree.2008.08.003>
- Navascués, M., Vaxevanidou, Z., Gonzalez-Martinez, S. C., Climent, José., Gil, L., & Emerson, B. C. (2006). Chloroplast microsatellites reveal colonization and metapopulation dynamics in the Canary Island pine. *Molecular Ecology*, 15(10), 2691–2698. <https://doi.org/10.1111/j.1365-294X.2006.02960.x>
- Nei, M., & Roychoudhury, A. K. (1974). Sampling variances of heterozygosity and genetic distance. *Genetics*, 76(2), 379–390. <https://doi.org/10.1093/genetics/76.2.379>
- Nieto-Blázquez, M. E., Quiroga, M. P., Premoli, A. C., & Roncal, J. (2022). *Podocarpus* in the palaeogeographically complex island of Hispaniola: A stepping-stone colonization and conservation recommendations. *Diversity and Distributions*, 28(2), 214–226. <https://doi.org/10.1111/ddi.13444>
- Nilsson, C., Brown, R. L., Jansson, R., & Merritt, D. M. (2010). The role of hydrochory in structuring riparian and wetland vegetation. *Biological Reviews*, no-no. <https://doi.org/10.1111/j.1469-185X.2010.00129.x>
- Ochoa-Zavala, M., Jaramillo-Correa, J. P., Piñero, D., Nettel-Hernanz, A., & Núñez-Farfán, J. (2019). Contrasting colonization patterns of black mangrove

(*Avicennia germinans* (L.) L.) gene pools along the Mexican coasts. *Journal of Biogeography*, 46(5), 884–898. <https://doi.org/10.1111/jbi.13536>

O’Connell, J., Allen, J., & Hawkes, K. (2010). Pleistocene Sahul and the origins of seafaring. In A. Anderson, J. Barrett, & K. Boyle (Eds.), *The global origins and development of seafaring* (1st ed., Vol. 1, pp. 57–68). MacDonald Institute for Archeological Research.

Oliver, P. M., Iannella, A., Richards, S. J., & Lee, M. S. Y. (2017). Mountain colonisation, miniaturisation and ecological evolution in a radiation of direct-developing New Guinea Frogs (Choerophryne, Microhylidae). *PeerJ*, 5, e3077. <https://doi.org/10.7717/peerj.3077>

Parker, I. M., Rodriguez, J., & Loik, M. E. (2003). An Evolutionary Approach to Understanding the Biology of Invasions: Local Adaptation and General-Purpose Genotypes in the Weed *Verbascum thapsus*. *Conservation Biology*, 17(1), 59–72. <http://www.jstor.org/stable/3095273>

Péros, J., Cousins, P., Launay, A., Cubry, P., Walker, A., Prado, E., Peressotti, E., Wiedemann-Merdinoglu, S., Laucou, V., Merdinoglu, D., This, P., Boursiquot, J., & Doligez, A. (2021). Genetic diversity and population structure in *Vitis* species illustrate phylogeographic patterns in eastern North America. *Molecular Ecology*, 30(10), 2333–2348. <https://doi.org/10.1111/mec.15881>

Petit, R. J. (2004). Biological invasions at the gene level. *Diversity Distributions*, 10(3), 159–165. <https://doi.org/10.1111/j.1366-9516.2004.00084.x>

Ptak, R. (1992). The Northern Trade Route to the Spice Islands : South China Sea - Sulu Zone - North Moluccas (14th to early 16th century). *Archipel*, 43(1), 27–56. <https://doi.org/10.3406/arch.1992.2804>

Purugganan, M. D. (2022). What is domestication? *Trends in Ecology & Evolution*, 37(8), 663–671. <https://doi.org/10.1016/j.tree.2022.04.006>

Pybus, O. G., Rambaut, A., & Harvey, P. H. (2000). An Integrated Framework for the Inference of Viral Population History From Reconstructed Genealogies. *Genetics*, 155(3), 1429–1437. <https://doi.org/10.1093/genetics/155.3.1429>

Qiu, T., Andrus, R., Aravena, M.-C., Ascoli, D., Bergeron, Y., Berretti, R., Berveiller, D., Bogdziewicz, M., Boivin, T., Bonal, R., Bragg, D. C., Caignard, T., Calama, R., Camarero, J. J., Chang-Yang, C.-H., Cleavitt, N. L., Courbaud, B., Courbet, F., Curt, T., ... Clark, J. S. (2022). Limits to reproduction and seed size-number trade-offs that shape forest dominance and future recovery. *Nature Communications*, 13(1), 2381. <https://doi.org/10.1038/s41467-022-30037-9>

- R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- Raes, N., & Van Welzen, P. C. (2009). The demarcation and internal division of Flora Malesiana: 1857--present. *Blumea-Biodiversity, Evolution and Biogeography of Plants*, 54(1–2), 6–8.
- Rajora, O. P., & Zinck, J. W. R. (2021). Genetic Diversity, Structure and Effective Population Size of Old-Growth vs. Second-Growth Populations of Keystone and Long-Lived Conifer, Eastern White Pine (*Pinus strobus*): Conservation Value and Climate Adaptation Potential. *Frontiers in Genetics*, 12. <https://doi.org/10.3389/fgene.2021.650299>
- Razgour, O., Forester, B., Taggart, J. B., Bekaert, M., Juste, J., Ibáñez, C., Puechmaille, S. J., Novella-Fernandez, R., Alberdi, A., & Manel, S. (2019). Considering adaptive genetic variation in climate change vulnerability assessment reduces species range loss projections. *Proceedings of the National Academy of Sciences*, 116(21), 10418–10423. <https://doi.org/10.1073/pnas.1820663116>
- Reid, A. (1995). Humans and Forests in Pre-colonial Southeast Asia. *Environment and History*, 1(1), 93–110. <http://www.jstor.org/stable/20722966>
- Rico, Y., León-Tapia, M. Á., Zurita-Solís, M., Rodríguez-Gómez, F., & Vásquez-Morales, S. G. (2021). Influence of Pleistocene climatic oscillations on the phylogeography and demographic history of endemic vulnerable trees (section *Magnolia*) of the Tropical Montane Cloud Forest in Mexico. *PeerJ*, 9, e12181. <https://doi.org/10.7717/peerj.12181>
- Rifany, M. A. (2016). *Hubungan bentuk biji dengan karakteristik morfologi bibit tanaman pala (Myristica fragrans Houtt)*. Universitas Andalas.
- Rockhill, W. W. (1914). Notes on the Relations and Trade of China with the Eastern Archipelago and the Coast of the Indian Ocean during the Fourteenth Century: Part i. *T'oung Pao*, 15(3), 419–447. <http://www.jstor.org/stable/4526419>
- 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>

- Sánchez-Robles, J. M., García-Castaño, J. L., Balao, F., García, C., Terrab, A., & Talavera, S. (2022). Anthropogenic deforestation and climate dryness as drivers of demographic decline and genetic erosion in the southernmost European fir forests. *European Journal of Forest Research*, *141*(4), 649–663. <https://doi.org/10.1007/s10342-022-01467-3>
- Sang, T., & Ge, S. (2007). Genetics and phylogenetics of rice domestication. *Current Opinion in Genetics & Development*, *17*(6), 533–538. <https://doi.org/10.1016/j.gde.2007.09.005>
- Shackleton, C. M., Drescher, A., & Schlesinger, J. (2020). Urbanisation reshapes gendered engagement in land-based livelihood activities in mid-sized African towns. *World Development*, *130*, 104946. <https://doi.org/https://doi.org/10.1016/j.worlddev.2020.104946>
- Shackleton, C. M., Ticktin, T., & Cunningham, A. B. (2018). Nontimber forest products as ecological and biocultural keystone species. *Ecology and Society*, *23*(4). <https://www.jstor.org/stable/26796885>
- Slate, J., David, P., Dodds, K. G., Veenvliet, B. A., Glass, B. C., Broad, T. E., & McEwan, J. C. (2004). Understanding the relationship between the inbreeding coefficient and multilocus heterozygosity: theoretical expectations and empirical data. *Heredity*, *93*(3), 255–265. <https://doi.org/10.1038/sj.hdy.6800485>
- Smith, S. A., & Donoghue, M. J. (2008). Rates of Molecular Evolution Are Linked to Life History in Flowering Plants. *Science*, *322*(5898), 86–89. <https://doi.org/10.1126/science.1163197>
- Soares, L. A. S. S., Cazetta, E., Santos, L. R., França, D. de S., & Gaiotto, F. A. (2019). Anthropogenic Disturbances Eroding the Genetic Diversity of a Threatened Palm Tree: A Multiscale Approach. *Frontiers in Genetics*, *10*. <https://doi.org/10.3389/fgene.2019.01090>
- Soepadmo, E., Saw, L. G., Chung, R. C. K., & others. (2004). Tree flora of Sabah and Sarawak: Volume 5. *Tree Flora of Sabah and Sarawak: Volume 5*.
- Solís-Hernández, W., & Fuchs, E.-J. (2019). Effective gene flow patterns across a fragmented landscape in southern Costa Rica for *Symphonia globulifera* (Clusiaceae); a species with mobile seed and pollen dispersers. *Revista de Biología Tropical*, *67*(2SUPL), S95–S111. <https://doi.org/10.15517/rbt.v67i2SUPL.37209>
- Steeves, R. A. D. (2011). *An Intrageneric and Intraspecific Study of Morphological and Genetic Variation in the Neotropical Componeura and Virola (Myristicaceae)*. University of Guelph.

- Stetter, M. G., Gates, D. J., Mei, W., & Ross-Ibarra, J. (2017). How to make a domesticate. *Current Biology*, 27(17), R896–R900. <https://doi.org/10.1016/j.cub.2017.06.048>
- Stibig, H.-J., Achard, F., Carboni, S., Raši, R., & Miettinen, J. (2014). Change in tropical forest cover of Southeast Asia from 1990 to 2010. *Biogeosciences*, 11(2), 247–258. <https://doi.org/10.5194/bg-11-247-2014>
- Stuessy, T. F., Takayama, K., López-Sepúlveda, P., & Crawford, D. J. (2014). Interpretation of patterns of genetic variation in endemic plant species of oceanic islands. *Botanical Journal of the Linnean Society*, 174(3), 276–288. <https://doi.org/10.1111/boj.12088>
- Sujii, P. S., Tambarussi, E. V., Grando, C., de Aguiar Silvestre, E., Viana, J. P. G., Brancalion, P. H. S., & Zucchi, M. I. (2021). High gene flow through pollen partially compensates spatial limited gene flow by seeds for a Neotropical tree in forest conservation and restoration areas. *Conservation Genetics*, 22(3), 383–396. <https://doi.org/10.1007/s10592-021-01344-3>
- Sung, W., Ackerman, M. S., Miller, S. F., Doak, T. G., & Lynch, M. (2012). Drift-barrier hypothesis and mutation-rate evolution. *Proceedings of the National Academy of Sciences*, 109(45), 18488–18492. <https://doi.org/10.1073/pnas.1216223109>
- Taim, E. A. P. (2022). Persebaran Arca Buddha abad ke 7-9 Masehi: Hubungannya dengan kawasan awal Kerajaan Sriwijaya. *Naditira Widya*, 16(2), 95–106. <https://doi.org/10.24832/nw.v16i2.501>
- Tamaki, I., Kuze, T., Hirota, K., & Mizuno, M. (2016). Genetic variation and population demography of the landrace population of *Camellia sinensis* in Kasuga, Gifu Prefecture, Japan. *Genetic Resources and Crop Evolution*, 63(5), 823–831. <https://doi.org/10.1007/s10722-015-0286-7>
- Thompson, G. D., Richardson, D. M., Wilson, J. R. U., Bellstedt, D. U., & Le Roux, J. J. (2016). Genetic diversity and structure of the globally invasive tree, *Paraserianthes lophantha* subspecies *lophantha*, suggest an introduction history characterised by varying propagule pressure. *Tree Genetics & Genomes*, 12(2), 27. <https://doi.org/10.1007/s11295-016-0984-0>
- Timko, J. A., Waeber, P. O., & Kozak, R. A. (2010). The socio-economic contribution of non-timber forest products to rural livelihoods in Sub-Saharan Africa: knowledge gaps and new directions. *The International Forestry Review*, 12(3), 284–294. <http://www.jstor.org/stable/24310584>
- Toczydlowski, R. H., & Waller, D. M. (2019). Drift happens: Molecular genetic diversity and differentiation among populations of jewelweed (*Impatiens capensis* Meerb.) reflect fragmentation of floodplain forests. *Molecular Ecology*, 28(10), 2459–2475. <https://doi.org/10.1111/mec.15072>

- Trethowan, L. A., Arvidsson, C., & Bramley, G. L. C. (2022). Environmental stress influences Malesian Lamiaceae distributions. *Ecology and Evolution*, *12*(11). <https://doi.org/10.1002/ece3.9467>
- Tsumura, Y., & Ohba, K. (1993). Genetic structure of geographical marginal populations of *Cryptomeria japonica*. *Canadian Journal of Forest Research*, *23*(5), 859–863. <https://doi.org/10.1139/x93-112>
- USDA, A. R. S. N. P. G. S. (2023). *Germplasm Resources Information Network (GRIN Taxonomy)*.
- van Heerwaarden, J., Doebley, J., Briggs, W. H., Glaubitz, J. C., Goodman, M. M., de Jesus Sanchez Gonzalez, J., & Ross-Ibarra, J. (2011). Genetic signals of origin, spread, and introgression in a large sample of maize landraces. *Proceedings of the National Academy of Sciences*, *108*(3), 1088–1092. <https://doi.org/10.1073/pnas.1013011108>
- Van Steenis, C. G. G. J. (1950). The delimitation of Malaysia and its main plant geographical divisions. In C. G. G. J. Van Steenis (Ed.), *Flora Malesiana* (1st ed., Vol. 1, pp. cvii–cxvi). Noordhoff-Kolff n.v.
- Van Welzen, P. C., Parnell, J. A. N., & Slik, J. W. F. (2011). Wallace’s Line and plant distributions: two or three phytogeographical areas and where to group Java? *Biological Journal of the Linnean Society*, *103*(3), 531–545. <https://doi.org/10.1111/j.1095-8312.2011.01647.x>
- van Welzen, P. C., Turner, H., & Hovenkamp, P. (2003). Historical biogeography of Southeast Asia and the West Pacific, or the generality of unrooted area networks as historical biogeographic hypotheses. *Journal of Biogeography*, *30*(2), 181–192. <https://doi.org/10.1046/j.1365-2699.2003.00827.x>
- Vanden Abeele, S., Janssens, S. B., Piñeiro, R., & Hardy, O. J. (2021). Evidence of past forest fragmentation in the Congo Basin from the phylogeography of a shade-tolerant tree with limited seed dispersal: *Scorodophloeus zenkeri* (Fabaceae, Detarioideae). *BMC Ecology and Evolution*, *21*(1), 50. <https://doi.org/10.1186/s12862-021-01781-1>
- Vandergast, A. G., Gillespie, R. G., & Roderick, G. K. (2004). Influence of volcanic activity on the population genetic structure of Hawaiian Tetragnatha spiders: fragmentation, rapid population growth and the potential for accelerated evolution. *Molecular Ecology*, *13*(7), 1729–1743. <https://doi.org/10.1111/j.1365-294X.2004.02179.x>
- Wahlsteen, E., Avramidou, E. V., Bozic, G., Mediouni, R. M., Schuldt, B., & Sobolewska, H. (2023). Continental-wide population genetics and post-Pleistocene range expansion in field maple (*Acer campestre* L.), a subdominant temperate broadleaved tree species. *Tree Genetics & Genomes*, *19*(2), 15. <https://doi.org/10.1007/s11295-023-01590-1>

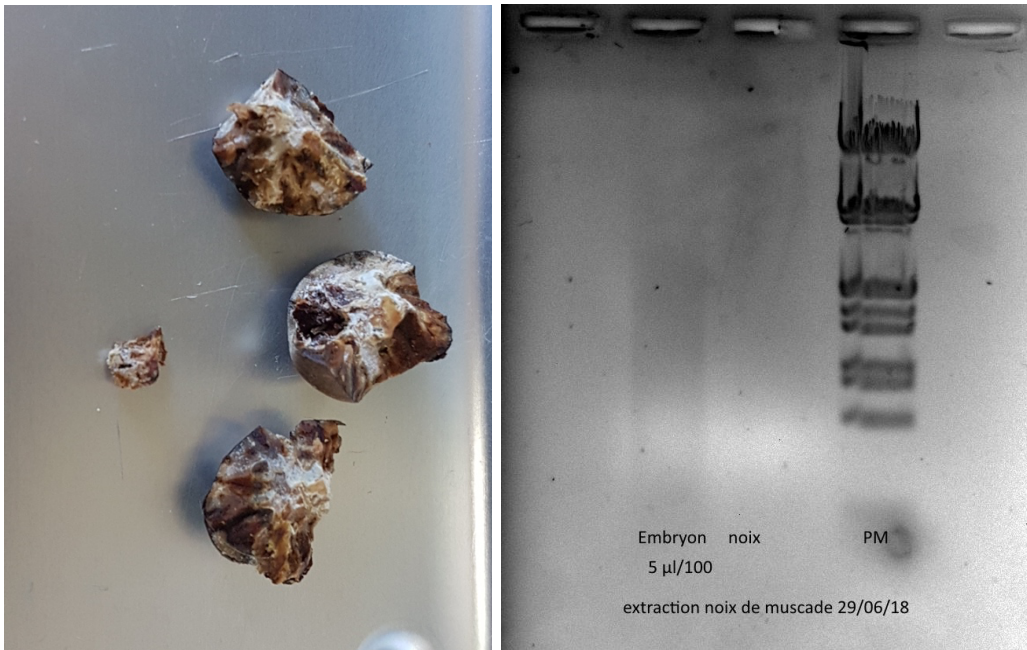
- Wahyuni, S., & Bermawie, N. (2020). Yield and fruit morphology of selected high productive Papua nutmeg trees (*Myristica argentea* Warb.). *IOP Conference Series: Earth and Environmental Science*, 418(1), 012032. <https://doi.org/10.1088/1755-1315/418/1/012032>
- Wallace, A. R. (1862). On the Physical Geography of the Malay Archipelago. *Proceedings of the Royal Geographical Society of London*, 7(5), 206–213.
- Wallace, A. R. (1880). *Island life* (1st ed., Vol. 1). Macmillan and Co.
- Wallace, A. R. (1886). *The Malay Archipelago : the land of the orang-utan and the bird of paradise ; a narrative of travel, with studies of man and nature*. Macmillan. <https://www.biodiversitylibrary.org/item/133732>
- Wang, L., Ji, Y., Hu, Y., Hu, H., Jia, X., Jiang, M., Zhang, X., Zhao, L., Zhang, Y., Jia, Y., Qin, C., Yu, L., Huang, J., Yang, S., Hurst, L. D., & Tian, D. (2019). The architecture of intra-organism mutation rate variation in plants. *PLOS Biology*, 17(4), e3000191. <https://doi.org/10.1371/journal.pbio.3000191>
- Waqar, Z., Moraes, R. C. S., Benchimol, M., Morante-Filho, J. C., Mariano-Neto, E., & Gaiotto, F. A. (2021). Gene Flow and Genetic Structure Reveal Reduced Diversity between Generations of a Tropical Tree, *Manilkara multifida* Penn., in Atlantic Forest Fragments. *Genes*, 12(12), 2025. <https://doi.org/10.3390/genes12122025>
- Watson, J. D., & Crick, F. H. C. (1953). Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature*, 171(4356), 737–738. <https://doi.org/10.1038/171737a0>
- Watts, J. (2021, January 12). *A sustainable pathway for Papuan nutmeg*. Kaleka. <https://medium.com/inobu/a-sustainable-pathway-for-papuan-nutmeg-d75dec0e087f>
- Weber, M. (1902). *Siboga Expeditie 1899-1900. Volume 1. Introduction et description de l'expédition*. <https://archimer.ifremer.fr/doc/00000/4745/>
- Weiland, R. J., & Cloos, M. (1996). Pliocene-Pleistocene asymmetric unroofing of the Irian fold belt, Irian Jaya, Indonesia: Apatite fission-track thermochronology. *GSA Bulletin*, 108(11), 1438–1449. [https://doi.org/10.1130/0016-7606\(1996\)108<1438:PPAUOT>2.3.CO;2](https://doi.org/10.1130/0016-7606(1996)108<1438:PPAUOT>2.3.CO;2)
- Weiss, E. (2002). *Spice crops*. CABI.
- Welzen, P. C., & Raes, N. (2011). The floristic position of Java. *Gardens' Bulletin Singapore*, 63, 335–345.
- Welzen, P. C., Slik, F., & Alahuhta, J. (2005). Plant distribution patterns and plate tectonics in Malesia. *Biol. Skr.*, 55, 199–217.

- Wessinger, C. A. (2021). From pollen dispersal to plant diversification: genetic consequences of pollination mode. *New Phytologist*, 229(6), 3125–3132. <https://doi.org/10.1111/nph.17073>
- Wicaksono, S. A., Russell, J. M., Holbourn, A., & Kuhnt, W. (2017). Hydrological and vegetation shifts in the Wallacean region of central Indonesia since the Last Glacial Maximum. *Quaternary Science Reviews*, 157, 152–163. <https://doi.org/https://doi.org/10.1016/j.quascirev.2016.12.006>
- Wilding, J. D., Zhu, W., Ross, Z. E., & Jackson, J. M. (2023). The magmatic web beneath Hawai'i. *Science*, 379(6631), 462–468. <https://doi.org/10.1126/science.ade5755>
- Williams, C. G. (2010a). Long-distance pine pollen still germinates after meso-scale dispersal. *American Journal of Botany*, 97(5), 846–855. <https://doi.org/10.3732/ajb.0900255>
- Williams, C. G. (2010b). Long-distance pine pollen still germinates after meso-scale dispersal. *American Journal of Botany*, 97(5), 846–855. <https://doi.org/10.3732/ajb.0900255>
- Williams, C. G., & Barnéoud, P. (2021). Live pine pollen in rainwater: reconstructing its long-range transport. *Aerobiologia*, 37(2), 333–350. <https://doi.org/10.1007/s10453-021-09697-5>
- Woodruff, D. S. (2003). Neogene marine transgressions, palaeogeography and biogeographic transitions on the Thai-Malay Peninsula. *Journal of Biogeography*, 30(4), 551–567. <https://doi.org/10.1046/j.1365-2699.2003.00846.x>
- 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. <http://www.jstor.org/stable/41503143>
- Wright, S. (1949). The genetical structure of a population. *Annals of Eugenics*, 15(1), 323–354. <https://doi.org/10.1111/j.1469-1809.1949.tb02451.x>
- Yoichi, W., Ono, E., Tsunamoto, Y., Matsuo, A., Suyama, Y., & Uehara, K. (2021). How co-distribution of two related azaleas (*Rhododendron*) developed in the Japanese archipelago: insights from evolutionary and demographic analyses. *Tree Genetics & Genomes*, 17(3), 30. <https://doi.org/10.1007/s11295-021-01512-z>
- Zeibig, F., Kilian, B., & Frei, M. (2022). The grain quality of wheat wild relatives in the evolutionary context. *Theoretical and Applied Genetics*, 135(11), 4029–4048. <https://doi.org/10.1007/s00122-021-04013-8>

- Zhu, S., Comes, H. P., Tamaki, I., Cao, Y., Sakaguchi, S., Yap, Z., Ding, Y., & Qiu, Y. (2020). Patterns of genotype variation and demographic history in *Lindera glauca* (Lauraceae), an apomict-containing dioecious forest tree. *Journal of Biogeography*, 47(9), 2002–2016. <https://doi.org/10.1111/jbi.13874>
- Zhu, Y. O., Siegal, M. L., Hall, D. W., & Petrov, D. A. (2014). Precise estimates of mutation rate and spectrum in yeast. *Proceedings of the National Academy of Sciences*, 111(22). <https://doi.org/10.1073/pnas.1323011111>
- Zohary, D., Hopf, M., & Weiss, E. (2012). *Domestication of Plants in the Old World*. Oxford University Press.
<https://doi.org/10.1093/acprof:osobl/9780199549061.001.0001>

CHAPTER 1

Microsatellites marker development for Indonesian nutmeg (*Myristica fragrans* Houtt.) and transferability to other *Myristicaceae* spp.



At IRD Montpellier, the genomic DNA purification procedure for developing microsatellite markers from *Myristica fragrans* involved the extraction from both seeds and embryos (left), followed by quality assessment (right). The seeds were sourced from Ambon, Indonesia, and transported to Montpellier, France in 2018 by Jérôme Duminil under a valid research permit issued by the National Research Agency and Innovation of the Republic of Indonesia. Throughout the sampling process in Indonesia, strict adherence to the Nagoya Protocol was maintained to ensure compliance. Pictures provided by Leïla Zekraoui (2018).

Chapter 1

Microsatellites marker development for Indonesian nutmeg (*Myristica fragrans* Houtt.) and transferability to other *Myristicaceae* spp.

(Published in **Molecular Biology Reports** volume 47, pages 4835–4840 (2020))

Jakty Kusuma^{1,2*}, Nora Scarcelli², Marie Couderc², Cédric Mariac², Leila
Zekraoui², Jérôme Duminil²

¹ Department of Plant Science, Politeknik Negeri Lampung, Lampung, Indonesia

² DIADE, IRD, Univ Montpellier, Montpellier, France

ABSTRACT

Myristica fragrans (Myristicaceae) is a tropical evergreen tree that yields the two famous spices: nutmeg and mace. Despite its socio-economic importance, the spatial distribution of its genetic diversity is barely documented. In this aim, 48 nuclear microsatellite markers were isolated of which 14 were polymorphic in *M. fragrans*. Number of alleles per locus ranged from two to six. The levels of observed and expected heterozygosity ranged from 0.038 to 0.929 and from 0.038 to 0.795 across loci, respectively. Transferability of these microsatellites to other *Myristica* species (*M. fatua*, *M. argentea*, and *M. crassipes*) and *Myristicaceae* species (*Horsfieldia palauensis*) was tested and successful. These new microsatellites will be useful for future investigation on genetic diversity and population structure of *M. fragrans* and phylogenetically-related species.

Key words: *Myristica*, nutmeg, microsatellites, fruit tree species, spices.

INTRODUCTION

Myristica fragrans Houtt, belongs to the Myristicaceae family. This dioecious tropical tree is considered as one of the most important food tree species to Indonesian people. It is utilized as spices, medicine [1] and raw material for perfume industries [2]. The species is most probably native from Banda islands in the Moluccas (the so-called ‘spice islands’) in Indonesia [3]. Given its economic value, the species has been widely propagated and cultivated across tropical regions, particularly in India, Malaysia, Sri Lanka and the Caribbean [4]. Despite its socio-economic importance, the spatial distribution of its genetic diversity remains poorly known, which impedes the development of conservation and sustainable management and use strategies. In order to achieve such objectives, polymorphic molecular markers need to be available. In this study, 14 polymorphic nuclear microsatellite markers (SSR) were developed and characterized. Additionally, cross-amplification of these markers was validated in *M. fatua*, *M. argentea*, *M. crassipes* and another Myristicaceae species, *Horsfieldia palauensis*.

METHODS AND RESULTS

Plant material and extraction

A total of 89 samples were collected as fresh leaves or from New York Botanical Garden (NY) and Herbier de l’Université de Montpellier (MPU): 75 individuals of *M. fragrans*, the cultivated species, five individuals of *M. fatua*, five individuals of *M. argentea* and two individuals of *M. crassipes*, and two individuals of *H. palauensis*, a species belonging to the Myristicaceae family (Supplementary File 1). Fresh leaves were collected and immediately dried in silica gel. *Myristica*

fragrans, *M. fatua*, and *M. argentea* samples were collected in three different localities, corresponding to three different islands (Ambon, Banda Besar, and Banda Naira) in Moluccas, Indonesia. All DNA were extracted following the protocol of Mariac et al [5].

Microsatellite development

A genomic library was prepared following Mariac *et al.* [6] on one *M. fragrans* sample (JK 131, Appendix 1), and then paired-end sequenced using Illumina MiSeq v2 reagents and 2×250 bp at the CIRAD Genotyping Platform (Montpellier, France). Script used for the following activities are available on Appendix 2. Illumina adapters were removed using CUTADAPT 1.8 [7] with parameters as follow: overlap length = 7, minimum length = 35 and quality = 20. Reads under a quality threshold of 30 were discarded using a PERL script (https://github.com/SouthGreenPlatform/arcad-hts/blob/master/scripts/arcad_hts_2_Filter_Fastq_On_Mean_Quality.pl). R1 and R2 reads were then paired using a PERL script (https://github.com/SouthGreenPlatform/arcad-hts/blob/master/scripts/arcad_hts_3_synchronized_paired_fastq.pl). We obtained 634906 paired reads that were then joined using FLASH v1.2.11 [8] and produced a total of 182869 joined reads. Microsatellite repeat units were identified and primers were designed using default setting in QDD software v3.1.2 [9]. From the 2374 SSR identified, 48 loci were selected according to the following criteria: (1) perfect (i.e., not compound) microsatellite di- or trinucleotide repeat units; (2)

repeats from 10 to 20; (3) primers situated at least 20 bp from the microsatellite region; (4) PCR product length between 90 bp and 300 bp.

Microsatellite amplification test

The 48 selected primer pairs were tested using three samples of *M. fragrans* (samples JK 131, JK 161, and JK 216; Appendix 1). A labeled tail (with either 6-FAM, NED, VIC or PET dye) was added to the 5' end of the forward primer of each locus in order to allow co-amplification of the loci in multiplex (Table 1). PCR reactions were done in a total volume of 10 μ L as follows: 2 μ L of 5X Colorless GoTaq[®] buffer (Promega, France) with final concentration at 1X (1.5 mM MgCl₂), 0.2 μ L of dNTP (Promega, France) at 10 mM, 0.05 μ L of GoTaq[®] polymerase (Promega, France) at 5 U/ μ L, 0.5 μ L of each primer at 10 mM, 2 μ L DNA template (of ca. 15-20 ng/ μ L), and water qs 10 μ L. PCR conditions were as follows: initial denaturation at 95 °C (5 min), 30 cycles of denaturation at 95 °C (23 s), annealing at 57 °C (1 min 30 s), elongation at 72 °C (1 min), and a final extension step at 72 °C (10 min). The quality of the PCR products was checked using a QIAxcel Advanced system (QIAGEN). Only the 40 loci showing consistent amplification in all three samples of *M. fragrans* were kept for the following analyses.

The polymorphism of the 40 loci was first tested on six *M. fragrans* individuals, i.e. on two individuals per population (three populations corresponding to three different islands; see details in Appendix 1). PCR amplifications were done in four multiplexes (details not shown) in a total volume of 10 μ L for each multiplex: 0.15 μ L of fluorescent labeled primers [10] at 10 μ M, 0.15 μ L of the reverse and 0.10 μ L of the forward at 10 μ M for both primers, 7.5 μ L Type-it Microsatellite PCR Master Mix (Qiagen), 1 μ L DNA template and water qs 10 μ L.

PCR conditions were identical to the amplification test (see above). The genotyping was done on an ABI 3500 XL (Applied Biosystems) at CIRAD (Montpellier, France) using 1 μ L of PCR product, 12 μ L of Hi-Di Formamide (Applied Biosystems) and 0.3 μ L GeneScan™ 500 LIZ™ dye Size Standard (Applied Biosystems). Peaks on each locus were analyzed using Geneious Prime® 2019.2.3 [11]. Loci that were difficult to read as well as monomorphic loci were excluded. In total, 14 primer pairs were kept and were combined in two multiplex reactions (the list of loci grouped in each multiplex is provided in Table 2) using Multiplex Manager 1.0 software [12].

Microsatellite diversity analysis

An analysis of genetic diversity was then performed on a larger set of samples of the three *M. fragrans* populations (22 to 29 samples per population; Appendix 1). The 14 primer pairs were amplified in two multiplexes (details provided in Table 2) in a final reaction volume of 15 μ L, consisted with 0.15 μ L of fluorescent labeled primers at 10 μ M, 0.15 μ L of the reverse and 0.10 μ L of the forward at 10 μ M for both microsatellite primers, 7.5 μ L Type-it Multiplex PCR Master Mix (Qiagen), 1 μ L DNA template and water qs 10 μ L. PCR conditions were identical to the amplification test (see above).

Genotypes were analyzed locus per locus using the automatic procedure implemented in Geneious Prime® 2019.2.3 [11] and manually corrected. Locus Myr42 showed two duplicated loci and was divided into two different loci (Myr42a and Myr42b) to be included in the following analyses. The genetic diversity was evaluated in each population by estimating the number of alleles per locus (A),

expected and observed heterozygosity (H_E and H_O , respectively) and the fixation index F_{IS} using the program SPAGeDi 1.5 [13], whereas frequencies of null alleles (r) were calculated using INEST 2.2 [14].

For each locus, the number of alleles ranged between two to six in all populations investigated (Table 2). The expected heterozygosity of the 14 loci ranged from 0.038 to 0.696 in the Ambon population, from 0.304 to 0.768 in the Banda Besar population and from 0.344 to 0.795 in the Banda Naira population. Mean expected heterozygosity was 0.539, 0.589, and 0.611 for each population respectively. Observed heterozygosity varied between 0.038 to 0.929 (Table 2). Some loci were presenting an excess of homozygosity (F_{IS} significantly superior to zero), but this was mainly due to the presence of null alleles.

Transferability of the markers in M. fatua, M. argentea, M. crassipes, and H. palauensis

Cross-amplification of the markers was tested in three other *Myristica* species (*M. fatua*, *M. argentea*, and *M. crassipes*) and another Myristicaceae species (*Horsfieldia palauensis*), using the same PCR conditions as before. Fifteen individuals from five different species were successfully genotyped (Table 3). All of these markers were found polymorphic in species *M. fatua* and *M. argentea*. Four and three markers did not amplified respectively in *M. crassipes* and *H. palauensis*, although the number of samples were low. Moreover, four loci were monomorphic in *H. palauensis*.

CONCLUSIONS

Fourteen polymorphic microsatellite markers were developed for an economically important food tree species, *Myristica fragrans*. These microsatellite markers will be useful for assessing species' genetic diversity and population structure as well in phylogenetically-related species given the high transferability of these markers.

REFERENCES

- [1] Arumugam G, Purushotham B, and Swamy MK (2019) *Myristica fragrans* Houtt.: Botanical, Pharmacological, and Toxicological Aspects. *Natural Bio-active Compounds* 81-106.
- [2] Periasamy G, Karim A, Gibrelibanos M, Gebremedhin G, and Gilani AH (2016) Nutmeg (*Myristica fragrans* Houtt.) Oils. *Essential Oils in Food Preservation, Flavor and Safety* pp. 607–616.
- [3] Rumphius GE (1741–1750). *Herbarium Amboinense ... pars prima (- sexta)*, ed. J. Burman. Amsterdam: Franciscum Changuion, Joannem Catuffe, Hermannum Uytwerf.
- [4] Swetha VP, Parvathy VA, Sheeja TE, and Sasikumar B (2017) Authentication of *Myristica fragrans* Houtt. using DNA barcoding. *Food Control*, vol. 73, pp. 1010–1015.
- [5] Mariac C, Luong V, Kapran I *et al.* (2006) Diversity of wild and cultivated pearl millet accessions (*Pennisetum glaucum* [L.] R. Br.) in Niger assessed by microsatellite markers. *Theor. Appl. Genet.*, vol. 114, no. 1, pp. 49–58
- [6] Mariac C, Scarcelli N, Pouzadou J *et al.* (2014) Cost-effective enrichment

- hybridization capture of chloroplast genomes at deep multiplexing levels for population genetics and phylogeography studies. *Mol. Ecol. Resour.*, vol. 14, no. 6, pp. 1103–1113.
- [7] Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, vol. 17, no. 1, p. 10.
- [8] Magoc T and Salzberg SL (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics*, vol. 27, no. 21, pp. 2957–2963.
- [9] Meglécz E, Costedoat C, Dubut V *et al.* (2010) QDD: a user-friendly program to select microsatellite markers and design primers from large sequencing projects. *Bioinformatics*, vol. 26, no. 3, pp. 403–404.
- [10] Culley TM, Weller SG, Sakai AK, Putnam KA (2008) Characterization of microsatellite loci in the Hawaiian endemic shrub *Schiedea adamantis* (Caryophyllaceae) and amplification in related species and genera. *Mol. Ecol. Resour.*, vol. 8, no. 5, pp. 1081–1084.
- [11] Kearse M *et al* (2012) Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, vol. 28, no. 12, pp. 1647–1649.
- [12] Holleley CE and Geerts PG (2009) Multiplex Manager 1.0: a cross-platform computer program that plans and optimizes multiplex PCR. *Biotechniques*, vol. 46, no. 7, pp. 511–517.
- [13] Hardy OJ, Vekemans X (2002) Spagedi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol. Ecol. Notes*, vol. 2, no. 4, pp. 618–620.

- [14] Chybicki IJ, Burczyk J (2009) Simultaneous Estimation of Null Alleles and Inbreeding Coefficients. *J. Hered.*, vol. 100, no. 1, pp. 106–113.
- [15] Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. *Nature biotechnology*, 18, 233-234.

FUNDING/ACKNOWLEDGEMENT

The authors would like to thank Gregory M. Plunkett, Sean Thackurdeen and Michael J. Balick from New York Botanical Garden (NY) for having provided samples of *H. palauensis*. The authors also thank Herbier de l'Université de Montpellier (MPU) through Jérôme Munzinger for having provided samples of *M. crassipes*. Authors also greatly appreciate the support from the Indonesia Endowment Fund for Education (LPDP). This project was supported by Agropolis Fondation under the reference ID 1502-503 and 1502-504 through the « Investissements d'avenir » programme (Labex Agro: ANR-10-LABX-0001-01), under the frame of I-SITE MUSE (ANR-16-IDEX-0006).

CONFLICTS OF INTEREST

The authors declare that they have no conflict of interest

ETHICS APPROVAL

The authors declare no ethics approval needed.

CONSENT TO PARTICIPATE

The authors declare no consent to participate needed.

CONSENT FOR PUBLICATION

The authors guarantee that the contribution to the work has not been previously published elsewhere, or that if it has been published in whole or in part, any permission necessary to publish it in the work has been obtained and provided.

AVAILABILITY DATA AND MATERIAL

DNA sequences of microsatellites described in this paper were deposited in GenBank (Table 1).

CODE AVAILABILITY

Scripts used in this paper were described in Supplementary File 2.

AUTHORS' CONTRIBUTIONS

All authors contributed to the study. Material preparation and data collection were performed by JK and JD. JK, MC, LZ, and CM undertook the experiment. JK, JD and NS analyzed the data. The first draft of the manuscript was written by JK and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Tables

Table 1. Characteristics of 14 microsatellite loci developed in *Myristica fragrans*.

Locus	Primer sequences (5'-3') ^A	Labelled primer	Tm (°C)	Allele size range (bp)	Amplification size (bp)	Genebank accession
Myr10	F: TAGGAGTGCAGCAAGCATTACTCCCGCGACCACTTAAA R: AGCAGAACCTACAAGAAGCAA	NED	57	132-164	135	MT175707
Myr12	F: CTAGTTATTGCTCAGCGGTAGAGTGACGGGTTGTATGCA R: AGGAATGACGATCGAGGATAGA	PET	57	155-161	140	MT175708
Myr13	F: TGTA AACGACGGCCAGTCCTTTCACATGGGTTGAAATGC R: TCCATCAAGTCCGCAGAACG	6-FAM	57	146-164	141	MT175709
Myr18	F: TAGGAGTGCAGCAAGCATTACGCCGATTGGACCTCATT R: ACTCGTCTGGATGACTACAGC	NED	57	173-191	148	MT1757010
Myr23	F: CACTGCTTAGAGCGATGCGCCTCCTGAAACCCGTTGTA R: TGATGAGAGCTACTGGCCTTG	VIC	57	153-175	156	MT1757011
Myr26	F: TAGGAGTGCAGCAAGCATAGTAAGTGTATCCAGCCGCC R: CGACCTTTCGTCTCCTGGTA	NED	57	173-185	160	MT1757012
Myr27	F: CACTGCTTAGAGCGATGCAGTGGGTTGGCGATTTAGGT R: TGGCCAAGGAGATACCATTCT	VIC	57	175-197	161	MT1757013
Myr28	F: CTAGTTATTGCTCAGCGGTTCCAAGATCCACAGACCAGC R: CAAAGCGAACGAACACCCGAG	PET	57	169-203	163	MT1757014
Myr29	F: TGTA AACGACGGCCAGTAGTTGCAATTTGAGCTCCAGA R: ATGAAGCACACGCACATTGT	6-FAM	57	172-198	165	MT1757015
Myr32	F: CTAGTTATTGCTCAGCGGTTGGCCATTAACACATCACTCCA R: GCATCCAAACAGAGCCTCAG	PET	57	187-211	170	MT1757016
Myr36	F: CTAGTTATTGCTCAGCGGTTGGGCAAACCTTTCAAGGACA R: CTAGGGCATCTGCTGGTGAG	PET	57	218-242	180	MT1757017
Myr42*	F: TAGGAGTGCAGCAAGCATCAACTGTGGGATCTGCCAGA R: GGCCGCTTGACTGTCAACTT	NED	57	223-237 241-251	203	MT1757018
Myr43	F: CACTGCTTAGAGCGATGCCACATGCAGGTAGTCCTCGT R: ACGCTAGAGATGGTTGCCAG	VIC	57	214-242	212	MT1757019
Myr47	F: CACTGCTTAGAGCGATGCGGATCACGATATAAATGGCAGC R: GGGTATAAACGGGAAAGGGC	VIC	57	247-269	240	MT1757020

^A The universal linkers attached to the forward primers are underlined (TGTAAAACGACGGCCAGT, [15]; TAGGAGTGCAGCAAGCAT, CACTGCTTAGAGCGATGC and CTAGTTATTGCTCAGCGGT Culley et al. [10]); *duplicated locus
Table 2. Results of initial screening of 14 microsatellite loci developed in *Myristica fragrans*.

Multiplex	Locus	Ambon (<i>N</i> = 26)					Banda Besar (<i>N</i> = 22)					Banda Naira (<i>N</i> = 27)				
		A	H _o	H _e	<i>F</i> _{IS}	<i>r</i>	A	H _o	H _e	<i>F</i> _{IS}	<i>r</i>	A	H _o	H _e	<i>F</i> _{IS}	<i>r</i>
Multiplex- 1	Myr10	3	0.308	0.347	0.117	0.01±0.03	2	0.273	0.304	0.106	0.02±0.06	4	0.444	0.533	0.169	0.06±0.06
	Myr12	2	0.269	0.237	-0.136	0.00±0.00	2	0.591	0.459	-0.294	0.00±0.00	2	0.370	0.452	0.185	0.05±0.06
	Myr23	5	0.577	0.671	0.144	0.06±0.05	5	0.773	0.765	-0.010	0.00±0.00	6	0.741	0.795	0.071	0.00±0.01
	Myr26	4	0.654	0.578	-0.133	0.00±0.00	4	0.273	0.324	0.163	0.05±0.06	4	0.593	0.584	-0.015	0.00±0.00
	Myr29	4	0.577	0.696	0.174	0.04±0.06	5	0.864	0.745	-0.163	0.00±0.00	5	0.929	0.719	-0.295**	0.00±0.00
	Myr32	4	0.462	0.676	0.322*	0.11±0.05	3	0.455	0.512	0.116	0.02±0.06	3	0.778	0.544	-0.441*	0.00±0.00
	Myr43	4	0.615	0.670	0.084	0.03±0.05	6	0.773	0.746	-0.036	0.00±0.00	6	0.778	0.735	-0.059	0.00±0.00
Multiplex- 2	Myr13	4	0.615	0.684	0.102	0.05±0.06	4	0.636	0.737	0.140	0.00±0.00	4	0.704	0.712	0.012	0.00±0.00
	Myr18	2	0.038	0.038	-0.000	0.00±0.00	2	0.386	0.371	0.008	0.28±0.10	3	0.333	0.344	0.034	0.25±0.09
	Myr27	3	0.731	0.584	-0.257	0.00±0.00	4	0.636	0.627	-0.014	0.00±0.00	3	0.741	0.677	-0.096	0.00±0.00
	Myr28	5	0.769	0.677	-0.139	0.00±0.00	6	0.810	0.768	-0.054	0.08±0.07	6	0.630	0.690	0.090	0.04±0.04
	Myr36	3	0.231	0.548	0.584**	0.20±0.05	4	0.227	0.700	0.681**	0.27±0.04	4	0.481	0.724	0.340**	0.13±0.05
	Myr42a	3	0.385	0.561	0.320	0.09±0.05	3	0.727	0.576	-0.270	0.00±0.00	3	0.481	0.552	0.131	0.04±0.06
	Myr42b	3	0.500	0.518	0.036	0.00±0.00	3	0.636	0.586	-0.087	0.00±0.00	3	0.556	0.634	0.127	0.05±0.05
	Myr47	4	0.577	0.595	0.031	0.02±0.04	4	0.628	0.606	-0.127	0.00±0.00	3	0.370	0.469	0.215	0.08±0.06
Multilocus average		3.53	0.487	0.539	0.098*	0.04±0.02	3.80	0.582	0.589	0.013	0.04±0.02	3.93	0.595	0.611	0.027	0.04±0.03

N: number of individuals sampled, *A* number of alleles, H_o observed heterozygosity, H_e expected heterozygosity, *F*_{IS} inbreeding coefficient, *r* frequency of null alleles.

Significant deviation from Hardy–Weinberg equilibrium at **p* < 0.05, ***p* < 0.01

Table 3. Alleles size ranges observed in cross-amplification of microsatellite loci isolated from *Myristica fragrans* and tested in four additional species.

Locus	<i>M. fatua</i> (<i>N</i> = 5)	<i>M. argentea</i> (<i>N</i> = 5)	<i>M. crassipes</i> (<i>N</i> = 2)	<i>H. palauensis</i> (<i>N</i> = 2)
Myr10	126-154 bp	132-162 bp	134-152 bp	128 bp
Myr12	155-161 bp	153-161 bp	155-161 bp	155 bp
Myr23	153-173 bp	165-189 bp	149-171 bp	175 bp
Myr26	177-185 bp	177-185 bp	181-191 bp	181-183 bp
Myr29	182-198 bp	178-194 bp	172-182 bp	174-190 bp
Myr32	193-209 bp	195-245 bp	193-199 bp	-
Myr43	214-240 bp	214-246 bp	218-224 bp	216-236 bp
Myr13	146-160 bp	146-160 bp	146-152 bp	148-150 bp
Myr18	173-191 bp	175-183 bp	173-183 bp	173 bp
Myr27	179-191 bp	179-197 bp	-	185-205 bp
Myr28	169-203 bp	191-207 bp	-	-
Myr36	202-238 bp	228-242 bp	-	-
Myr42a	223-235 bp	223-233 bp	217-235 bp	205-223 bp
Myr42b	241-245 bp	235-245 bp	-	245-251 bp
Myr47	247-267 bp	263-275 bp	269-285 bp	251-273 bp

N = number of individuals tested; - = no amplification

Supplementary Online Materials

Supplementary File 1. Voucher and locality information for the samples used in this study

Sample ID	Species	Sample type	Herbarium voucher ID	Country	Locality	Latitude	Longitude	Collector	Collection date	Used for amplification test	Used for diversity analysis	Used for transferability test
JK 115	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Ambon	-3.75236	128.01841	J. Duminil & J. Kusuma	20-05-2018	No	Yes	No
JK 117	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Ambon	-3.75260	128.01863	J. Duminil & J. Kusuma	20-05-2018	No	Yes	No
JK 116	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Ambon	-3.75246	128.01862	J. Duminil & J. Kusuma	20-05-2018	No	Yes	No
JK 133	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Ambon	-3.75309	128.02175	J. Duminil & J. Kusuma	20-05-2018	No	Yes	No
JK 131	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Ambon	-3.75303	128.02163	J. Duminil & J. Kusuma	20-05-2018	Yes	Yes	No
JK 127	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Ambon	-3.75327	128.02144	J. Duminil & J. Kusuma	20-05-2018	No	Yes	No
JK 132	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Ambon	-3.75301	128.02175	J. Duminil & J. Kusuma	20-05-2018	No	Yes	No
JK 109	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Ambon	-3.75425	128.02058	J. Duminil & J. Kusuma	20-05-2018	No	Yes	No
JK 112	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Ambon	-3.75265	128.01838	J. Duminil & J. Kusuma	20-05-2018	No	Yes	No
JK 123	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Ambon	-3.75327	128.01997	J. Duminil & J. Kusuma	20-05-2018	No	Yes	No
JK 124	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Ambon	-3.75328	128.02014	J. Duminil & J. Kusuma	20-05-2018	No	Yes	No
JK 130	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Ambon	-3.75310	128.02151	J. Duminil & J. Kusuma	20-05-2018	No	Yes	No

JK 119	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Ambon	-3.75289	128.01848	J. Duminil & J. Kusuma	20-05-2018	No	Yes	No
JK 120	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Ambon	-3.75313	128.01936	J. Duminil & J. Kusuma	20-05-2018	No	Yes	No
JK 108	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Ambon	-3.75423	128.02038	J. Duminil & J. Kusuma	20-05-2018	No	Yes	No
JK 114	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Ambon	-3.75252	128.01827	J. Duminil & J. Kusuma	20-05-2018	No	Yes	No
JK 122	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Ambon	-3.75304	128.01942	J. Duminil & J. Kusuma	20-05-2018	No	Yes	No
JK 134	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Ambon	-3.75277	128.02163	J. Duminil & J. Kusuma	20-05-2018	No	Yes	No
JK 126	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Ambon	-3.75322	128.02133	J. Duminil & J. Kusuma	20-05-2018	No	Yes	No
JK 121	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Ambon	-3.75312	128.01960	J. Duminil & J. Kusuma	20-05-2018	No	Yes	No
JK 135	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Ambon	-3.75265	128.02176	J. Duminil & J. Kusuma	20-05-2018	No	Yes	No
JK 118	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Ambon	-3.75295	128.01851	J. Duminil & J. Kusuma	20-05-2018	No	Yes	No
JK 128	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Ambon	-3.75354	128.02121	J. Duminil & J. Kusuma	20-05-2018	No	Yes	No
JK 125	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Ambon	-3.75306	128.02030	J. Duminil & J. Kusuma	20-05-2018	No	Yes	No
JK 113	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Ambon	-3.75257	128.01838	J. Duminil & J. Kusuma	20-05-2018	No	Yes	No
JK 110	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Ambon	-3.75198	128.01570	J. Duminil & J. Kusuma	20-05-2018	No	Yes	No
JK 198	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Besar	-4.32538	129.52543	J. Kusuma & J. Duminil	14-11-2018	No	Yes	No
JK 209	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Besar	-4.32536	129.52558	J. Kusuma & J. Duminil	14-11-2018	No	Yes	No
JK 213	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Besar	-4.32535	129.52385	J. Kusuma & J. Duminil	14-11-2018	No	Yes	No
JK 206	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Besar	-4.32520	129.52559	J. Kusuma & J. Duminil	14-11-2018	No	Yes	No

JK 232	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Besar	-4.32450	129.52449	J. Kusuma & J. Duminil	14-11-2018	No	Yes	No
JK 208	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Besar	-4.32532	129.52568	J. Kusuma & J. Duminil	14-11-2018	No	Yes	No
JK 207	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Besar	-4.32523	129.52566	J. Kusuma & J. Duminil	14-11-2018	No	Yes	No
JK 227	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Besar	-4.32554	129.52388	J. Kusuma & J. Duminil	14-11-2018	No	Yes	No
JK 231	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Besar	-4.32450	129.52418	J. Kusuma & J. Duminil	14-11-2018	No	Yes	No
JK 217	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Besar	-4.32533	129.52366	J. Kusuma & J. Duminil	14-11-2018	No	Yes	No
JK 224	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Besar	-4.32562	129.52369	J. Kusuma & J. Duminil	14-11-2018	No	Yes	No
JK 223	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Besar	-4.32566	129.52364	J. Kusuma & J. Duminil	14-11-2018	No	Yes	No
JK 211	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Besar	-4.32539	129.52387	J. Kusuma & J. Duminil	14-11-2018	No	Yes	No
JK 216	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Besar	-4.32535	129.52371	J. Kusuma & J. Duminil	14-11-2018	Yes	Yes	No
JK 201	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Besar	-4.32518	129.52552	J. Kusuma & J. Duminil	14-11-2018	No	Yes	No
JK 215	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Besar	-4.32534	129.52380	J. Kusuma & J. Duminil	14-11-2018	No	Yes	No
JK 205	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Besar	-4.32520	129.52552	J. Kusuma & J. Duminil	14-11-2018	No	Yes	No
JK 203	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Besar	-4.32541	129.52538	J. Kusuma & J. Duminil	14-11-2018	No	Yes	No
JK 199	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Besar	-4.32535	129.52547	J. Kusuma & J. Duminil	14-11-2018	No	Yes	No
JK 210	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Besar	-4.32539	129.52389	J. Kusuma & J. Duminil	14-11-2018	No	Yes	No
JK 212	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Besar	-4.32540	129.52379	J. Kusuma & J. Duminil	14-11-2018	No	Yes	No
JK 202	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Besar	-4.32535	129.52535	J. Kusuma & J. Duminil	14-11-2018	No	Yes	No

JK 192	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Naira	-4.32591	129.54389	J. Kusuma & J. Duminil	13-11-2018	No	Yes	No
JK 168	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Naira	-4.30555	129.54073	J. Kusuma & J. Duminil	13-11-2018	No	Yes	No
JK 164	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Naira	-4.32569	129.54059	J. Kusuma & J. Duminil	13-11-2018	No	Yes	No
JK 161	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Naira	-4.32568	129.25073	J. Kusuma & J. Duminil	13-11-2018	Yes	Yes	No
JK 159	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Naira	-4.32565	129.54071	J. Kusuma & J. Duminil	13-11-2018	No	Yes	No
JK 163	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Naira	-4.32566	129.54073	J. Kusuma & J. Duminil	13-11-2018	No	Yes	No
JK 162	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Naira	-4.32571	129.54068	J. Kusuma & J. Duminil	13-11-2018	No	Yes	No
JK 167	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Naira	-4.32551	129.54071	J. Kusuma & J. Duminil	13-11-2018	No	Yes	No
JK 195	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Naira	-4.32591	129.54380	J. Kusuma & J. Duminil	13-11-2018	No	Yes	No
JK 188	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Naira	-4.32582	129.54409	J. Kusuma & J. Duminil	13-11-2018	No	Yes	No
JK 171	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Naira	-4.32566	129.54079	J. Kusuma & J. Duminil	13-11-2018	No	Yes	No
JK 166	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Naira	-4.32558	129.54045	J. Kusuma & J. Duminil	13-11-2018	No	Yes	No
JK 165	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Naira	-4.32561	129.54052	J. Kusuma & J. Duminil	13-11-2018	No	Yes	No
JK 191	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Naira	-4.32593	129.54412	J. Kusuma & J. Duminil	13-11-2018	No	Yes	No
JK 172	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Naira	-4.32585	129.54089	J. Kusuma & J. Duminil	13-11-2018	No	Yes	No
JK 189	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Naira	-4.32598	129.54406	J. Kusuma & J. Duminil	13-11-2018	No	Yes	No
JK 169	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Naira	-4.32555	129.54083	J. Kusuma & J. Duminil	13-11-2018	No	Yes	No
JK 176	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Naira	-4.32540	129.54081	J. Kusuma & J. Duminil	13-11-2018	No	Yes	No

JK 158	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Naira	-4.32546	129.54075	J. Kusuma & J. Duminil	13-11-2018	No	Yes	No
JK 174	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Naira	-4.32556	129.54088	J. Kusuma & J. Duminil	13-11-2018	No	Yes	No
JK 196	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Naira	-4.32561	129.54341	J. Kusuma & J. Duminil	13-11-2018	No	Yes	No
JK 160	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Naira	-4.32568	129.54072	J. Kusuma & J. Duminil	13-11-2018	No	Yes	No
JK 177	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Naira	-4.32528	129.54077	J. Kusuma & J. Duminil	13-11-2018	No	Yes	No
JK 194	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Naira	-4.32588	129.54383	J. Kusuma & J. Duminil	13-11-2018	No	Yes	No
JK 190	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Naira	-4.32590	129.54408	J. Kusuma & J. Duminil	13-11-2018	No	Yes	No
JK 175	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Naira	-4.32553	129.54090	J. Kusuma & J. Duminil	13-11-2018	No	Yes	No
JK 173	<i>Myristica fragrans</i>	Fresh leaves	-	Indonesia	Banda Naira	-4.32562	129.54090	J. Kusuma & J. Duminil	13-11-2018	No	Yes	No
JK 197	<i>Myristica fatua</i>	Fresh leaves	-	Indonesia	Banda Naira	-4.54958	129.90955	J. Kusuma & J. Duminil	13-11-2018	No	No	Yes
JK 233	<i>Myristica fatua</i>	Fresh leaves	-	Indonesia	Banda Besar	-4.54569	129.87830	J. Kusuma & J. Duminil	13-11-2018	No	No	Yes
JK 182	<i>Myristica fatua</i>	Fresh leaves	-	Indonesia	Banda Naira	-4.54891	129.91013	J. Kusuma & J. Duminil	13-11-2018	No	No	Yes
JK 253	<i>Myristica fatua</i>	Fresh leaves	-	Indonesia	Pulau Ay	-4.52211	129.77863	J. Kusuma & J. Duminil	13-11-2018	No	No	Yes
JK 229	<i>Myristica fatua</i>	Fresh leaves	-	Indonesia	Banda Besar	-4.54844	129.87769	J. Kusuma & J. Duminil	13-11-2018	No	No	Yes
JK 153	<i>Myristica argentea</i>	Fresh leaves	-	Indonesia	Ambon	-3.74382	127.92984	J. Duminil & J. Kusuma	20-05-2018	No	No	Yes
JK 105	<i>Myristica argentea</i>	Fresh leaves	-	Indonesia	Ambon	-3.65085	127.96763	J. Duminil & J. Kusuma	20-05-2018	No	No	Yes
JK 005	<i>Myristica argentea</i>	Fresh leaves	-	Indonesia	Ambon	-3.59808	128.14028	J. Duminil & J. Kusuma	20-05-2018	No	No	Yes
JK 007	<i>Myristica argentea</i>	Fresh leaves	-	Indonesia	Ambon	-3.59813	128.14024	J. Duminil & J. Kusuma	20-05-2018	No	No	Yes

JK 107	<i>Myristica argentea</i>	Fresh leaves	-	Indonesia	Ambon	-3.65091	127.96523	J. Duminil & J. Kusuma	20-05-2018	No	No	Yes
JK 347	<i>Horsfieldia palauensis</i>	Herbarium	4371	Palau	Koror State	-	-	Gregory M. Plunkett	22-03-2018	No	No	Yes
JK 353	<i>Horsfieldia palauensis</i>	Herbarium	4677	Palau	Melekeok State	-	-	Michael J. Balick	12-11-2013	No	No	Yes
JK 317	<i>Myristica crassipes</i>	Herbarium	2996	Papua New Guinea	Madang	-	-	Molino <i>et al.</i>	-	No	No	Yes
JK 318	<i>Myristica crassipes</i>	Herbarium	6876	Papua New Guinea	Madang	-	-	Munzinger <i>et al.</i>	-	No	No	Yes

Supplementary File 2. Scripts used for microsatellite development

CODELINE

Remove adapter

for i in *.fastq; do cutadapt-1.2.1/bin/cutadapt \$(< adapters.conf) -o cut\$i \$i ; done

Contents of adapters.conf

-b

AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCT
CTTCCGATCT

-b ttactatgccgctgggtgctctagatgtgagaaagggatgtgctgcgagaaggctaga

-b

AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTC
GCCGTATCATT

-b agatcggaagagcgggtcagcaggaatgccgagaccgatctcgtatgccgtcttctgcttg

-b tctagccttctcgccaagtcgtccttacggctctggctagagcatacggcagaagacgaac

-O 7

-m 35

-q 20

Quality filtering

for i in cut* ; do perl Filter_Fastq_On_Mean_Quality.pl -f \$i -o filt\$i ; done

##Pairing

perl [compare_fastq_paired_v5.pl](#) -f filtcutt_R1_paired.fastq -r
filtcut_R2_paired.fastq -of Sfiltcut_R1_paired.fastq -or Sfiltcut_R2_paired.fastq -
os Sfiltcut_single.fastq

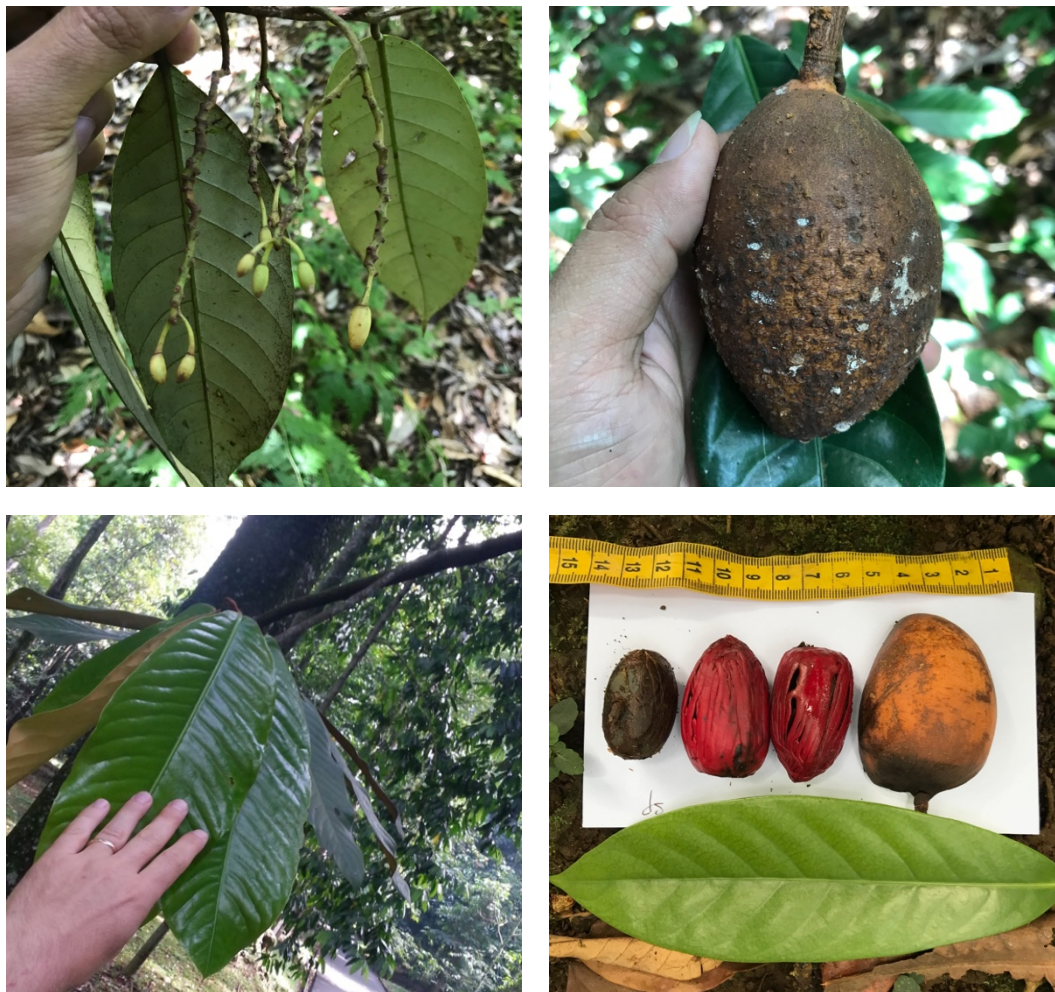
```
## Join reads
```

```
./flash Sfiltcut_R1_paired.fastq Sfiltcut_R2_paired.fastq
```

CHAPTER 2

Complete chloroplast genome of two Nutmeg species

Myristica argentea Warb 1891 and *Myristica fatua* Houtt 1774



This figure illustrates various aspects of *Myristica argentea* and *Myristica fatua*. It includes images of young male flowers (upper left) and mature fruit (upper right) of *M. argentea* collected in Fakfak, Papua, Indonesia. The lower left image showcases the leaves of *Myristica fatua* found in Bogor Botanical Garden, Bogor, Indonesia. The lower right image presents a composition of a young leaf and various components of the fruit, including the seed, arils, and flesh of *M. fatua* in North Sulawesi. All photographs of *M. argentea* and *M. fatua* were taken by Jaky Kusuma in 2019, while the photographs of *M. fatua* leaves were taken by Jérôme Duminil in 2018.

Chapter 2

Complete chloroplast genome of two Nutmeg species

Myristica argentea Warb. 1891 and *Myristica fatua* Houtt. 1774

(Myristicaceae)

(Published in **Mitochondrial DNA: Part B** VOL. 8, NO. 7, 751–755)

Jakty Kusuma^{1,2}, Marie Couderc¹, Nora Scarcelli¹, Jérôme Duminil¹

¹ DIADE, University of Montpellier, IRD, CIRAD, Montpellier, France

² Politeknik Negeri Lampung, Indonesia

ABSTRACT

Myristica argentea Warb. 1891 and *M. fatua* Houtt. 1774 are two South-East Asian food tree species. They are harvested from the wild or cultivated for local uses as a condiment (nutmeg and mace), medicine and source of wood. In this study, we reconstructed the complete chloroplast (cp) genomes of these two species from whole genome sequencing data using the Illumina NovaSeq platform. The genome sizes of *M. argentea* and *M. fatua* were respectively 155,871 base pairs (bp) and 155,898 bp, including 126 genes and an overall GC content of 39.20% in both species. Our study provides useful resources for future evolutionary research and diversity analysis of *Myristica* species.

Key words: chloroplast, phylogeography, evolution, genome

INTRODUCTION

The *Myristica* genus contains 172 species that are mainly distributed in Tropical Southeast Asia, the Malay Archipelago (Malesia) and the western Pacific [1-2]. *Myristica argentea* Warb. 1891, known as Papuan nutmeg or Fakfak nutmeg, is endemic to Papua island, Indonesia, and is named after the region where the species is occurring. The species is utilized for spices and medicinal purposes, by using the seed and its aril. In contrast with the relatively narrow distribution of *M. argentea*, *Myristica fatua* Houtt. 1774 presents a large distribution, from the Malesian archipelago to the Pacific islands [3]. Their distribution in these archipelagos makes them compelling subjects for investigating the biogeography of island plants in the region. Since their evolutionary history remains largely undocumented, the reconstruction of their chloroplast genome will help conducting phylogeographic studies.

MATERIALS

Leaf samples of *M. fatua* originated from Tafea, Vanuatu (19° 35' 52.8" S, 169° 20' 45.6" E) were collected from New York Botanical Garden herbaria (Figure 1). A specimen was deposited at New York Botanical Garden (<https://sweetgum.nybg.org/science/vh/specimen-details/?irn=3167800>; Gregory M. Plunkett, gplunkett@nybg.org) under the voucher number 03509312. Further, leaf samples of *M. argentea* (Figure 1) from a silica-dried sample collected in Ambon, Indonesia (3° 44' 38.112" S, 127° 55' 47.568" E). A specimen of *M. argentea* was deposited at UMR DIADE, IRD, France (https://dataverse.ird.fr/dataverse/umr_diade; jakty.kusuma@ird.fr) under the number JK0155.

METHODS

Genomic DNA were extracted using the MATAB protocol following Mariac *et al* [4]. We constructed genomic libraries following protocol described in Mariac *et al.* [5]. Libraries generated from real-time PCR were then sequenced on an Illumina NovaSeq (Illumina, San Diego, CA). Paired-end reads were filtered as described in Scarcelli *et al.* [6], and aligned to verified *M. fragrans* whole chloroplast genome (OL628762). Cleaned reads were then assembled using the *de novo* assembler GetOrganelle ver.1.7.6.1 [7; Supplementary Methods S1], and annotated using CPGAVAS2 (<http://www.herbalgenomics.org/cpgavas2/>) [8]. We visualized the circular genome using CHLOROPLLOT (<https://irscope.shinyapps.io/Chloroplot/>) [9]. All genomic information was deposited in GenBank (see Data availability statement).

We then constructed a phylogenetic tree including the complete cp genome sequences of these two species as well as seven available complete cp genome sequences of closely related species of the family Myristicaceae (*M. fragrans* OL628762 [10]; *M. yunnanensis* MN495964 [11]; *Knema tenuinervia* MN495962 [10]; *Knema furfuracea* MK285563 [12]; and *Horsfieldia amygdalina* MK285561 [13]), a species from Magnoliaceae (*Michelia maudiae* MK631950 [14]), and from Annonaceae, *Greenwayodendron suaveolens* MH924590 [15] as an outgroup. MAFFT (<https://mafft.cbrc.jp/alignment/server/index.html>) ver. 7 [16] and RAxML ver. 7.2.7 (<https://cme.h-its.org/exelixis/php/countSource727.php>) [17] were used to align the whole cpDNA sequences and infer a phylogenetic tree (Supplementary Methods S2).

RESULTS

The full-length chloroplast genome of *M. argentea* is 155,871 bp, while *M. fatua* is 155,898 bp and they exhibit respective assembly coverages of 41.2× and 31.8× (Figure 3 and Supplementary Figure S1). The cpDNA genomes includes a large single copy (LSC) region (87,065 bp and 87,108 bp), a small single-copy (SSC) region (20,670 bp and 20,646 bp), and a pair of inverted repeat (IR) region (24,068 bp and 24,072 bp) in *M. argentea* and *M. fatua* respectively. For both species, we identified 126 genes corresponding to 83 protein-coding genes, 35 tRNA, and 8 rRNA, less genes in comparison with *M. fragrans* (146 genes) [19], and slightly higher compared with *M. yunnanensis* (120 genes) [12].

Phylogenetic relations according to cpDNA genomes between species were well resolved (bootstrap support > 85) (Figure 4). Four *Myristica* species formed a monophyletic clade, showing a similar topology as previously reported [18].

DISCUSSIONS AND CONCLUSION

The chloroplast genome structures of *M. argentea* and *M. fatua* have been analyzed and found to exhibit typical structural characteristics of angiosperm chloroplast genomes, including one LSC and SSC region, as well as two IR regions. We identified a different length of *atpF*, *clpP*, and *ndhD*, genes in both species, with a difference from 1 to 3 bp. The cpDNA-based phylogenetic tree support *M. argentea* as a sister species to *M. fragrans*, and the close taxonomic relationship between genera *Knema* and *Myristica*. This study will underpin the relationships between *Myristica* species and provide important details for evolutionary research and intra-specific diversity analysis.

ACKNOWLEDGEMENTS

We thank Gregory M. Plunkett, Sean Thackurdeen and Michael J. Balick from the New York Botanical Garden (NYBG) for having provided samples of *M. fatua*. The authors acknowledge the ISO 9001 certified IRD i-Trop HPC (member of the South Green Platform) at IRD montpellier for providing HPC resources that have contributed to the research results reported within this paper. URL: <https://bioinfo.ird.fr/>- <http://www.southgreen.fr>

AUTHORS' CONTRIBUTION

JK and JD conducted the field work; JK, NS, and JD conceptualized and designed research; JK and MC performed lab experiment; JK and NS analyzed data; JK wrote original draft of the manuscript; All authors read and approved the final manuscript.

DISCLOSURE STATEMENT

The authors report there are no competing interests to declare.

ETHICS PERMISSION

This research was done with the authorization of Indonesian National Research and Innovation Agency (BRIN) with research permit to JD (144/SIP/FRP/E5/Ditk.KI/V/2018) and collecting permit to JK (B-3382/IPH.1/KS.02.04/IX/2019 and B-75/IV/KS.01.04/3/2022).

DATA AVAILABILITY STATEMENT

The genome sequence data on this study are openly available in GenBank of NCBI (<https://www.ncbi.nlm.nih.gov/>) indicated in Bioproject (PRJNA904151),

BioSample (*M. argentea*: SAMN31831556; *M. fatua*: SAMN31831555) and SRA (*M. argentea*: SRR22428069; *M. fatua*: SRR22428070). The species are registered under the accession numbers OP866724 (*M. argentea*) and OP866725 (*M. fatua*).

FUNDING

This work was supported by the Indonesian Endowment Fund for Education (LPDP) [0003719/AFR/D/BUDI-2018]; Agropolis Fondation [ID 1502-503 and 1502-504] through the «*Investissements d'avenir*» program (Labex Agro: ANR-10-LABX-0001-01), under the frame of I-SITE MUSE (ANR-16-IDEX-0006). Field work for obtaining *M. argentea* specimen were authorized by Indonesian National Research and Innovation Agency (BRIN) with research permit to JD (144/SIP/FRP/E5/Ditk.KI/V/2018) and collecting permit to JK (B-3382/IPH.1/KS.02.04/IX/2019 and B-75/IV/KS.01.04/3/2022).

REFERENCES

- [1] R. Govaerts *et al*, “The World Checklist of Vascular Plants, a continuously updated resource for exploring global plant diversity,” *Sci. Data*, vol.8, no. 1, p. 215, Aug. 2021.
- [2] W.J.J.O. de Wilde, *Flora Malesiana*. *Pensoft Publishers*, p. 360 – 362. 2014.
- [3] T. Lim, “*Myristica fatua*,” in *Edible Medicinal and Non-Medicinal Plants*, 1st ed., Springer Netherlands, 2012, pp. 572–574.
- [4] C. Mariac *et al.*, “Diversity of wild and cultivated pearl millet accessions (*Pennisetum glaucum* [L.] R. Br.) in Niger assessed by microsatellite markers,” *Theor. Appl. Genet.*, vol. 114, no. 1, pp.49–58, Nov. 2006.
- [5] C. Mariac *et al.*, “Cost-effective enrichment hybridization capture of chloroplast genomes at deep multiplexing levels for population genetics and phylogeography studies,” *Mol. Ecol. Resour.*, vol.14, no.6, pp.1103–1113, Nov. 2014.

- [6] N. Scarcelli *et al.*, “Intra-individual polymorphism in chloroplasts from NGS data: where does it come from and how to handle it?” *Mol. Ecol. Resour.*, vol.16, no. 2, pp. 434–445, Mar. 2016.
- [7] J. Jin *et al.*, “GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes,” *Genome. Biol.*, vol. 21, no.1, p.241, Dec. 2020.
- [8] L. Shi *et al.*, “CPGAVAS2, an integrated plastome sequence annotator and analyzer,” *Nucleic. Acids. Res.*, vol. 47, no. W1, pp. W65–W73, Jul. 2019.
- [9] S. Zheng *et al.*, “CHLOROPlot: An Online Program for the Versatile Plotting of Organelle Genomes,” *Front. Genet.*, vol.11, Sep. 2020.
- [10] C. Cai *et al.*, “Comparative phylogenetic analyses of Chinese *Horsfieldia* (Myristicaceae) using complete chloroplast genome sequences,” *J. Syst. Evol.*, vol. 59, no. 3, pp. 504–514, May. 2021.
- [11] C. L. Mao *et al.*, “The complete chloroplast genome sequence of *Myristica yunnanensis* (Myristicaceae),” *Mitochondrial DNA Part B*, vol. 4, no. 1, pp. 1871–1872, Jan.2019.
- [12] T. Yang *et al.*, “The complete chloroplast genome sequence of *Knema furfuracea* (Myristicaceae),” *Mitochondrial DNA Part B*, vol.5, no. 1, pp. 325–326, Jan. 2020.
- [13] F. L. Zhang *et al.*, “The complete chloroplast genome sequence of *Horsfieldia amygdalina* (Myristicaceae),” *Mitochondrial DNA Part B*, vol.4, no. 2, pp. 3923–3924, Jul.2019.
- [14] J. Wang *et al.*, “Characterization of the complete chloroplast genome of *Michelia maudiae* (Magnoliaceae),” *Mitochondrial DNA Part B*, vol.4, no. 2, pp. 2146–2147, Jul. 2019.
- [15] J. Migliore *et al.*, “Pre-Pleistocene origin of phylogeographical breaks in African rain forest trees: new insights from *Greenwayodendron* (Annonaceae) phylogenomics,” *J. Biogeogr.*, p.jbi.13476, Dec. 2018.
- [16] K. Katoh *et al.*, “MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization,” *Brief. Bioinform.*, vol. 20, no. 4, pp.1160–1166, Jul. 2019.

[17] A. Stamatakis, “RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies,” *Bioinformatics*, vol. 30, no. 9, pp 1312–1313, May. 2014.

[18] S. de Oliveira *et al.*, “Chloroplast genome of the nutmeg tree: *Myristica fragrans* Houtt. (Myristicaceae),” *bioRxiv*, Jan. 2020.

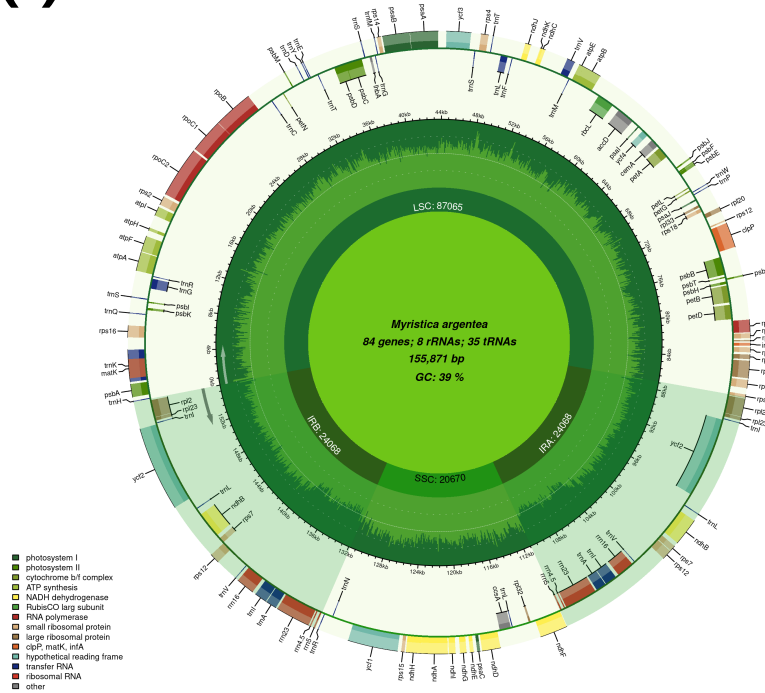
Figures

Figure 1. *Myristica argentea* grown in the forest of West Papua, Indonesia. (A) Tree trunk. (B) Mature fruit, with a description as follow: fruits solitary, \pm ellipsoid-fusiform, 6.5-7.5 by 3.5-4.5 cm, apex narrowed, base somewhat contracted into a pseudostalk, 3-8 mm long, early glabrescent, brown, usually set with conspicuous coarse pale pustules or lenticels, similar as reported by de Wilde (2018). (C) Longitudinal cut of the fruit showing red mace and nut (C). Pictures by Jaky Kusuma, Ambon, Indonesia, 2018. For the figure of *Myristica fatua*, picture is available in <https://sweetgum.nybg.org/images3/3009/671/03509312.jpg>



Figure 2. Annotation of the *M. argentea* (a) and *M. fatua* (b) chloroplast genomes inferred by CPGAVAS2. The outer circle represents the genes annotated, classified by color based on their function. The length of genome is presented in the inner circle. Large single copy (LSC), Single short copy (SSC) and inverted repeat regions (IRA and IRB) are marked.

(a)



(b)

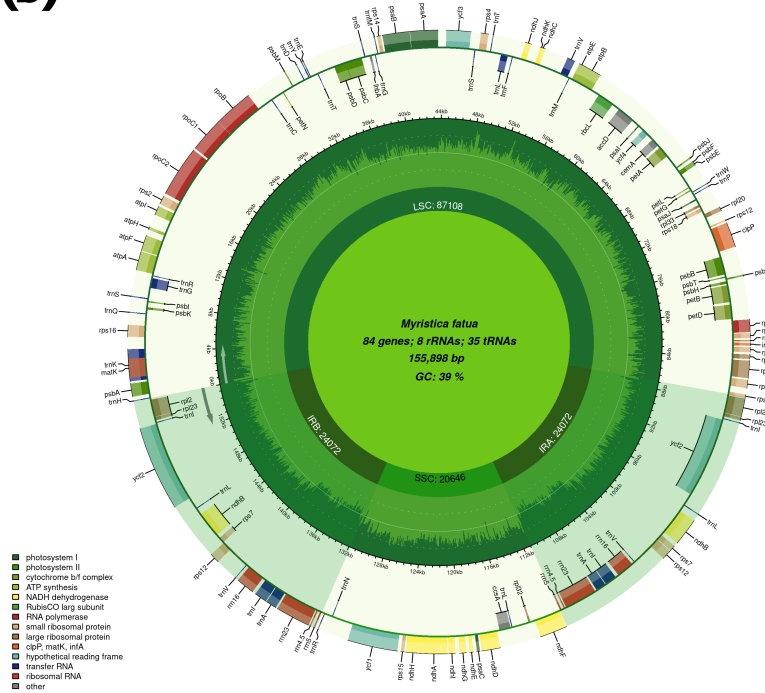
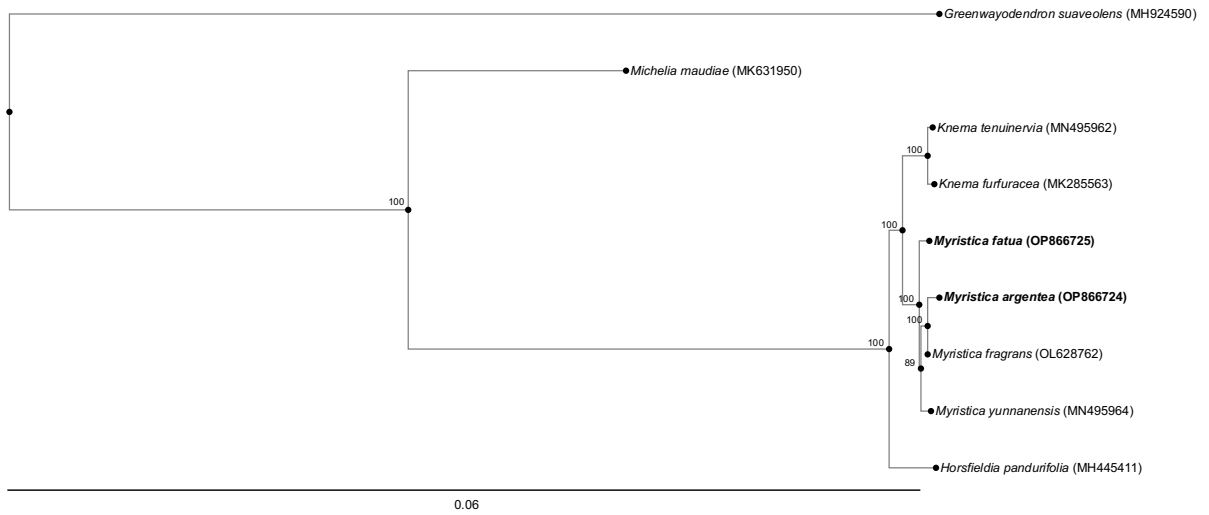


Figure 3. The best maximum-likelihood phylogenetic tree reconstructed using RAxML ver. 7.2.7 based on complete chloroplast genomes. The following sequences were used: *M. fragrans* (OL628762 [10]), *M. yunnanensis* (MN495964 [11]), *Knema tenuinervia* (MN495962 [10]), *Knema furfuracea* (MK285563 [12]), *Horsfieldia pandurifolia* (MH445411 [13]), as well as *Michelia maudiae* (MK631950 [14]) and *Greenwayodendron suaveolens* (MH924590 [15]) as outgroups. Numbers on branches indicate bootstrap values.



Supplementary Online Materials

Figure S1. Contig coverage depths of *Myristica argentea*.

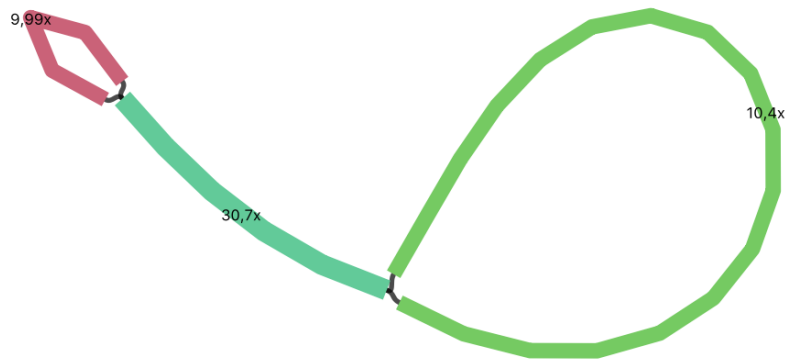
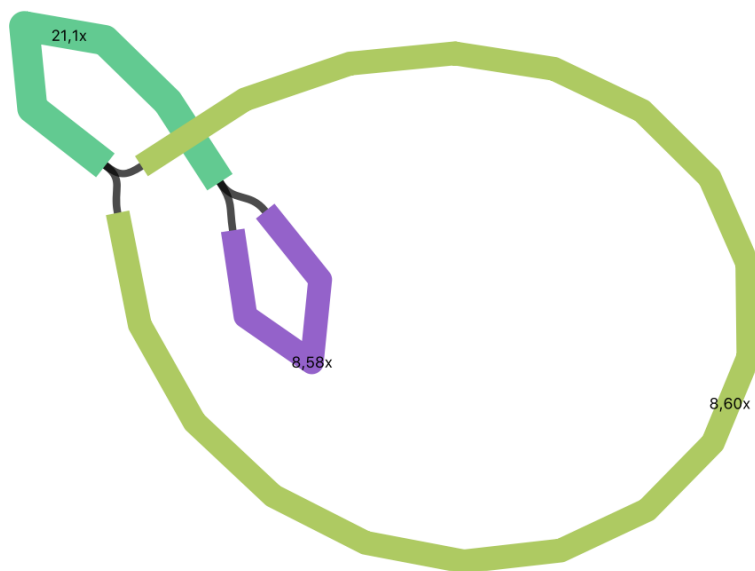


Figure S2. Contig coverage depths of *Myristica fatua*.



Supplementary methods

Methods S1. *De novo* assembly of whole chloroplast genome of *Myristica argentea* and *M. fatua* using GetOrganelle assembler [1]

We used `get_organelle_from_reads.py` script to assemble the chloroplast genome on both species, with this following parameter:

- -R 15 (maximum extension round)
- -k 21,45,65,85,105 (kmer setting)

We then used Bandage [2] to visually check if the sequence contains the target genome sequence.

Methods S2. Parameters used during aligning sequences of assembled *M. argentea* and *M. fatua* to closely related species of the family Myristicaceae and infer a phylogenetic tree.

- We first reordered all assemblies of focal species along with *M. fragrans* NC_060715 [11]; *M. yunnanensis* NC_060716 [12]; *Knema tenuinervia* NC_060714 [11]; *Knema furfuracea* MK285563 [13]; and *Horsfieldia amygdalina* MK285561 [14]), a species from Magnoliaceae (*Michelia maudiae* NC_047409 [15]), and from Annonaceae, *Greenwayodendron suaveolens* NC_042164 [16] using this parameter from MAFFT [3].

```
mafft --merge table --thread 3 --threadit 0 --reorder --anysymbol --auto input > output
```

- We then used all multifasta file of to align with these following parameters:

```
scoring matrix = 200, PAM k=2, gap open penalty = 1.53, offset value = 0.123
```

- Finally, we used RAxML to infer phylogenetic tree and using this following command to infer a phylogenetic tree:

```
raxmlHPC-SSE3 -f a -x 1237 -p 12345 -# 500 -m GTRGAMMA -s All_samples.fasta -n tree -o Greenwayodendron suaveolens (MH924590)
```

CHAPTER 3

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



Antique Map of the Moluccas by Plancius titled '*Insulae Molucae*', in the year of 1598. Engraved by Petrus Plancius, the great Dutch engraver and first Hydrographer to the *Dutch East India Company* (VOC). Compiled from the latest and highly secret Portuguese sources, the map provides the most detailed cartographic knowledge of the region at the end of the 16th century. From a European standpoint, this map stands out as one of the pioneering sources that revealed the geographical whereabouts of nutmeg, thus igniting the Age of Exploration. Source: barthelegallery.com

Chapter 3

Retracing the center of origin and evolutionary history of nutmeg

Myristica fragrans, an emblematic spice tree species

(Submitted to **Molecular Ecology**)

Jakty Kusuma^{1,2}, Nora Scarcelli¹, Miguel de Navascués^{3,4}, Marie Couderc¹, Pierre R. Gérard⁵, Jérôme Duminil¹

¹DIADE, Univ. Montpellier, IRD, CIRAD, Montpellier, France.

²Politeknik Negeri Lampung, Bandar Lampung, Indonesia.

³CBGP, INRAE, CIRAD, IRD, Institut Agro, Université de Montpellier, Montpellier, France.

⁴Human Evolution, Department of Organismal Biology, Uppsala University, Uppsala, Sweden.

⁵Génétique Quantitative et Evolution, INRAE, CNRS, Université Paris-Saclay, AgroParisTech, Gif-sur-Yvette, France

ABSTRACT

The identification of cultivated plant centers of origin is of growing interest since the pioneering work of Nikolai 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 showed patterns of a recent bottleneck, whereas North Moluccas 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.

Key words: food tree species, nutmeg, phylogeography, propagation history, spice trade.

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 (de Candolle, 1855, 1882). He synthesized data from different disciplines (botany, plant biogeography, archaeology, paleontology, history, linguistics, ethnology) to propose hypotheses on the geographic origin of 247 cultivated species. This work was later extended by Nikolai Vavilov, who proposed main centers of domestication for cultivated plants (Vavilov, 1926). Vavilov's work has since been revisited and updated by others (Burkill, 1953; Darlington, 1973; Hanson, 1952; Harlan, 1986; Harris, 1990; Sauer, 1952; Schery, 1952). Indeed, the centers of plant domestication (*sensu* Vavilov 1926) were confirmed by a number of archaeological findings (Smith, 1968, 1969). Furthermore, additional progress has been made on this topic through the use of genomic data to study the origin of cultivated plants (Schreiber et al., 2018). However, despite this 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 outside of its original distribution range often results in genetic bottlenecks (Baute et al., 2015; Price & Hadfield, 2014; Zhang et al., 2021). 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 (Hirota et al., 2021; Li et al., 2021). 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 (Burgarella et al., 2019; Cornille et al., 2012; Janzen et al., 2019).

Genetic studies based on extensive sampling across a species' distribution can contribute to tracing back the biogeographical history of that species (Avisé, 2009; Petit & Grivet, 2002; Ray & Excoffier, 2009). For non-model species with few genetic 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 (Meegahakumbura et al., 2018; Wang et al., 2019).

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 (Dayuan, 1339). 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 (Fish, 2012; Hostetler, 2005; Liao, 2015). The Portuguese were the first European explorers arriving in the region in search of the Spice Islands in the XVIth 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 (Rumpf (1741) English translation by Beekman (2011) 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 (Ellen, 2003). 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 (Keay, 2006; Wright, 1958). 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 the arrival of the Europeans (Lape et al., 2018). This study was conducted 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 2). 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.* (2006). We checked DNA quality and quantity using a NanoDrop™ One spectrophotometer (Thermo Scientific™). 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.* (2020). As one of the primer pair (*Myr42* locus) allowed for the amplification of two different genomic regions, corresponding information was split into two, namely *Myr42a* and *Myr42b* (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 μ L of PCR product, 12 μ L of Hi-Di Formamide (Applied Biosystems) and 0.3 μ L 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 (Kearse *et al.*, 2012) 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.* (2014). All sequenced individuals were analyzed as described in Scarcelli *et al.* (2016). In short, data were demultiplexed and cleaned using the scripts DEMULTADAPT (<https://github.com/Maillol/demultadapt>) and CUTADAPT v.1.2.1 (Martin, 2011). 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 (Li & Durbin, 2009) using the *M. fragrans* cpDNA genome (de Oliveira *et al.*, 2020) as

reference. Unmapped reads were discarded and mapped reads were sorted using SAMTOOLS v.1.9 (Danecek et al., 2021). Then VARSCAN v.2.3.7 (Koboldt et al., 2012) was used to call Single Nucleotide Polymorphisms (SNPs). All SNPs were filtered using VCFtools (Danecek et al., 2011), 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 (Danecek et al., 2011). 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* at nSSR loci*

The distribution of genetic diversity in *M. fragrans* ($N = 393$ individuals) was characterized at the population level ($N_{\text{POP}} = 5$) by estimating the following indices using SPAGeDi v.1.5 (Hardy & Vekemans, 2002): 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.2.3.4 (Pritchard et al., 2000) 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 steps in a single burn-in periods and 100,000 MCMC replicates. The value of K that best explained the

data was estimated using STRUCTURE HARVESTER (Earl & VonHoldt, 2012). We then summarized STRUCTURE results using CLUMPP (Jakobsson & Rosenberg, 2007) and projected the barplot using CLUMPAK (Kopelman et al., 2015).

*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 ver.6.12.03 (Librado & Rozas, 2009). 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 (Excoffier & Lischer, 2010) assessing significant differences with 1,000 permutations.

Haplotype network reconstruction

We assigned cpDNA SNPs to different haplotypes using the *pegas* R package ver.1.1 (Paradis, 2010). We constructed the haplotype network using the maximum parsimonious tree method (Kannan & Wheeler, 2012) available in PopART v.1.7 (Leigh & Bryant, 2015), with a 95% statistical parsimony criterion. To test for phylogeographic signal, we calculated the N_{ST} , and the G_{ST} (Pons & Petit, 1996), using Permut v.1.2.1 applying 1,000 permutation tests (Petit, 2012).

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 (Beaumont et al., 2002; Csilléry et al., 2010). Using the nSSR dataset, we characterized the demographic evolution through time for each population using the program DIYABCskylineplot (Navascués et al., 2017). Coalescent simulations and the computation of prior parameters and summary statistics were done using command line in DIYABC v.2.0 (Cornuet et al., 2014). 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 (Leblois et al., 2014) as implemented in MIGRAINE (Rousset et al., 2018). 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

We inferred the ancestry and divergence time of the genetic clusters defined by STRUCTURE (see Results section), using DIYABCRF v.1.1.1 (Collin et al., 2021). Populations assigned to a given genetic cluster according to STRUCTURE results were pooled into 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 (1991) and Lape (2000) 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 3), 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 (*Hap3*, ~60% of the individuals), that was present in all populations throughout the Moluccas

archipelago (Figure 4; Supplementary Figure S4). Two other common haplotypes (*Hap4* and *Hap17*) 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 *Hap3* 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 ($\Pi = 0.122$) than in the Northern Moluccas cluster ($\Pi = 0.280$). The AMOVA shows that the genetic variance is mainly explained by the cluster ($\Phi = 41.50\%$ 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 size 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 5). MIGRAINE revealed a population expansion in Ternate, with $\theta_{\mu} > \theta_{\mu}^{\text{anc}}$. 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 $\theta_{\text{ratio}} = \theta_0/\theta_A$ (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 S6, S7 and S8 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 6; 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 *Hap3*. The *Hap3* 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 *Hap3*. 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 genus *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 (Kitamura, 2011). 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 (Dick et al., 2007; Gandour et al., 2008; Hanaoka et al., 2014; Liu et al., 2021). 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 (Pigafetta, 1522; Rumpf, 1741b; Sonnerat, 1776). It is thus striking that at that time, most Europeans were focusing on getting material from Banda islands to circumvent the Dutch monopoly and propagate the species in their tropical colonies (Hanna, 1990). Why was the initial focus concentrated on getting material from the strictly controlled islands of Banda and Ambon, if plantation material was more easily accessible in the North Moluccas? We do not have a definitive answer to this question.

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 (Donkin, 2003), and South Moluccas for the cultivation of nutmegs (Brixius, 2018; Dhont, 2022). 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 has remained unimpacted in contrast to the South Moluccas stock that has 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 (Bentley & Trimmen, 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 are 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 (Ibrahim et al., 2020; van Ruth et al., 2019). 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 (de Villemereuil et al., 2016). 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

The authors would like to thank all farmers and land owners in Ambon, Banda, Ternate, Tidore, and Bacan for providing information on nutmeg history as well as leaves materials, as well as to our guide and Dina Tri Wahyuni for providing assistance during the field work. We thank Pongky Van den Broeke in Banda for all information on nutmeg history. Francis J. Nge for providing the English corrections. This study was authorized by Indonesian National Research and Innovation Agency (BRIN) with research permit to JD

(144/SIP/FRP/E5/Ditk.KI/V/2018) and collecting permit to JK (B-3382/IPH.1/KS.02.04/IX/2019 and B-75/IV/KS.01.04/3/2022). We also greatly appreciate the support from the Indonesia Endowment Fund for Education (LPDP) to JK with grant No. 0003719/AFR/D/BUDI-2018. This project was also supported by the Agropolis Fondation under the reference ID 1502-503 and 1502-504 through the « *Investissements d'avenir* » programme (Labex Agro: ANR-10-LABX-0001-01), under the frame of I-SITE MUSE (ANR-16-IDEX-0006). The authors acknowledge the ISO 9001 certified IRD i-Trop HPC (South Green Platform) at IRD Montpellier for providing HPC resources that have contributed to the research results reported within this paper.

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>
- 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>
- Beekman, E. M. (2011). *The Ambonese Herbal by Rumphius* (1st ed., Vol. 2). Yale University Press.
- Bentley, R., & 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. 1). J. & A. Churchill. <https://www.biodiversitylibrary.org/item/78161>
- 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.

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., Tenailon, 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>

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 géographique des plantes de l'époque 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 Wilde, W. J. J. O. (2014). *Flora Malesiana*. Pensoft Publishers. <https://doi.org/10.3897/ab.e1141>
- 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>
- 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. <http://www.jstor.org/stable/j.ctvvmxq8.6>
- 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>
- 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>
- 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/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>
- 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.
- 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>
- 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.
- 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). Clumpak : 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., Néron, 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). popart : 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., Bezançon, 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., Kougbéadjou, 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., Möller, 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>
- Navascués, M., Leblois, R., & Burgarella, C. (2017). Demographic inference through approximate-Bayesian-computation skyline plots. *PeerJ*, *5*, e3530. <https://doi.org/10.7717/peerj.3530>
- 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. <https://hal.inrae.fr/hal-02810373/document>
- Petit, R. J., & Grivet, D. (2002). Optimal randomization strategies when testing the existence of a phylogeographic structure. *Genetics*, *161*(1), 469–471. <https://doi.org/10.1093/genetics/161.1.469>
- Pigafetta, V. (1522). *Premier voyage autour du monde sur l'escadre de Magellan*. Librairie Ch. DeLagrange.
- Pons, O., & Petit, R. J. (1996). Measuring 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* and *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. <http://www.genetics.org/content/155/2/945.abstract>
- 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>
- 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 Française de Statistique*, *159*(3), 142–166. <https://hal.umontpellier.fr/hal-01998088>
- Rumpf, G. E. (1741a). *Herbarium amboinense* (1st ed., Vol. 2). Apud Franciscum Changuion, Joannem Catuffe, Hermannum Uytwerf. <https://doi.org/10.5962/bhl.title.569>
- Rumpf, G. E. (1741b). *Herbarium amboinense* (1st ed., Vol. 3). Apud Franciscum Changuion, Joannem Catuffe, Hermannum Uytwerf. <https://doi.org/10.5962/bhl.title.569>
- Sauer, C. O. (1952). *Agricultural origins and dispersals* (1st editio). American Geographical Society.
- 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 NGS 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 edition). 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. <http://www.jstor.org/stable/4252962>
- Smith, C. E. (1969). From Vavilov to the Present: A Review. *Economic Botany*, 23(1), 2–19. <http://www.jstor.org/stable/4253008>
- Sonnerat, P. (1776). *Voyage à la Nouvelle Guinée*. 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. <http://www.jstor.org/stable/24367594>
- 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. <http://www.jstor.org/stable/41503143>
- 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.

Tables

Table 1. Genetic diversity in *M. fragrans* populations at fifteen polymorphic microsatellite loci.

Populations	<i>N</i>	<i>N_P</i>	<i>N_A</i>	<i>N_M</i>	<i>AR_(k=46)</i>	<i>H_O</i>	<i>H_E</i>	<i>F_{IS}</i>
Ambon	161	7	75	5.00	3.85	0.512	0.547	0.068**
Banda	72	4	68	4.53	3.84	0.587	0.606	0.038
Ternate	55	6	154	10.26	8.63	0.711	0.770	0.086**
Tidore	30	3	154	10.26	9.74	0.764	0.812	0.076
Bacan	75	9	176	11.73	8.98	0.668	0.776	0.147**

Number of individuals analyzed (*N*), number of subpopulations in the population (*N_P*), number of alleles (*N_A*), mean number of allele per populations (*N_M*), rarefied allelic richness for k=46 gene copies (*AR_(k=46)*), observed heterozygosity (*H_O*), expected heterozygosity (*H_E*), inbreeding coefficient (*F_{IS}*) calculated overall loci on each population, where * indicates values significantly different from zero with $P < 0.01$.

Table 2. Genetic differentiation between five populations of *Myristica fragrans* at nSSR loci.

F_{ST}	Ambon	Banda	Ternate	Tidore	Bacan
Ambon	-				
Banda	0.028	-			
Ternate	0.106	0.064	-		
Tidore	0.182	0.126	0.019	-	
Bacan	0.106	0.065	0.012	0.023	-

Significant F_{ST} values are shown in bold ($P < 0.05$).

Table 3. Genetic diversity in *M. fragrans* populations at cpDNA loci.

Populations	<i>N</i>	<i>H</i>	Π	S	P
Ambon	20	8	0.116	56	14
Banda	15	7	0.117	50	10
Ternate	20	5	0.312	90	45
Tidore	8	6	0.188	59	0
Bacan	19	10	0.277	96	65
Cluster level	<i>N</i>	<i>H</i>	Π	S	P
South Moluccas	35	15	0.122	82	23
North Moluccas	47	21	0.280	113	77

Number of individuals analyzed (*N*), number of haplotype (*H*), nucleotide diversity (Π), number of segregating sites (S), number of parsimony-informative sites (P).

Table 4. Inference of divergence time (t) and time of the bottleneck event in the South Moluccas (tc) from DIYABC-RF (corresponding to scenario 1 in Figure S1), including the 0.025 and 0.975 quantiles (Q) of the posterior distribution. Values described below are the number of generations.

Parameter	Marker	Median	Q0.005	0.975
t	SSR	1713	223	10170
	cpDNA	4948	750	7625
	SSR and cpDNA	7433	1084	15241
tc	SSR	261	37	793
	cpDNA	332	25	915
	SSR and cpDNA	449	51	962

Figures

Figure 1. The Moluccas Archipelago, 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.

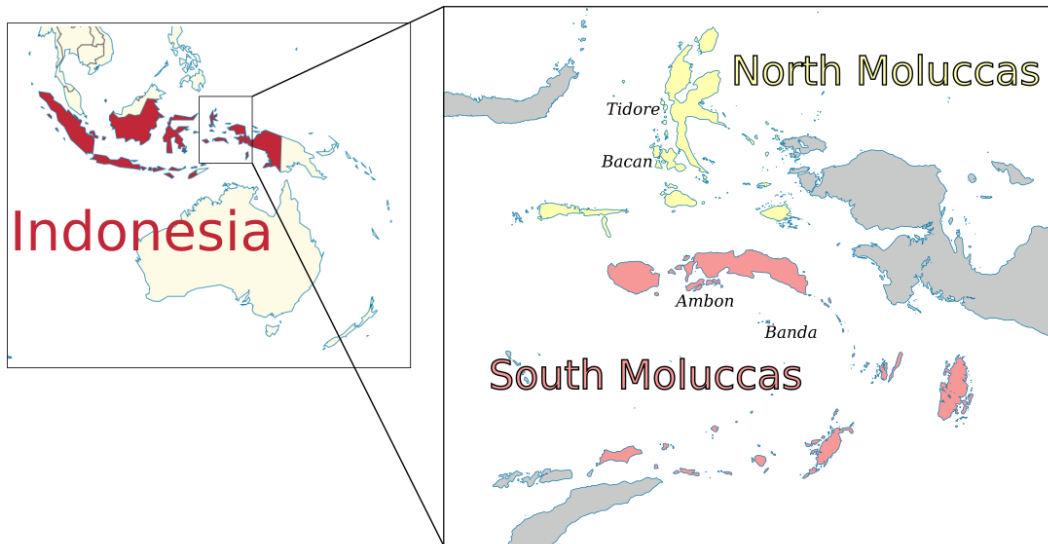


Figure 2. Moluccas archipelago in Indonesia (A), the inner caption is centered on Indonesia. (B) Ternate and Tidore island, (C) Bacan, (D) Ambon, and (E) Banda archipelago. Each dot corresponds to a sampled individual.

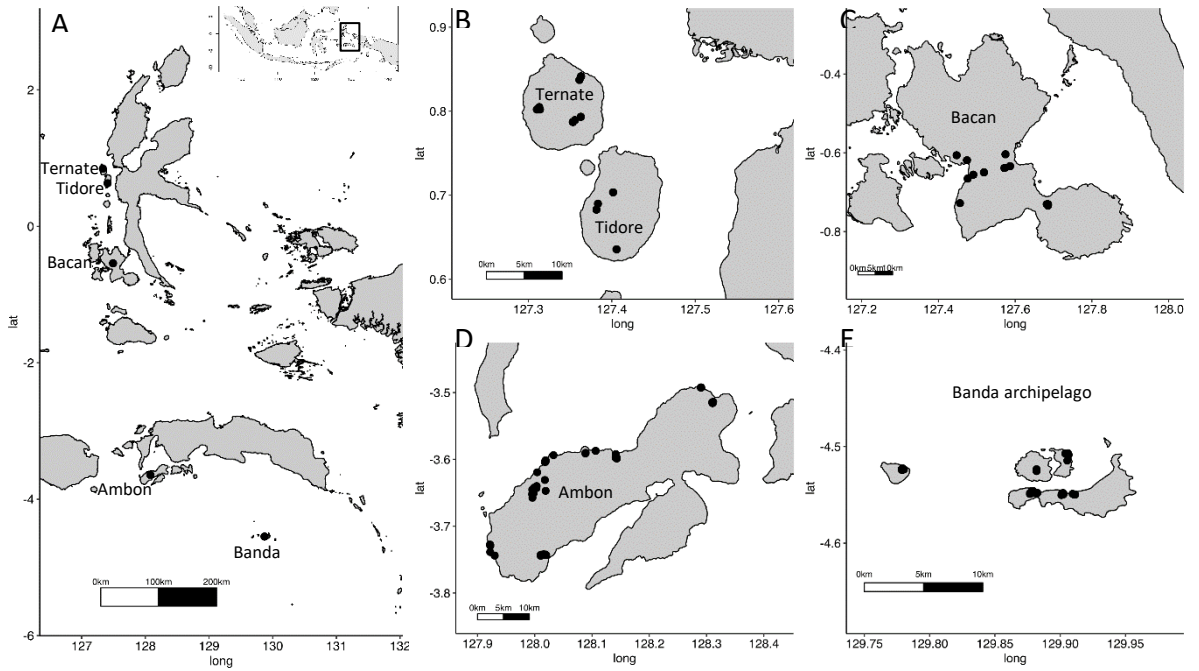


Figure 3. Intra-specific genetic structure in *M. fragrans*, obtained by a Bayesian clustering analysis on fifteen nSSR loci. Barplot of individual admixture proportions for each individual inferred by STRUCTURE for (A) $K = 2$ and (B) for $K = 3$, and geographical distribution of individuals inferred with STRUCTURE for (C) $K = 2$ and for (D) $K = 3$.

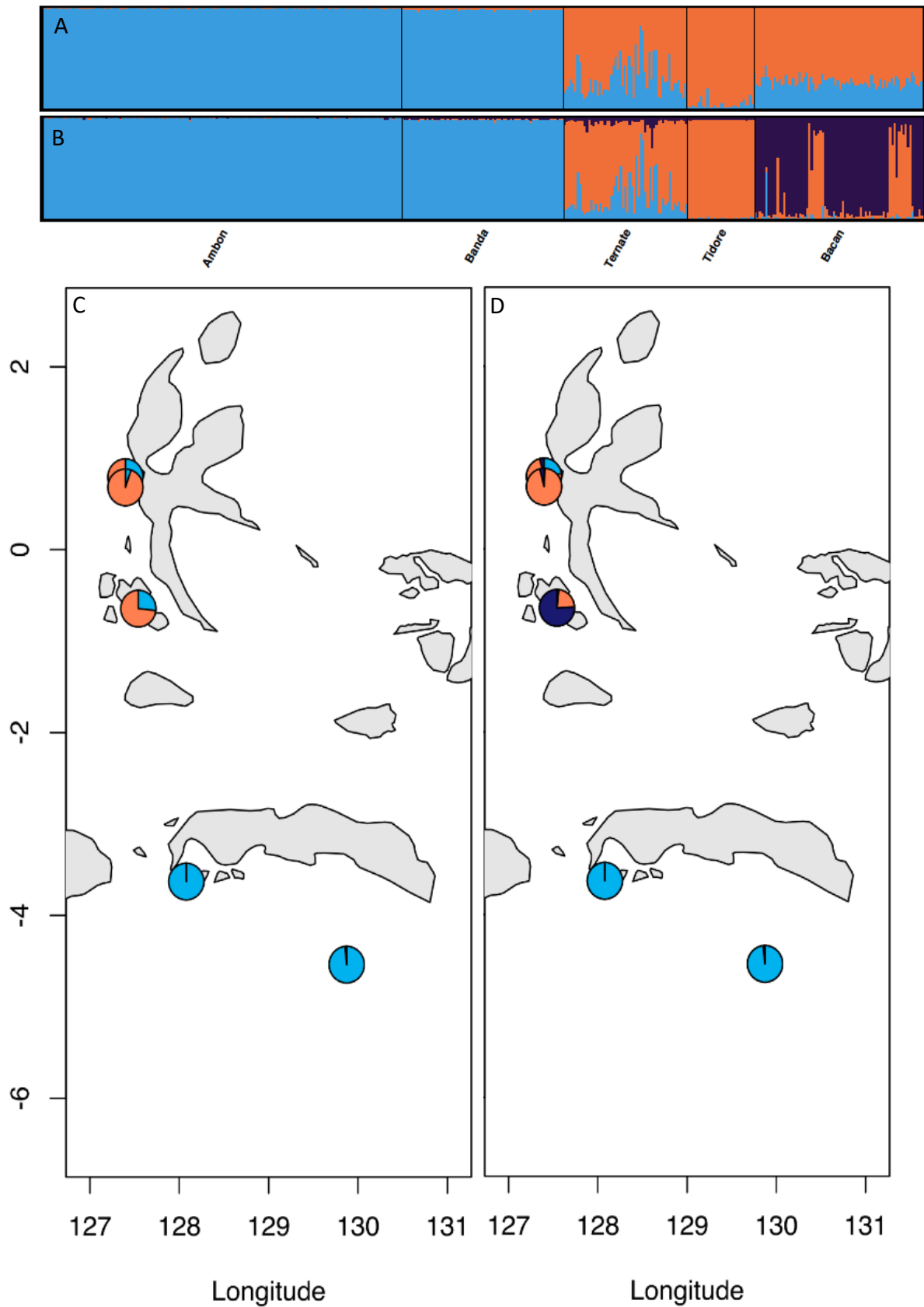


Figure 4. Maximum parsimonious network based on cpDNA haplotypes. The circles represent the different haplotypes. Blue haplotypes correspond to the Southern Moluccas genetic cluster as obtained from the Bayesian clustering analysis, and orange haplotypes correspond to the Northern Moluccas one. Each bar on the branch represents a mutation event. Hap5, Hap6, and Hap7 are *M. argentea* haplotypes represented in purple color.

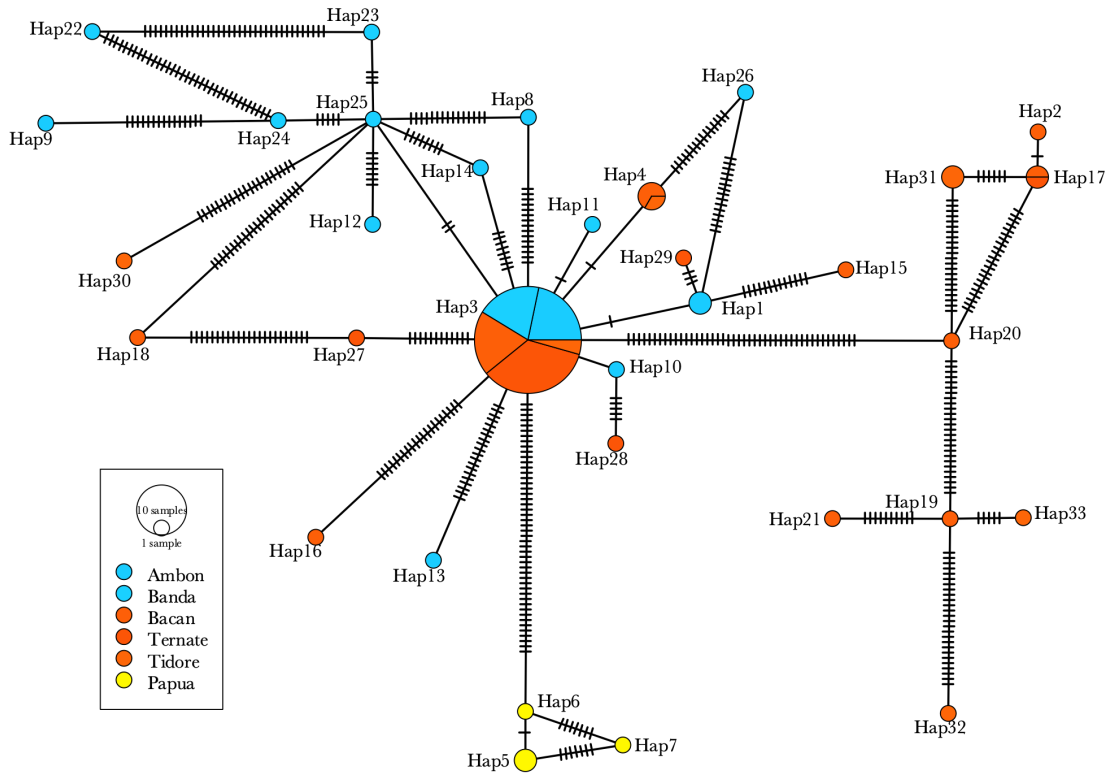


Figure 5. Demographic history of *M. fragrans* in the five islands under study using DIYABCskylineplots. The figures show changes in population size through time for Ambon, Banda, Ternate, Tidore, and Bacan. On each island's graph plot, median is shown in colored plain lines. Dashed lines on each box indicate the 95% highest posterior density (HPD).

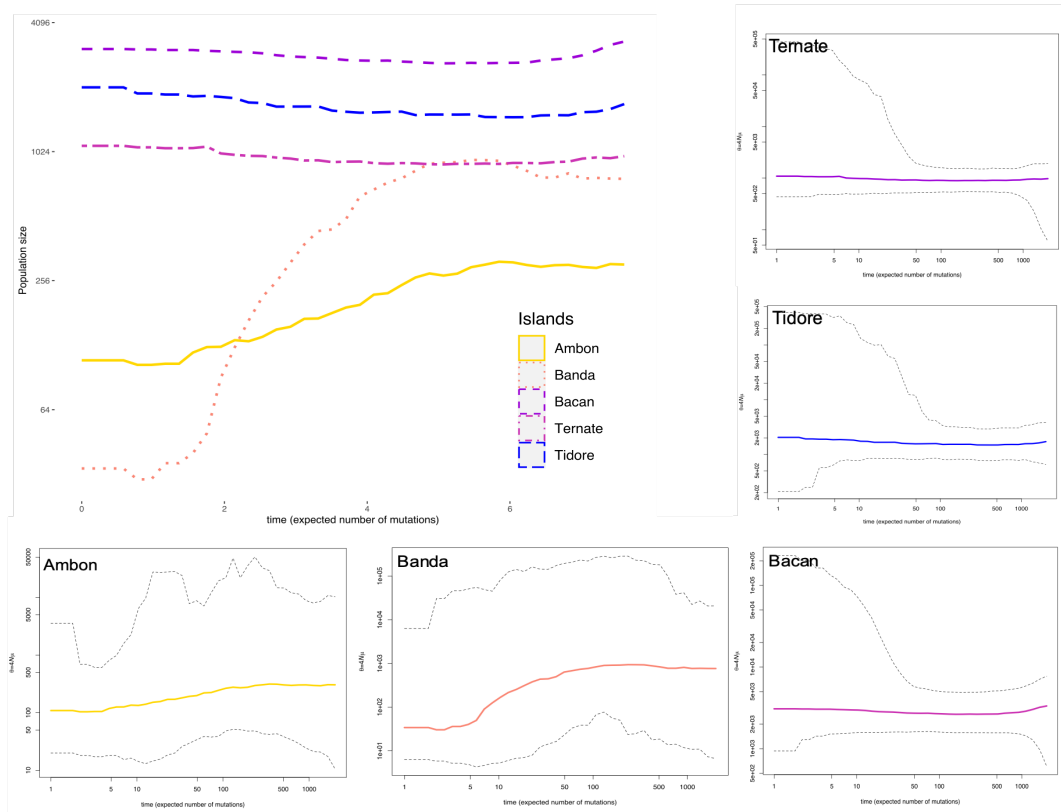
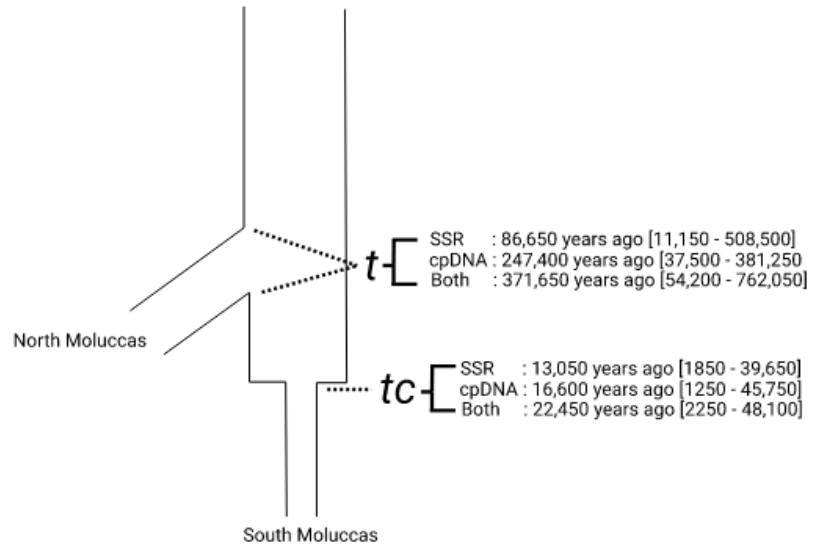


Figure 6. Best supported scenario according to the Random Forest analysis. (A) The most voted model estimates according to the random forest analysis, using nSSR and a combination of nSSR and cpDNA, shows that *M. fragrans* originated from South Moluccas (in blue line), then colonization occurred in North Moluccas (in green line) at t , followed by a population size reduction in South Moluccas (in orange line) at tc .



Supplementary Online Materials

Figure S1. Graphic representation of inferred scenarii of population history of *M. fragrans* in the Moluccas, Indonesia. *Pop_1* represents Southern populations and *Pop_2* represents Northern population. (A) Scenario 1 shows that the Southern populations were ancestral and the Northern populations diverged at time t , before a reduction in population size in the South at time tc ; and (B) scenario 2 illustrates that the Northern populations are ancestral and the South was then colonized at time t with a reduction in genetic diversity by a founder event.

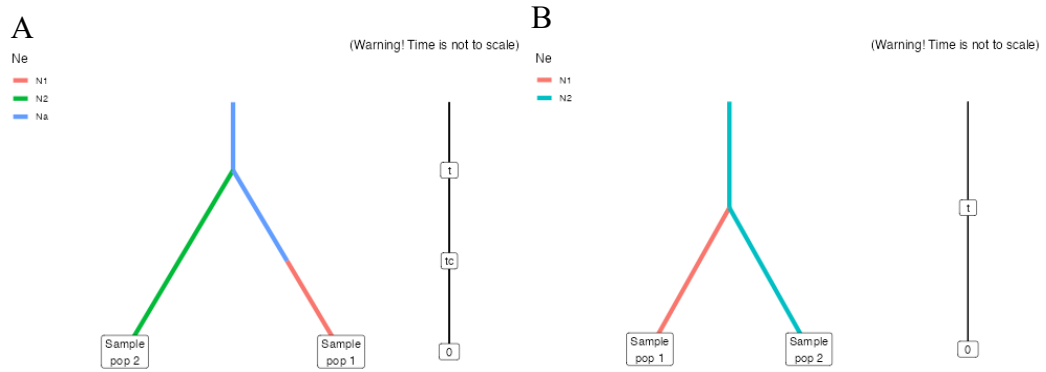


Figure S2. Plot of delta K values from the Structure analyses. The Evanno method (Evanno *et al.*, 2005) revealed that the most likely number of populations (K) is two.

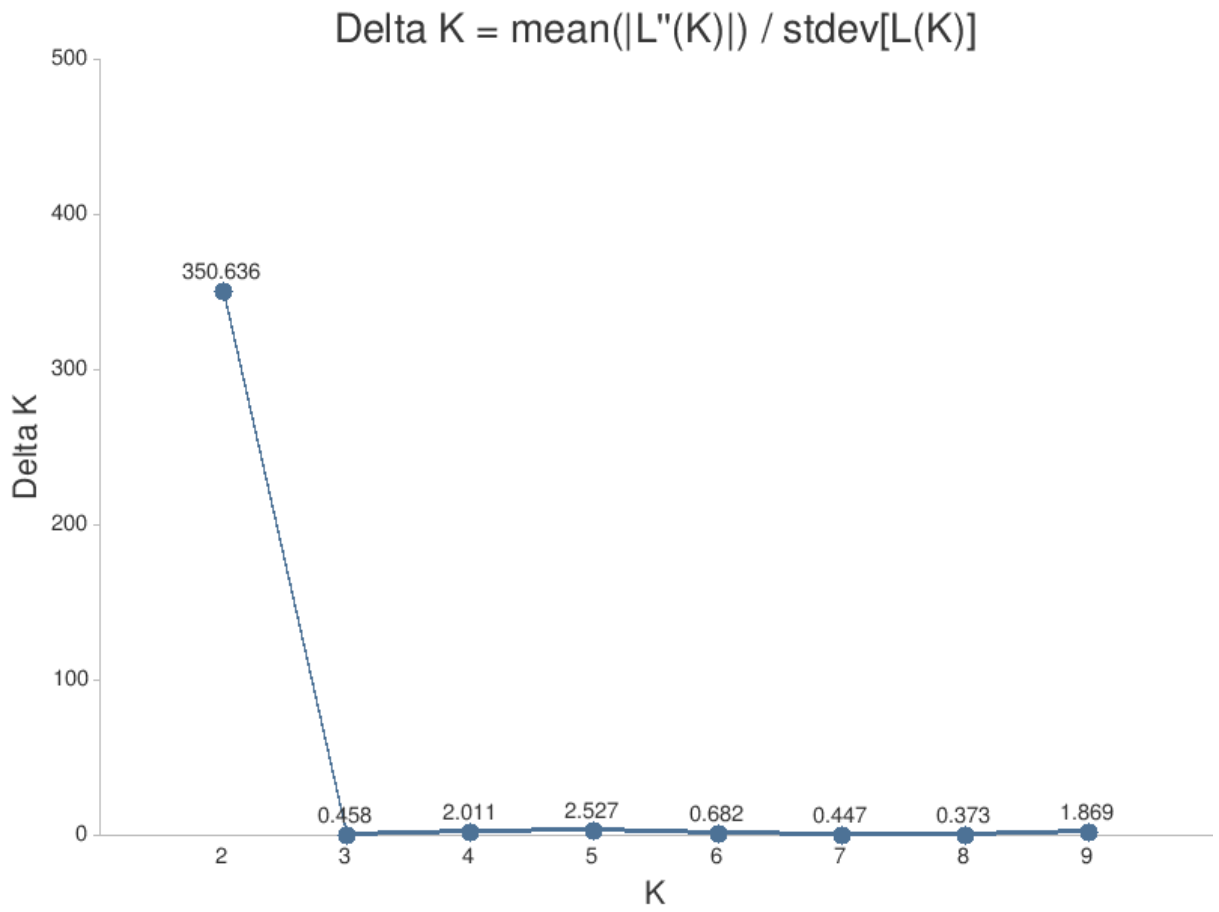


Figure S3. Confirmation of population structure subdivision within the North Moluccas group in Bacan island population. Assignment of $K = 2$ (A) in the North Moluccas group and (B) in Bacan island with indication of subpopulation names. © Spatial distribution of individual assignment in Bacan island inferred by TESS3 (Caye *et al.*, 2016). (D) Geographical distribution of individuals inferred by STRUCTURE with $K = 2$.

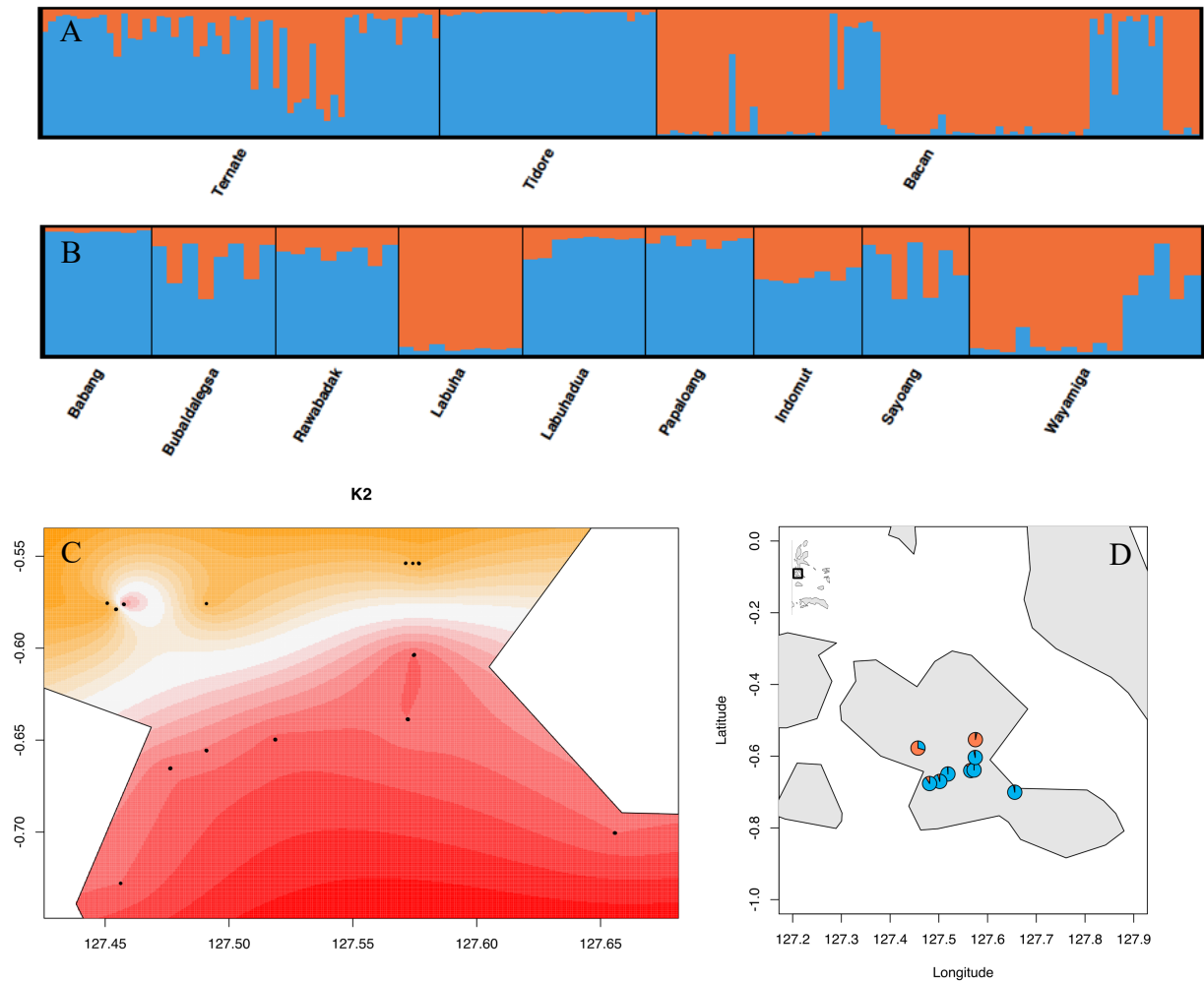


Figure S4. Haplotype distribution based on cpDNA sequences in five different islands in the Moluccas (A). Common haplotype (*Hap3*) is represented by a white color, and is found in all five islands. Two haplotypes (*Hap4* and *Hap17*, orange color) are present only in North Moluccas. Each island in the North Moluccas group (Ternate, Tidore, and Bacan) has three, four, and seven unique haplotypes, respectively (colors yellow, green and red). No shared haplotype exists in the South Moluccas group (except the common *Hap3*), but each island has seven unique haplotypes for Ambon (dark blue), and six unique haplotypes for Banda (light blue). Haplotypes of *M. argentea* are represented in purple color.

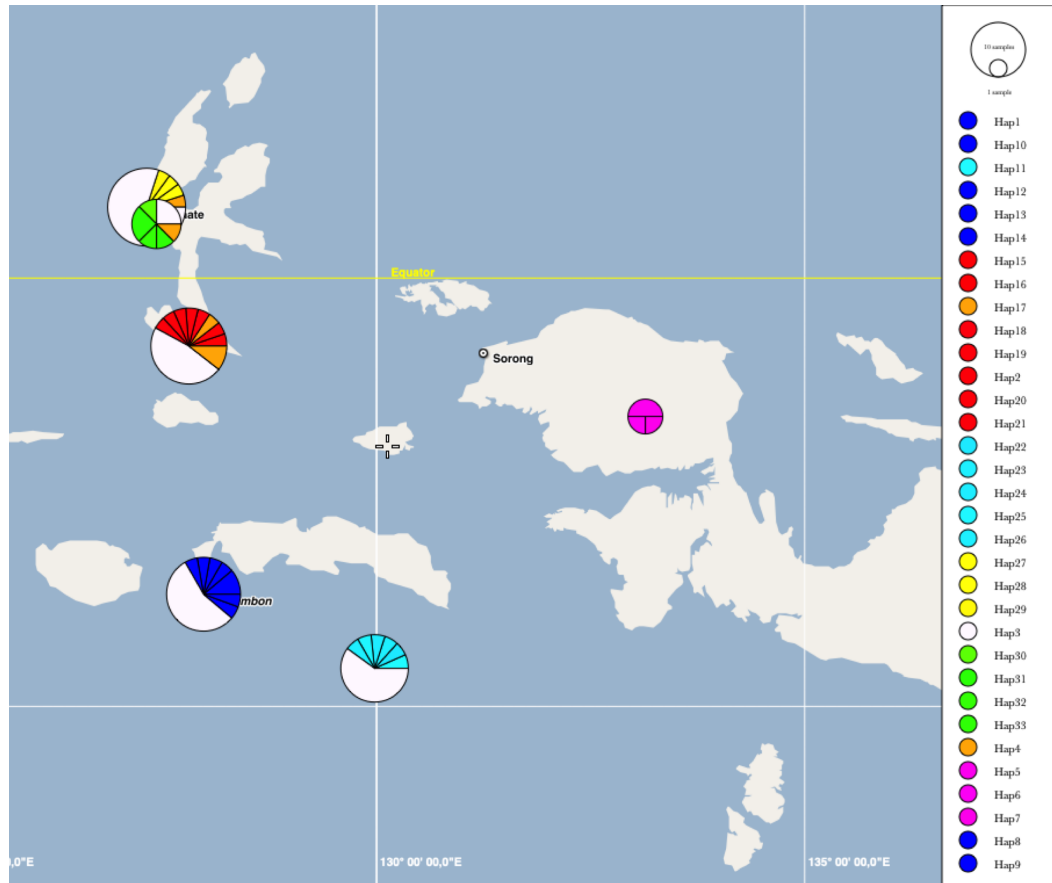


Figure S6. Principal Component Analysis (PCA) of the prior-scenario combinations and the observed dataset using nSSR. Scenario 1 and 2 priors are indicated with green and red dots respectively, and observed dataset is indicated by a yellow dot.

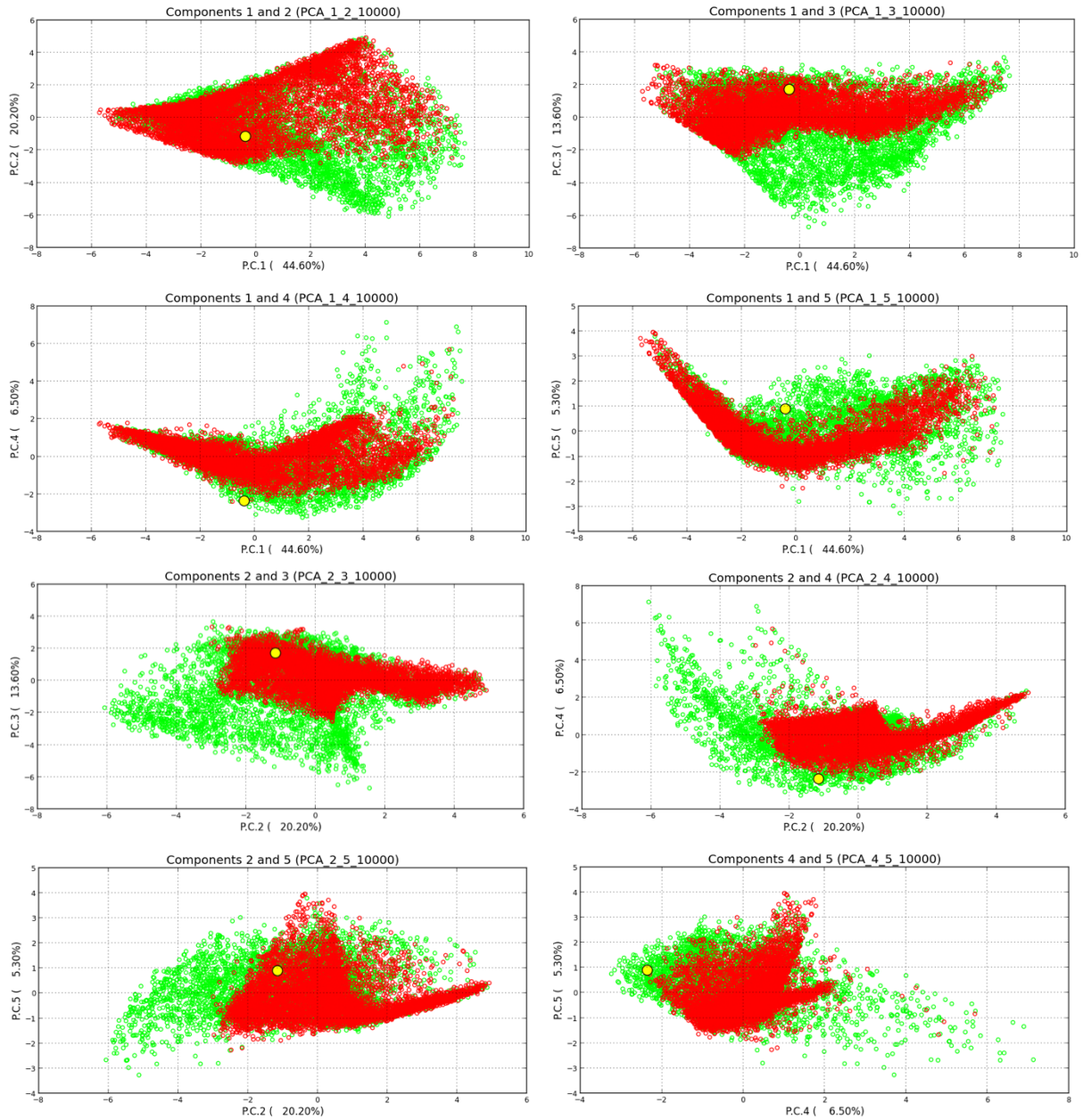


Figure S7. Principal Component Analysis (PCA) of the prior-scenario combinations and the observed dataset using cpDNA. Scenario 1 and 2 priors are indicated with green and red dots respectively, and observed dataset is indicated by a yellow dot.

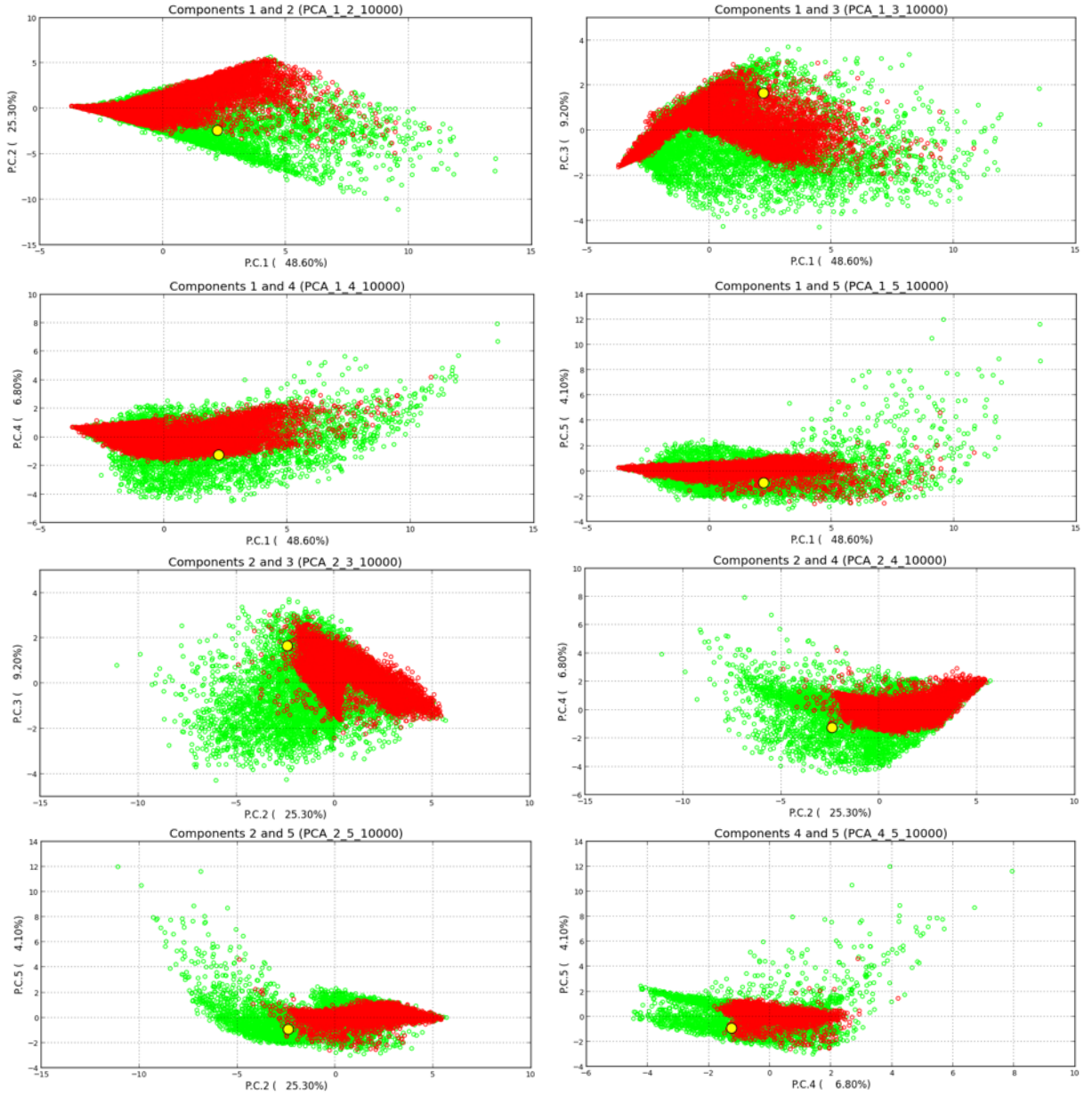


Figure S8. Principal Component Analysis (PCA) of the prior-scenario combinations and the observed dataset using a combination of nSSR and cpDNA. Scenario 1 and 2 priors are indicated with green and red dots respectively, and observed dataset is indicated by a yellow dot.

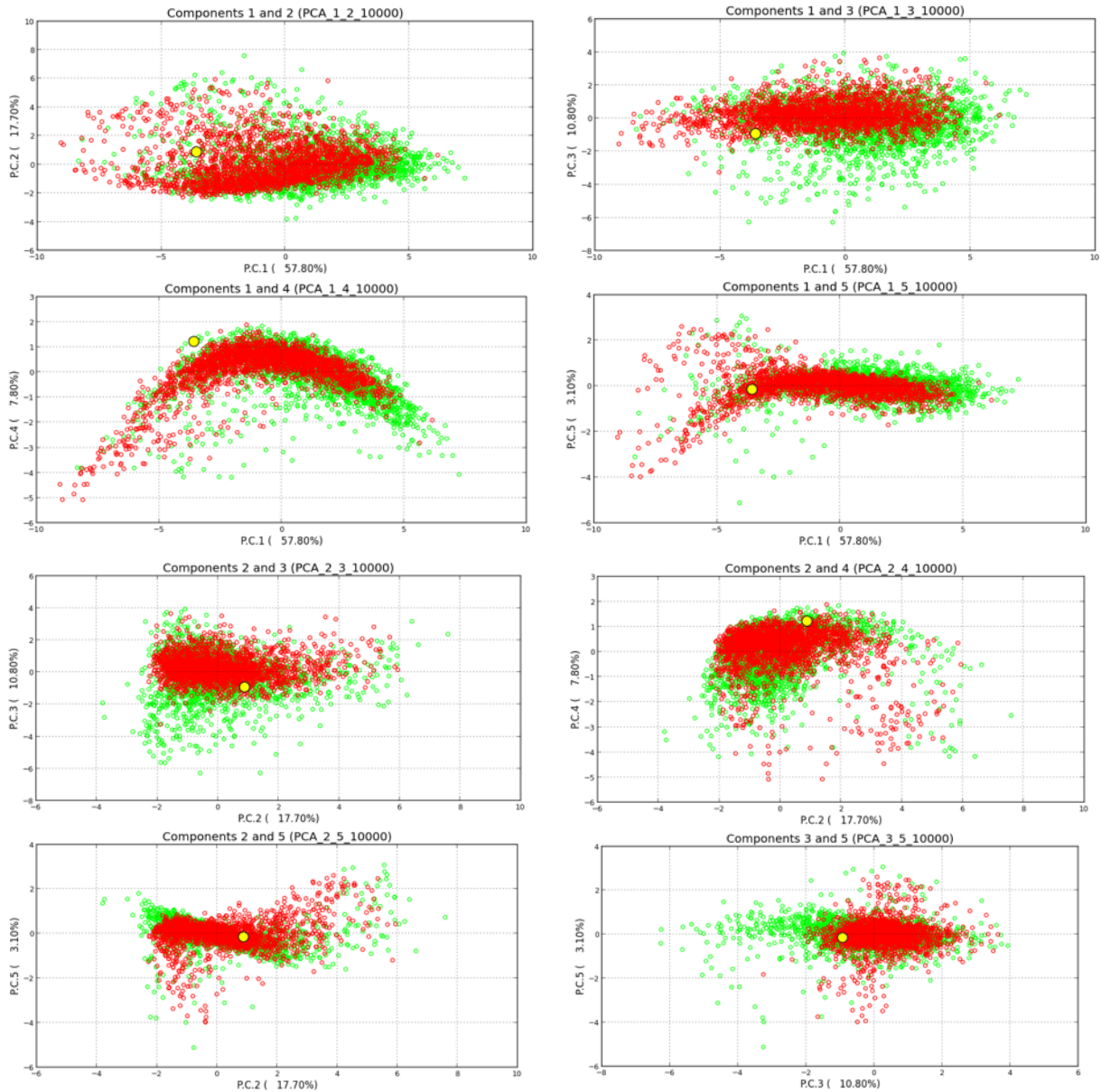


Table S1. Description of the historical parameter priors on the demographic inference of *M. fragrans* in ABC analysis.

Parameter type	Parameter name	Prior distributions
Population size [†]	Na	Log_U [10 - 50,000,000)
	N1	Log_U [10 - 5000]
	N2	Log_U [10 - 10,000,000]
Divergence time [‡]	t	U [10 - 16,000]
	tc	U [10 - 1000]
Conditions	t > tc	

[†]In number of reproductively effective diploid individuals. [‡]In number of generations assuming a generation time of X years. Distributions were considered either uniform (U) or Log-uniform (Log_U). Reference tables were constructed for 100,000 data sets per scenario.

Table S2. The genetic structure within island populations (F_{ST} among subpopulations) (Weir & Goudet, 2017).

No	Island	Num. of subpopulations	Num. of indiv	F_{ST}
1	Ambon	7	161	0.033
2	Banda	4	72	0.027
3	Ternate	6	55	0.015
4	Tidore	3	30	0.041
5	Bacan	9	75	0.009

Table S3. Composition of Bacan island population with nine sub-populations.

Population	<i>N</i>	Mean latitude	Mean longitude
Babang	7	-0.633868	127.586764
Bubaldalegsa	8	-0.603794	127.574543
Rawabadak	8	-0.618809	127.474496
Labuha	8	-0.649834	127.518740
Labuhadua	8	-0.665360	127.476460
Papaloang	7	-0.727851	127.456238
Indomut	7	-0.655589	127.490832
Sayoang	7	-0.633702	127.586541
Wayamiga	15	-0.638741	127.572245

Table S4. Variance components from AMOVA analyses. Sigma represents the variance for each hierarchical level and to the right is the percent of the total variation.

Sum of variations	Sigma	% variations	Phi	<i>p</i> value
Among groups	245.38	41.50	0.47	0.001
Among populations	38.35	6.48	0.11	0.045
Within populations	307.42	52	0.41	0.021
Total	591.15			

Table S5. Results of the ABC model selection using DIYABC-RF.

Marker	Selected scenario	Prior error (%)	Posterior probability	Bayes factor (K)*
nSSR	1	14.07	0.737	2.802
cpDNA	1	13.36	0.881	7.403
Combination nSSR and cpDNA	1	14.90	0.799	3.975

* A value of $K > 1$ means that Bayes factor is strongly supported according to Bayes factor in Jeffrey's scale (Jeffreys, 1998). $K = \text{post.prob} / (1 - \text{post.prob})$.

Table S6. Description of demographic parameters in *M. fragrans* populations as inferred by MIGRAINE.

Population s	p_{GSM}^{\dagger}	θ_{μ}	$\theta_{\mu}^{\text{anc}}$	Dg_{μ}	$\theta_{\text{ratio}} (\theta/\theta_{\text{anc}})$	T_{years}	Remarks
		0.851				5700	
Ambon	0.186	[0.651 – 1.094]	42.07 [29.19 – 42.1]	0.114 [0.037 – 0.899]	0.0202 [0.009 – 0.033]	[1850 – 44,950]	Decline
Banda	0.025	[0.31 – 0.859]	38.06 [21.23 – 49.71]	0.769 [0.427 – 1.356]	0.0142 [0.007 – 0.027]	38,450 [21,350 – 67,800]	Decline
Ternate	0.592	[4.413 - 6.606]	2.805 [2.765 - 4.917]	0.572 [0.422 - 0.899]	1.984 [1.505 - 2.16]	26,350 [21,100 – 44,950]	Expansion
Tidore	0.497	[1.31 - 9.245]	8.448 [2.872 -14.89]	0.866 [0.00012 -19.09]	1.761 [7.01e-05 - 9.387]	43,000 [6 – 954,500]	Constant
Bacan	0.563	[5.509 - 6.996]	6.992 [3.203 - 3.293]	7.16 [6.716 - 14.74]	2.183 [0.145 - 6.848]	358,000 [335,800 – 737,00]	Constant

p_{GSM} , parameter of geometric distribution of mutation steps; θ_{μ} , recent effective population size; $\theta_{\mu}^{\text{anc}}$, ancestor effective population size (individual number); Dg_{μ} , scaled parameter of time; $\theta_{\text{ratio}} = \theta/\theta_{\text{anc}}$, the present and ancestral scaled population sizes in mutation rate ($\theta = 4N_{\mu}$, $\theta_{\text{anc}} = 4N_{\mu}^{\text{anc}}$); Here we assume $D = G * \mu$, where G is time measured by generations, μ is the mutation rate per locus per generation. We assumed a generation time (g) of 50 years to calculate T_{years} : $T_{\text{years}} = D * g / \mu$. If $\theta_{\text{ratio}} < 1$, the population has experienced contraction; if $\theta_{\text{ratio}} > 1$, the population has experienced expansion.

Supplementary Files

Supplementary File S1. The following is a comprehensive inventory of the samples utilized in this study, encompassing species information, geographic coordinates, population details, as well as their application for SSR and cpDNA data analysis. However, due to the longevity of the list and save some paper space for sustainability, we upload them in our GitHub page, which easily accessible in this following link: <https://github.com/jaktykusuma/myristica/tree/master/fragrans>.

Supplementary File S2. This script serves the purpose of filtering fastq files based on desired quality parameters. We have made it available on our GitHub page for convenient access. You can find the script easily by following this link: <https://github.com/jaktykusuma/myristica/tree/master/scripts>.

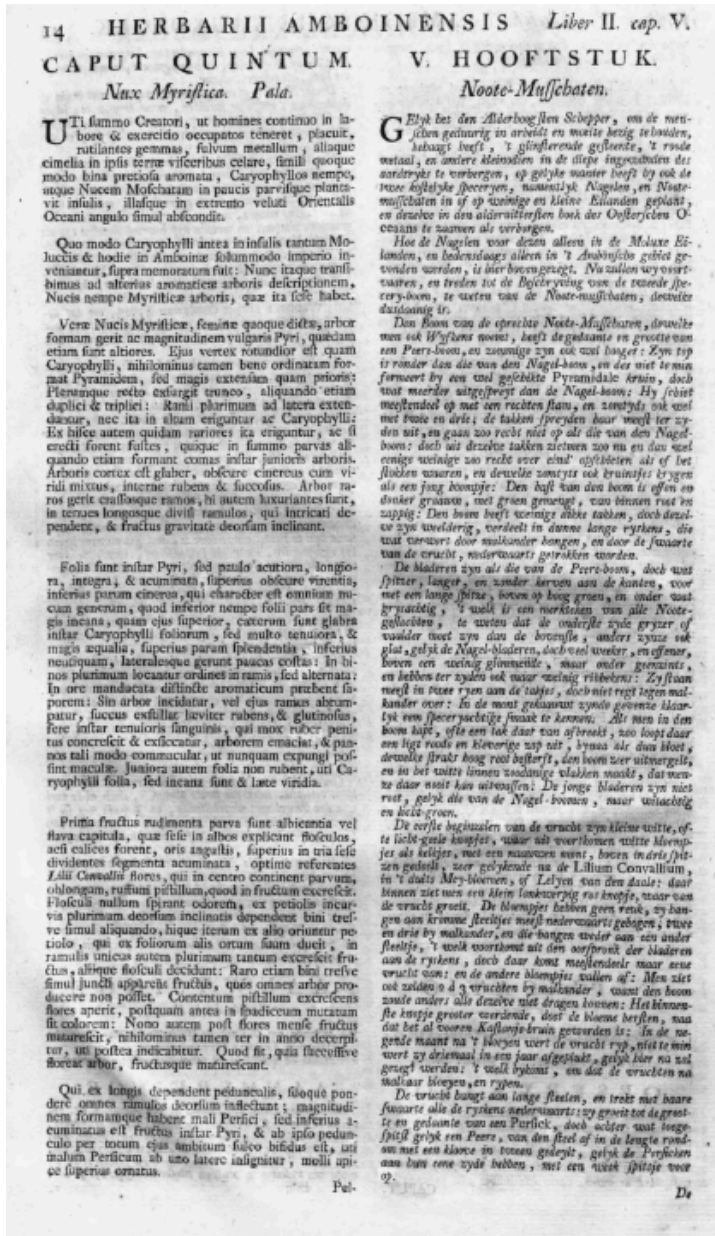
Supplementary File S3. This script serves the purpose of cpDNA analysis, from mapping to SNP calling. We have made it available on our GitHub page for convenient access. You can find the script easily by following this link: <https://github.com/jaktykusuma/myristica/tree/master/scripts>.

Supplementary File S4. Historical text describing the presence of *Myristica fragrans* in the Moluccas. Antonio Pigafetta - *Premier voyage autour du monde sur l'escadre de Magellan*. In this manuscript, Pigafetta documented the presence of cloves in the islands of Moluccas; however, he noted that the fruit quality was subpar. Additionally, he made a similar observation regarding nutmeg, stating that it too faced the same unfavorable conditions as previously mentioned.

girofler réussit mieux sur les coteaux que dans la plaine. La feuille, l'écorce et la partie ligneuse même de l'arbre ont une odeur aussi forte et autant de saveur que le fruit même. Si ce dernier n'est pas cueilli dans sa juste maturité, il devient si gros et si dur qu'il n'y reste de bon que l'écorce¹. Il n'y a de giroflers que dans les montagnes des cinq îles Malucco, et quelques arbres dans l'île de Giailolo et sur l'île de Mare, entre Tadore et Mutir; mais leurs fruits ne sont pas si bons. On prétend que le brouillard leur donne un certain degré de perfection; ce qu'il y a de certain, c'est que nous vîmes chaque jour un brouillard en forme de petits nuages environner tantôt l'une et tantôt l'autre des montagnes de ces îles. Chaque habitant possède quelques giroflers, auxquels il veille lui-même, et dont il va cueillir les fruits, mais sans en soigner la culture. Dans chaque île on donne un nom différent aux clous de girofle, on les appelle *ghomodes* à Tadore, *bougalavan* à Sarangani, et *chianche* aux îles Malucco.

Cette île produit aussi la noix muscade², qui res-

Supplementary File S4. Historical text describing the presence of *Myristica fragrans* in the Moluccas. Georg Rumphius - *Herbarium amboinense*. Rumphius documented the presence of nutmeg in the North Moluccas, but its accuracy was questioned as it is possible that the species observed was not *Myristica fragrans*. The significant aspect of his account lies in his mention of the existence of 'laki-laki' nutmeg, referring to male trees, in the North Moluccas.



Supplementary File S4. Historical text describing the presence of *Myristica fragrans* in the Moluccas. Pierre Sonnerat - *Voyage à la Nouvelle Guinée*. Sonnerat also mention the presence of male trees of nutmeg, but it might refer to *Myristica argentea*.

Je finirai cet article par la description des différentes especes d'Epicerics que me procurerent les Papoux.

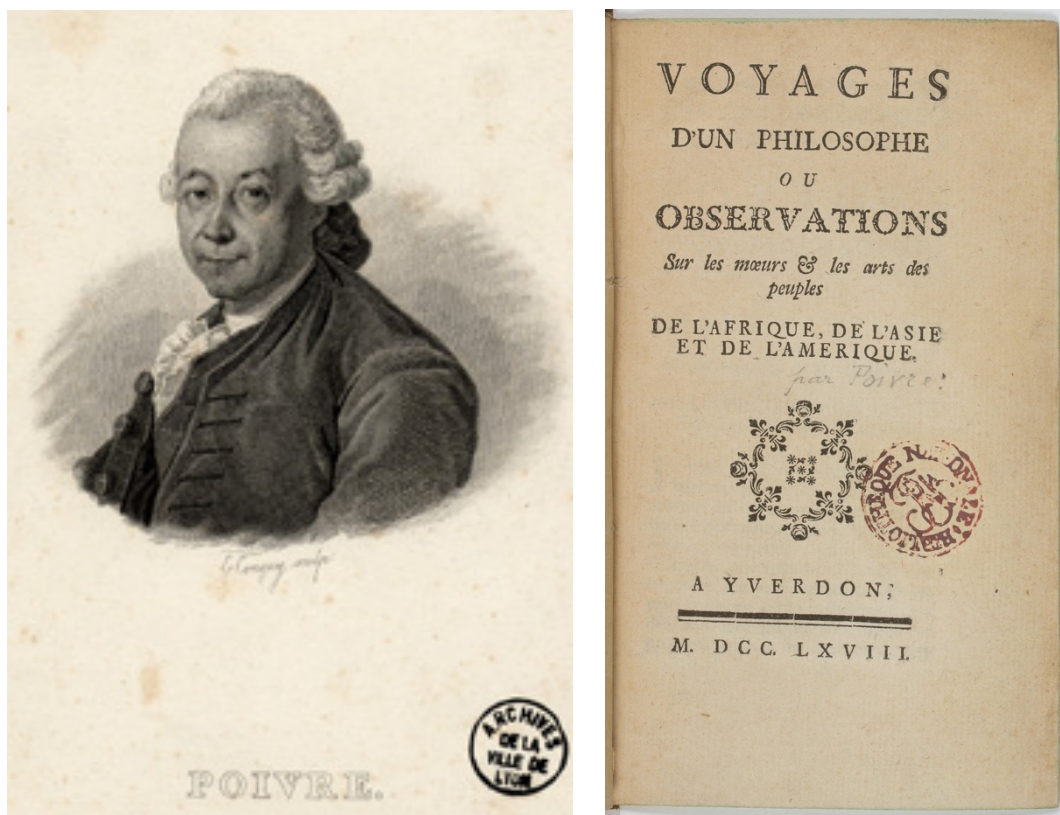
Parmi ces différentes especes, j'en trouvai trois de Muscades, dont une n'étoit point aromatique, & trois especes de Gerofliers, qui ne l'étoient pas non plus. A, dans la Planche ci à côté, représente la Muscade mâle, & B la Muscade femelle. La premiere est plus estimée en Europe, & la seconde est préférée dans l'Inde. Les Indiens l'appellent Kilaki. Indépendamment de ces deux variétés de Muscades, il y en a une troisieme qui se trouve à Salvaty, qui est beaucoup plus longue & plus aromatique que la Muscade appelée Kilaki : c'est cette derniere espece sur-tout qui est transportée par les Macassards à la presqu'Isle Malaye; d'où elle se répand à la Chine, à la côte de Coromandel, dans le Bengale, &c.

Pl. CXVI,
CXVII &
CXVIII.

Le Muscadier s'éleve jusqu'à trente pieds de haut; ses feuilles sont alternes, vertes & lissées en-dessus, blanchâtres en - dessous, attachées à un pédicule

CHAPTER 4

Propagation history and worldwide genetic diversity distribution of the Banda nutmeg (*Myristica fragrans* Houtt.)



The left side of this depiction showcases a portrait of **Pierre Poivre** (1719 - 1786), a notable French horticulturist and botanist. Poivre was the first person who was responsible for the introduction of nutmeg outside of its native habitat in the Moluccas. He presented a proposition to the *Compagnie des Indes Orientales* (CIO) to acquire spices from various regions in the East. Poivre initially struggled to cultivate nutmeg successfully, even resorting to smuggling. His rival, Jean Baptiste Christophore Fusée-Aublet, contended that Poivre's nutmeg was not authentic. However, through establishing amicable relations with Moluccan kings, Poivre ultimately acquired propagation materials and planted them in La Réunion and Mauritius, achieving triumph in his endeavors. On the right side of the illustration is his book, *Voyages d'un philosophe, ou observations sur les mœurs et les arts des peuples de l'Afrique, de l'Asie et de l'Amérique*, published in 1769, which tells the story about his journey throughout the world.

Chapter 4

Propagation history and worldwide genetic diversity distribution of the Banda nutmeg (*Myristica fragrans* Houtt.)

(In preparation)

Jakty Kusuma^{1,2}, Nora Scarcelli¹, Marie Couderc¹, Jérôme Duminil¹

¹ DIADE, Univ. Montpellier, IRD, CIRAD, Montpellier, France.

² Politeknik Negeri Lampung, Bandar Lampung, Indonesia.

ABSTRACT

Nutmeg (*Myristica fragrans*) is a tropical tree species that produces the eponymous spice which is widely used in the culinary and pharmaceutical industries. The species, that originates from Moluccas, Indonesia, has been widely introduced to other tropical parts of the world. The history of this introduction, as well as the level of genetic diversity contained in introduced populations remain poorly documented. In this study, we collected 882 nutmeg individuals from various locations worldwide, including its native range in the Moluccas, as well as Asia, America, Africa, and the Pacific. We analyzed their genetic composition using nuclear microsatellites (SSR), whole chloroplast genomes (cpDNA), and nuclear DNA (nrDNA) sequences. Our results revealed that most of the introduced populations were derived from South Moluccas. Our findings provide valuable insights into the population history of nutmeg and have significant implications for its conservation and management. The observed genetic diversity and structure of nutmeg populations can be used to inform the development of sustainable management programs and the identification of potentially valuable genetic resources. Furthermore, the identification of genetically distinct populations and evidence of admixture highlight the need for strategic management and monitoring of nutmeg populations, particularly in the Moluccas, to prevent the loss of genetic diversity and preserve the unique genetic heritage of this important species.

Key words: nutmeg, spice trade, genetic diversity, introduction

INTRODUCTION

The Banda nutmeg (*Myristica fragrans* Houtt.) is a spice tree species originating from the Moluccas, Indonesia. Once known as a “green gold”, it has captured the attention of the world for centuries. Archeobotanical findings from Pulau Ay (one of the 10 Banda Islands, South Moluccas, Indonesia) suggest that *M. fragrans* is used locally in a food context at least since the early Neolithic (*ca* 1000 B.C.; Lape et al., 2018) . Since ancient times, the species was an important object of trade in South-East Asia, involving notably Chinese, Malay, Javanese, and Arabo-Persian merchants. Arab use of nutmeg as medicine dates back to the first centuries A.D. (Warburg, 1897). The spice is known in Europe since the Middle age. The search for this elusive spice tree spanned from the 6th century B.C. (Weiss, 2002) to the 14th century A.D. when its origin was unknown to the Europeans, making it a highly priced commodity (Zohary et al., 2012).

In the 16th century, in order to source these precious aromatic nuts directly from their place of origin, Europeans explored sea routes in search of the “Spice Islands” (the Moluccas). In 1511, the Portuguese established a base at Malacca. In 1511-1512, they sailed along the coasts of Java, Bali, Lombok, Sumbawa, and Flores to the Moluccas, where they found the place of origin of the species. From 1609 on, the Dutch, through the Vereenigde Oost-indische Compagnie (VOC, Dutch East India Company), controlled the region and established a strong, military-defended monopoly of this world’s sole source of nutmeg in South Moluccas: the Banda archipelago (Dhont, 2022; Geerken, 2021; Hanna, 1991). The maintenance of this monopoly over the production of nutmeg led to wars and

conflicts, not only between the local Bandanese but also among the European nations.

Despite the Dutch monopoly on nutmeg production and trade, Pierre Poivre, a French colonial administrator and horticulturist, and his collaborators (Provost in particular) were the first to break the control over this spice by obtaining plants and seeds of nutmeg (different tentative in the period 1769–1771) and by introducing the species in the ‘Isle de France’ (Mauritius) (Brixius, 2015). The species was then probably first introduced in the Americas around 1773, with a transfer of plant material from the Isle of France to the ‘Isle de Cayenne’ (French Guiana) (Zumbroich, 2005) The English explorers, Frank Gurney also managed to introduced the nutmeg with seeds brought from Banda in 1843 (Steele, 2006). From then on, the nutmeg industry began to flourish in other tropical parts of the world. Food and Agriculture Organization (FAO) reported that the main nutmeg producers today are Indonesia, India, Granada, and Guatemala (2020). Today, nutmeg is readily available and is widely used for its unique flavour and aroma in various culinary applications, including in baking, drinks, and traditional dishes.

Characterizing the history of introduction of nutmeg and spatial patterns of genetic diversity worldwide can contribute to the sustainable management and use of the species, notably in relation to its adaptation to climate change. In this study, we utilized a set of molecular markers (nuclear microsatellites genotyping —SSR for short—, whole chloroplast genomes sequencing —cpDNA—, and nuclear-targeted sequencing—nrDNA—) to investigate history of introduction of the species. Our findings shed new light on the fascinating story of this valuable spice and its journey from its native range to the rest of the world.

RESULTS

We collected a total of 882 nutmeg samples from various locations across the world. Among them, 811 were leaves collected in the field and dried in silica gel. These samples were collected from native populations in the Moluccas (group of South Moluccas: Ambon, Banda, and Seram; and group of North Moluccas: Ternate, Tidore, Bacan, Halmahera, Obi, and Makian), and from introduced populations in Indonesia (Bali, Java, Sulawesi, Sumatera, Papua, Siau, and Simeulue; see Table 1). This sampling was completed with herbarium samples ($N = 71$) from MNHN Paris (P) and Naturalis Leiden (L) in order to cover the maximum places of species' introduction in the Rest of the World (Brazil, Dominica, Equatorial Guinea, French Guyana, French Polynesia, French Reunion, Guadeloupe, Hawaii, India, Madagascar, Malaysia, Martinique, Mauritius, Papua New Guinea, Singapore, Sri Lanka, Tanzania, and Thailand; Table 1). Individuals were analyzed in two different steps. We first characterized all 882 individuals using 14 polymorphic microsatellite (SSR) markers (Kusuma et al., 2020), and then select part of these individuals ($N = 269$ covering the distribution of our sampling) to conduct (1) whole chloroplast genome sequencing (cpDNA) and (2) targeted enrichment and sequencing of hundreds of nuclear loci (nrDNA) using baits developed by Couvreur et al. (2019).

Tracing the origin of introduced nutmeg

We first assessed the genetic structure of nutmeg population using nuclear markers. The global genetic structure among populations as measured by the F_{ST} (Goudet, 2005) was 0.145 in SSR and 0.136 in nrDNA. The pairwise F_{ST} between populations ranged from 0.009 (between Obi and Halmahera) to 0.327 (between

Bali and Papua) for SSR dataset (Supplementary Figure S1) and from 0 (between Simeulue and Sumatera) to 0.161 (between Bali and Seram) (Supplementary Figure S2).

We further inferred the population structure using Bayesian clustering analysis on nuclear markers as implemented in fastStructure (Pritchard et al., 2000; Raj et al., 2014). We found that both nuclear datasets (SSR and nrDNA) showed the presence of two distinct genetic clusters ($K = 2$) within native populations: a South Moluccas genetic cluster comprising all populations from the region, and a North Moluccas genetic cluster containing mainly individuals from Halmahera (Figures 1-2). We observed a pattern of admixture between the two clusters in other North Moluccas populations (Ternate, Tidore, Bacan, Obi, and Makian, with these populations being genetically intermediate between the South and North Moluccas groups). The majority of the introduced populations were assigned to the South Moluccas group, except for Hawaii, India, Malaysia, Papua New Guinea and Thailand, which were assigned to the North Moluccas group. In contrast, the nrDNA sequence data revealed a clearer population structure within native populations, with less evidence of admixture in the majority of the North Moluccas populations. Furthermore, we found that Indonesian introduced populations were globally assigned to the South Moluccas group, but present an admixture pattern with the North Moluccas group. The pattern observed at $K = 3$ (Figure 2) suggests that introduced population are mainly of South Moluccas origin, but some individuals are also assigned to the North Moluccas group, which suggest that propagation material from both groups was used to create the introduced populations.

Finally, we investigated the genetic structure by reconstructing a cpDNA haplotype network using a maximum parsimonious method (Kannan & Wheeler, 2012). We did not observe any clear phylogeographical structure among populations or groups based on cpDNA data (Figure 3). The constructed haplotype network revealed a dominant haplotype (*Hap1*) that did not allow any diagnosis on the origin of introduced populations, as this haplotype is widespread across all native populations. The sharing of haplotypes between native and Indonesian introduced populations (Hap 5, 8, 23, 35, 36) supports the use of propagation material from both South and North Moluccas for the planting of new populations. The *Hap36* (composed from the population of Banda, Bali, Java, Papua, Siau, and Singapore) is the only common haplotype between native and introduced populations from the Rest of Indonesia and the Rest of the World, which suggests a South Moluccas origin of propagation material.

Genetic diversity: native and introduced populations

We found high genetic diversity in the North Moluccas for each of the three genetic datasets (SSR, cpDNA, nrDNA; Tables 2-3). From SSR data, the North Moluccas showed a mean number of alleles of 13.73, ranging from 4.60 in Makian to 11.87 in Bacan. In contrast, the South Moluccas presented a low mean number of alleles ($N_A = 5.53$), ranging from 3.33 in Seram to 4.93 in Ambon. This is lower than the introduced populations from Indonesia ($N_A = 10.40$) and the Rest of the World populations ($N_A = 6.16$). These results are further supported by the observed number of private alleles per group (South Moluccas, North Moluccas, the Rest of Indonesia, and the Rest of the World). Specifically, the North Moluccas exhibit a

higher number of private alleles than the South Moluccas. Interestingly, we observe of the presence private alleles within Indonesian introduced populations, particularly in Sumatra. In contrast, no private allele is detected in the group representing the Rest of the World (see Supplementary Figure S3).

Using cpDNA genome sequence, we found a total of 48 haplotypes, with 39 haplotypes present in the North Moluccas, 19 in the South Moluccas, 17 in the Rest of Indonesia, and 19 in the Rest of the World (Table 3). Only one haplotype is unique to an introduced population (Hap 37), all other haplotypes found in introduced populations are also found in native populations. The levels of nucleotide diversity were equivalent in the different groups investigated.

The nrDNA sequences revealed 1,164 single-nucleotide polymorphisms (SNPs). Allelic richness ranged from 1,494 in the North Moluccas to 1,300 in the South Moluccas. Our results also showed that the North Moluccas group exhibits the highest level of genetic diversity, as indicated by the highest expected heterozygosity value (0.532). On the other hand, the lowest level of genetic diversity was found in the group of the Rest of Indonesia, with the lowest expected heterozygosity value of 0.324.

DISCUSSION

History of introduced nutmegs: what is their geographic origin?

Our findings contribute to the understanding of the introduction history of nutmeg outside its native range. Before the 16th century, the geographic origin of nutmeg was unknown to Europeans (Bickmore, 1867). The high demand for this valuable spice played a significant role in driving the Age of Exploration and initiating the intensive Spice Trade (Dhont, 2022; Geerken, 2021). However, the

origin of this spice remained shrouded in mystery, as Europeans had limited knowledge of the distant lands where nutmeg was cultivated. The Portuguese were among the pioneers in this endeavor, successfully reaching the Spice Islands in the late 15th century (Warburg, 1896). This sparked a surge of interest and competition among European powers to establish their presence and control over the lucrative spice trade.

The VOC played a significant role in the spice trade, particularly in the Banda Islands. Under the VOC's strict regulations, nutmeg cultivation on the Banda Islands was closely monitored and limited to specific areas (Wright, 1958). Despite the VOC's efforts to maintain a monopoly, Pierre Poivre and collaborators smuggled nutmeg plants and seeds from Gilolo (or Jailolo, a kingdom in Halmahera Island) in the 18th century, transporting them first to Mauritius and later on to French Guiana. However, our research findings contradict this historical fact. Genetic analysis reveals that the nutmeg clusters in Mauritius and French Guiana, and other former French colonies, align with the South Moluccas rather than the North Moluccas.

One possible explanation is that the nutmeg material found in Mauritius and French Guiana, despite being sourced in North Moluccas, would actually have originated from the South Moluccas. To explain this, a reference in Garcia de Orta's *Colóquios dos simples e drogas da India*, states that "nutmeg also grew in the Moluccas and Ceylon but did not bear fruit like the nutmeg trees on the Banda Islands" (Orta, 1563). It is important to note that during the 16th century, the term "Moluccas" referred to a specific set of five islands: Ternate, Tidore, Halmahera, Bacan, and Makian (Leirissa, 1977). This is different from the present-day

definition of the Moluccas, which encompasses all islands in the Wallacean region (Andaya, 1991). We hypothesize that the trees in question were likely male trees, as they did not bear any fruit (Purwiyanti et al., 2020). However, in certain instances, a small quantity of fruit was produced, indicating the possibility of hermaphroditism under specific climatic conditions (Kaihatu et al., 2021; Karmanah et al., 2016). During that time, it was reported that the local inhabitants made no attempt to cultivate them (Orta, 1563). Moreover, the Gilolo kingdom, was frequently visited by *orang kaya*—rich man, who sometimes brought slaves (and small amount of nutmeg for local use) from Banda (Villiers, 1981). We then speculate that the nutmeg propagation material obtained by Poivre was probably originated from Banda, carried by visitor to Gilolo. These historical pieces further support our suggestion that can explain the connection between the genetic clusters of the populations in Mauritius and French Guiana, as well as their former French colonies with the South Moluccas.

A different scenario seems to have occurred with former British colonies, as introduced individuals from India, Malaysia, Papua New Guinea, and Thailand used in our study are more related to the North Moluccas than to the South Moluccas. We only found little historical facts on the origin of planting material used by the British (Steele, 2006), which explains that the propagation material was originated from Banda. This includes the common knowledge that the British once had a possession in one of the Banda islands (Ai Island). There is also historical account that the British were creating a nutmeg plantation in Singapore, whose seeds were taken from *Bencoolen*—Bengkulu (a province in Sumatera Island) (Gouger et al., 1843). But the genetic evidence suggested that the Singapore

population is more related structure to the South Moluccas than to the North Moluccas.

Meanwhile, in all introduced Indonesian populations, it has been suggested that propagation material predominantly came from the South Moluccas. This is expected due to the historical efforts to increase nutmeg production since the 19th century (Juwita & Tsuchida, 2017; Madiki et al., 2015), with the new plantations primarily established using seeds from Ambon or Banda (Simanjorang et al., 2020). Given the presence of a signal of admixture within the genome of Indonesian introduced populations, it is also likely that a tiny part of introduced material was sourced from the North Moluccas. This result is further supported by the cpDNA data, with the presence of typical North Moluccas haplotypes in introduced Indonesian population.

Levels of genetic diversity in native and introduced populations

The highest level of genetic diversity has been observed in the North Moluccas, and the lowest in the South Moluccas, whereas the introduced Indonesian populations show an intermediate value. We found that this reduction of allelic richness in South Moluccas is more than 50%, compared to the Northern Moluccas ($AR = 10.42$). This pattern suggests that the historical events during the VOC era, marked by extensive extirpation, have played a significant role in shaping the genetic diversity of the South Moluccas population. It is well understood that the population in the North Moluccas were remain untouched by human activities, as indicated by the presence of private alleles, which found to be abundant in the Bacan and Halmahera Island. The intermediate value observed in introduced

Indonesian populations supports the mix origin of these plantations, with material sourced both in the South and in the North Moluccas. Our sampling scheme outside of Indonesia, based on the sampling of few herbarium samples per country, did not allowed us estimating levels of genetic diversity at the population level. We tentatively estimated the levels of diversity at the country level, but the results obtained needs to be interpreted with caution given the very low effectives. For countries represented by at least five individuals, we estimated the rarefied allelic richness, and results suggest a low level of genetic diversity in these countries, except in Singapore. Additional sampling in the different countries where the species has been introduced are absolutely needed to estimate the levels of genetic diversity in these populations and to compare them with the level of genetic diversity from native populations.

CONCLUSION

Globally, our results underscore the dynamic interplay between human management practices and the genetic composition of introduced populations. It highlights the importance of considering genetic diversity as a key factor in conservation and management strategies, particularly in the native range of the species. Maintaining and protecting the genetic diversity of the North and South Moluccas populations is crucial for its long-term viability and resilience, while further studies are needed to monitor and assess the potential of introduced populations in different geographical regions. The genetic examination of historically introduced species enhancing our understanding of the evolutionary consequences of human management practices on agrobiodiversity (Rius &

Darling, 2014). Furthermore, these analyses underscore the significance of preserving genetic diversity within the native range of species, as it serves as a crucial reservoir for evolutionary potential and resilience in the face of ongoing environmental changes.

MATERIALS AND METHODS

Sampling, laboratory and bioinformatic activities

This study is based on a comprehensive dataset consisting of 882 nutmeg individuals that span the entirety of *M. fragrans* global distribution (Table 1 and Supplementary Table S1). Material was either collected directly in the field ($N = 811$; fresh young leaves in silica gel) or from herbarium specimens ($N = 71$; Table 1 and Supplementary Table S1) from two botanical institutions, MNHN Paris (P) and Naturalis Leiden (L). We used a modified MATAB protocol to extract DNA (Mariac et al., 2006). The 882 individuals were first genotyped using fifteen polymorphic nSSR markers (Kusuma et al., 2020). We then selected 269 individuals from all populations (3 to 5 individuals per population when possible) to construct genomic libraries. These libraries were prepared for whole genome sequencing to recover whole cpDNA genome data, and for a nuclear-targeted enrichment, using Annonaceae baits as defined in Couvreur et al. (2019). Both types of libraries were sequenced using Novaseq, Illumina.

For cpDNA, we demultiplexed, filtered, and paired the raw reads according to Scarcelli et al. (2016). We then mapped the reads on the *M. fragrans* cpDNA genome (de Oliveira et al., 2020) using BWA v.0.7.12 (Li & Durbin, 2009). We performed SNP calling using VarScan2 (Koboldt et al., 2012).

For nrDNA sequences, we built a pseudo-reference using the pipeline SECAPR v.1.14 (Andermann et al., 2018). BWA v.0.7.12 (H. Li & Durbin, 2009) was used for mapping and GATK v.4 (McKenna et al., 2010) was used to remove duplicates and to call SNPs. We filtered the data by keeping only biallelic SNPs with a mapping quality >40%, depth >25, and quality by depth >2 using BCFtools (H. Li, 2011). Additionally, we removed SNPs with allele frequency <0.01 and excluded monomorphic sites.

Tracing the origin of introduced populations

We used a Bayesian clustering approach as implemented in STRUCTURE software (Pritchard et al., 2000) for nSSRs 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. We also used fastStructure (Raj et al., 2014) for nuclear DNA sequence dataset with the same parameters. The value of K on both nuclear markers were estimated using StructureSelector (Y.-L. Li & Liu, 2018).

Comparing the level of genetic diversity in native and introduced populations

In nSSRs, we characterized the distribution of the genetic diversity at the population level and group level using SPAGeDI 1.5 (Hardy & Vekemans, 2002), with the following indices: number of alleles (N_A), observed and expected heterozygosity (H_O and H_E , respectively), rarefied allelic richness (A_R), and the inbreeding coefficient (F_{IS}).

For cpDNA, we used SNPs to calculate haplotype number (H) and nucleotide diversity (π), using DnaSP 6.12.03 (Librado & Rozas, 2009). For the nrDNA sequence dataset, we computed individual and population measurements of expected and observed heterozygosity (H_E and H_O) as well as allelic richness using ‘*adegenet*’ (Jombart & Ahmed, 2011) and ‘*vcfR*’ (Knaus & Grünwald, 2017) packages available in R (R Core Team, 2022).

ACKNOWLEDGEMENTS

The authors would like to thank all farmers and nutmeg’s land owners in Indonesia for providing information on nutmeg history as well as leaves materials. We also thank Roxali Bijmoer and Marnel Scherrenberg from Naturalis Biodiversity Center, Leiden, The Netherlands as well as Myriam Gaudeul from National Museum of Natural History, Paris, France for their invaluable help during the specimen collection. We further thank Pierre Mournet from CIRAD, Montpellier, France and Sylvain Santoni, Audrey Weber from INRAe, Montpellier, France for providing assistance during lab analysis. We also greatly appreciate the support from the Indonesia Endowment Fund for Education (LPDP) to JK with grant No. 0003719/AFR/D/BUDI-2018. This project was also supported by the Agropolis Fondation under the reference ID 1502-503 and 1502-504 through the « *Investissements d’avenir* » programme (Labex Agro: ANR-10-LABX-0001-01), under the frame of I-SITE MUSE (ANR-16-IDEX-0006). The authors acknowledge the ISO 9001 certified IRD i-Trop HPC (South Green Platform) at IRD Montpellier for providing HPC resources that have contributed to the research results reported within this paper.

ETHICS PERMISSION

This research was done with the authorization of Indonesian National Research and Innovation Agency (BRIN) with research permit to JD (144/SIP/FRP/E5/Ditk.KI/V/2018) and collecting permit to JK (B-3382/IPH.1/KS.02.04/IX/2019 and B-75/IV/KS.01.04/3/2022).

AUTHORS' CONTRIBUTION

JK and JD conducted the field work; JK, NS, and JD conceptualized and designed research; JK and MC performed laboratory experiments; JK and NS analyzed data; JK wrote the original draft of the manuscript; All authors read and approved the final manuscript.

REFERENCES

- Andaya, L. (1991). Local Trade Networks in Maluku in the 16th, 17th and 18th Centuries. *Cakalele*, 2(2), 71–96.
- Andermann, T., Cano, Á., Zizka, A., Bacon, C., & Antonelli, A. (2018). SECAPR—a bioinformatics pipeline for the rapid and user-friendly processing of targeted enriched Illumina sequences, from raw reads to alignments. *PeerJ*, 6, e5175. <https://doi.org/10.7717/peerj.5175>
- Bickmore, A. S. (1867). A Description of the Banda Islands. *Proceedings of the Royal Geographical Society of London*, 12(5), 324–334. <https://doi.org/10.2307/1799078>
- Brixius, D. (2015). A Pepper acquiring Nutmeg: Pierre Poivre, The French Spice Quest and the Role of Mediators in Southeast Asia, 1740s to 1770s. *Journal of the Western Society for French History*, 43(1).
- Couvreur, T. L. P., Helmstetter, A. J., Koenen, E. J. M., Bethune, K., Brandão, R. D., Little, S. A., Sauquet, H., & Erkens, R. H. J. (2019). Phylogenomics of the Major Tropical Plant Family Annonaceae Using Targeted Enrichment of Nuclear Genes. *Frontiers in Plant Science*, 9. <https://doi.org/10.3389/fpls.2018.01941>

- 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>
- 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>
- Geerken, H. H. (2021). *The Gold of the Bandas: The History of the Nutmeg: The forgotten islands that once made world history*. BoD--Books on Demand.
- 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>
- Gouger, H., Montgomerie, Wm., Church, J., Barber, N., Cobb, & Bland, W. J. (1843). Cultivation of nutmegs at Singapore. *Transactions of the Society, Instituted at London, for the Encouragement of Arts, Manufactures, and Commerce*, 54, 38–50. <http://www.jstor.org/stable/41326885>
- 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.
- 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>
- Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics*, 27(21), 3070–3071. <https://doi.org/10.1093/bioinformatics/btr521>
- Juwita, R., & Tsuchida, S. (2017). Current conditions and profitability of the nutmeg industry in Bogor regency, Indonesia. *Journal of the International Society for Southeast Asian Agricultural Sciences*, 23, 33–44.
- Kaihatu, S., Pesireron, M., Ohorella, I., Senewe, R. E., Waas, E. D., & Gaffar, A. (2021). Effect of natural male and female (sex ratio) on nutmeg (*Myristica fragrans*) production on Maluku. *E3S Web of Conferences*, 306, 01019. <https://doi.org/10.1051/e3sconf/202130601019>
- 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>

- Karmanah, Maslahat, M., & Nurhayati, L. (2016). Morphology and isozyme band-profile as sexual determinant of nutmeg plant (*Myristica fragrans* Houttj). *IOP Conference Series: Earth and Environmental Science*, 31, 012037. <https://doi.org/10.1088/1755-1315/31/1/012037>
- Knaus, B. J., & Grünwald, N. J. (2017). vcfR: a package to manipulate and visualize variant call format data in R. *Molecular Ecology Resources*, 17(1), 44–53. <https://doi.org/10.1111/1755-0998.12549>
- 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>
- 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>
- Leirissa, R. Z. (1977). Local potentates and the competition for cloves in early seventeenth century ternate [North Moluccos]: a preliminary study. *Proceedings of the Seventh Conference of the IAHA*.
- Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, 27(21), 2987–2993. <https://doi.org/10.1093/bioinformatics/btr509>
- 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, Y.-L., & Liu, J.-X. (2018). StructureSelector: A web-based software to select and visualize the optimal number of clusters using multiple methods. *Molecular Ecology Resources*, 18(1), 176–177. <https://doi.org/10.1111/1755-0998.12719>
- 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>

- Madiki, A., Guritno, B., Syekhfani, S., & Aini, N. (2015). The Relationship between Plant Density and Microclimate and Nutmeg (*Myristica fragrans* Houtt) Production in Nutmeg and Coconut Mixed-Planting System in Wakatobi District in Indonesia. *Journal of Agricultural Science*, 7(12), 187. <https://doi.org/10.5539/jas.v7n12p187>
- Mariac, C., Luong, V., Kapran, I., Mamadou, A., Sagnard, F., Deu, M., Chantereau, J., Gerard, B., Ndjeunga, J., Bezançon, 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>
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20(9), 1297–1303. <https://doi.org/10.1101/gr.107524.110>
- Orta, G. da. (1563). *Coloquios dos simples e drogas da India* (Vol. 1). Maxtor.
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, 155(2), 945 LP – 959. <http://www.genetics.org/content/155/2/945.abstract>
- Purwiyanti, S., Wahyu, Y., Rostiana, O., & Sudarsono. (2020). Correlation between morphological characters and the sex phenotypes of *Myristica fragrans* Houtt Trees. *IOP Conference Series: Earth and Environmental Science*, 418(1), 012033. <https://doi.org/10.1088/1755-1315/418/1/012033>
- R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- Raj, A., Stephens, M., & Pritchard, J. K. (2014). fastSTRUCTURE: Variational Inference of Population Structure in Large SNP Data Sets. *Genetics*, 197(2), 573–589. <https://doi.org/10.1534/genetics.114.164350>
- 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 NGS 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>
- Simanjorang, T., Irham, Waluyati, L. R., & Mulyo, J. H. (2020). Comparative and competitive advantages of nutmeg farming in two regions in Maluku Province, Indonesia. *Biodiversitas Journal of Biological Diversity*, 21(3). <https://doi.org/10.13057/biodiv/d210342>

- Steele, B. (2006). *Grenada: A History of Its People* (B. Steele, Ed.; 1st ed., Vol. 1). Macmillan Caribbean.
- Villiers, J. (1981). Trade and Society in the Banda Islands in the Sixteenth Century. *Modern Asian Studies*, 15(4), 723–750. <http://www.jstor.org/stable/312170>
- Warburg, O. (1896). The Discovery of the Moluccas. *The Geographical Journal*, 7(4), 365–366. <https://doi.org/10.2307/1773549>
- Warburg, O. (1897). *Die Muskatnuss ihre Geschichte, Botanik, Kultur, Handel und Verwerthung sowie ihre Verfälschungen und Surrogate. Zugleich ein Beitrag zur Kulturgeschichte der Banda-Inseln*. Leipzig, W. Engelmann, 1897. <https://www.biodiversitylibrary.org/item/52549>
- Weiss, E. (2002). *Spice crops*. CABI.
- 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. <http://www.jstor.org/stable/41503143>
- Zohary, D., Hopf, M., & Weiss, E. (2012). *Domestication of Plants in the Old World*. Oxford University Press. <https://doi.org/10.1093/acprof:osobl/9780199549061.001.0001>
- Zumbroich, T. (2005). The introduction of nutmeg (*Myristica Fragrans* Houtt.) and cinnamon (*Cinnamomum Verum* J.Presl) to America. *Acta Botánica Venezuelica*, 28(1), 155–160. http://ve.scielo.org/scielo.php?script=sci_arttext&pid=S0084-59062005000100010&lng=en&nrm=iso&tlng=

Tables

Table 1. Samples of *Myristica fragrans* analyzed. Both fresh leaves, later dried in Silica-gel, and herbarium samples were collected.

Group	Population	Number of samples	Status
South Moluccas		242	
	Ambon	161	Silica-gel
	Banda	72	Silica-gel
	Seram	9	Silica-gel
North Moluccas		238	
	Ternate	55	Silica-gel
	Tidore	30	Silica-gel
	Bacan	79	Silica-gel
	Halmahera	51	Silica-gel
	Obi	18	Silica-gel
	Makian	5	Silica-gel
Rest of Indonesia		333	
	Bali	9	Silica-gel
	Java	31	Silica-gel
	Sulawesi	1	Herbarium
	Sumatera	96	Silica-gel
	Papua	7	Silica-gel
	Siau	121	Silica-gel
	Simeulue	68	Silica-gel
Rest of the World		69	
	Brazil	5	Herbarium
	French Reunion	3	Herbarium
	Dominica	1	Herbarium
	Equatorial Guinea	2	Herbarium
	French Guyana	5	Herbarium
	French Polynesia	2	Herbarium
	Guadeloupe	3	Herbarium
	Hawaii	2	Herbarium
	India	6	Herbarium
	Madagascar	2	Herbarium
	Malaysia	5	Herbarium
	Martinique	9	Herbarium
	Mauritius	2	Herbarium
	Papua New Guinea	6	Herbarium
	Singapore	5	Herbarium
	Srilanka	3	Herbarium
	Tanzania	3	Herbarium
	Thailand	5	Herbarium

Table 2. Genetic diversity of *Myristica fragrans* inferred by 15 nSSRs in all populations and group categories.

Group	Population	N_{INDV}	N_A	$AR_{(k=5)} / AR_{(k=69)}$	H_E	H_O
South Moluccas		242	5.53	4.09	0.566	0.534
	Ambon	161	4.93	2.59	0.539	0.511
	Banda	72	4.47	2.77	0.601	0.586
	Seram	9	3.33	2.71	0.583	0.529
North Moluccas		238	13.73	10.42	0.797	0.700
	Ternate	55	10.27	4.11	0.777	0.712
	Tidore	30	10.27	4.53	0.826	0.765
	Bacan	79	11.87	4.17	0.776	0.669
	Halmahera	51	7.27	3.48	0.797	0.776
	Obi	18	8.40	4.40	0.821	0.801
	Makian	5	4.60	3.91	0.774	0.707
Rest of Indonesia		333	10.40	6.16	0.646	0.579
	Bali	9	2.87	2.42	0.445	0.347
	Java	31	5.20	3.09	0.654	0.502
	Sulawesi	1	-	-	-	-
	Sumatera	96	7.40	3.04	0.642	0.589
	Papua	7	2.80	2.41	0.501	0.603
	Siau	121	7.20	2.87	0.593	0.581
	Simeulue	68	4.93	2.81	0.606	0.612
Rest of the World		69	6.60	6.07	0.667	0.610
	Brazil	5	1.73	1.65	0.296	0.467
	French Reunion	3	1.47	-	0.280	0.467
	Dominica	1	-	-	-	-
	Equatorial Guinea	2	1.67	-	0.411	0.567
	French Guyana	5	1.87	1.84	0.431	0.653
	French Polynesia	2	1.67	-	0.411	0.567
	Guadeloupe	3	1.73	-	0.391	0.601
	Hawaii	2	1.67	-	0.466	0.667
	India	6	2.43	1.87	0.542	0.352
	Madagascar	2	1.73	-	0.488	0.733
	Malaysia	5	2.60	2.44	0.554	0.753
	Martinique	9	1.73	1.67	0.294	0.489
	Mauritius	2	1.67	-	0.444	0.667
	Papua New Guinea	6	1.87	1.86	0.444	0.802
	Singapore	5	3.53	3.16	0.663	0.513
	Srilanka	3	1.93	-	0.622	0.533
	Tanzania	3	2.07	1.65	0.457	0.644
	Thailand	5	1.73	1.43	0.432	0.445

Number of individuals analyzed (N_{INDV}), mean number of alleles per populations (N_A), rarefied allelic richness for k=5 gene copies for all populations ($AR_{(k=5)}$), rarefied allelic richness for k=69 gene copies for group categories ($AR_{(k=69)}$), as well as expected heterozygosity (H_E), and observed heterozygosity (H_O).

Table 3. Genetic diversity of *Myristica fragrans* populations and group categories at cpDNA loci and nrDNA sequences.

Group	Populations	N_{INDV}	cpDNA			nrDNA sequences		
			H	π	T_D	nrH_E	nrH_O	AR
South Moluccas		46	19	0.165	0.045	0.255	0.132	1.300
	Ambon	24	14	0.176	-0.393	0.245	0.129	1.300
	Banda	19	4	0.235	-0.155	0.250	0.131	1.303
	Seram	3	1	0.084	1.033	0.271	0.138	1.298
North Moluccas		79	39	0.177	0.349	0.280	0.163	1.494
	Ternate	22	9	0.114	-1.014	0.278	0.184	1.352
	Tidore	10	10	0.276	0.566	0.262	0.156	1.362
	Bacan	21	13	0.155	-0.351	0.268	0.147	1.649
	Halmahera	22	4	0.311	0.076	0.282	0.195	1.869
	Obi	3	2	0.210	1.123	0.285	0.143	1.333
	Makian	1	1	0	-	0.310	0.155	1.400
Rest of Indonesia		73	17	0.176	0.049	0.218	0.145	1.272
	Bali	1	2	0.028	1.763	0.152	0.138	1.221
	Java	14	3	0.136	-0.873	0.219	0.165	1.278
	Sulawesi	1	1	0.243	1.671	0.183	0.158	1.200
	Sumatera	24	5	0.137	-0.760	0.247	0.137	1.304
	Papua	4	3	0.225	-0.482	0.221	0.146	1.278
	Siau	19	2	0.118	-0.973	0.241	0.139	1.294
	Simeulue	10	1	0.347	0.724	0.264	0.138	1.332
Rest of the world		69	19	0.157	-0.892	0.219	0.158	1.306
	Brazil	5	1	0.002	-0.892	0.219	0.154	1.292
	French Reunion	3	1	0.074	-	0.192	0.155	1.305
	Dominica	1	1	0	0.018	0.165	0.219	1.200
	Equatorial Guinea	2	1	0.139	-	0.203	0.166	1.257
	French Guyana	5	1	0.012	1.938	0.169	0.180	1.218
	French Polynesia	2	1	0.031	-	0.261	0.151	1.378
	Guadeloupe	3	1	0.034	-	0.254	0.134	1.347
	Hawaii	2	1	0.011	-	0.283	0.155	1.378
	India	6	1	0.446	1.893	0.248	0.135	1.326
	Madagascar	2	1	0.306	0.642	0.213	0.204	1.310
	Malaysia	5	1	0.473	0.446	0.213	0.125	1.305
	Martinique	9	1	0.093	0.737	0.210	0.181	1.311
	Mauritius	2	1	0.007	-	0.160	0.147	1.236
	Papua New Guinea	6	1	0.350	0.019	0.188	0.137	1.254
	Singapore	5	2	0.334	0.113	0.249	0.132	1.355
	Srilanka	3	1	0.008	0.379	0.218	0.151	1.280
	Tanzania	3	1	0.003	-	0.281	0.171	1.428
	Thailand	5	1	0.517	-	0.233	0.149	1.332

Number of individuals analyzed (N), number of haplotypes (H), nucleotide diversity (π), Tajima D's value (T_D), expected heterozygosity (nrH_E), observed heterozygosity (nrH_O) and allelic richness (AR) per population using nrDNA sequences.

Figures

Figure 1. Population structure in all 34 populations using Bayesian clustering in STRUCTURE based on SSR markers. Delta K value according to Evanno (2005) method indicating $K = 2$ is the most likely. Barplot of individual admixture proportions for all populations inferred by STRUCTURE for (A) $K = 2$ to $K = 4$, and a zoom on the Rest of the World populations (B). Global geographical distribution of individuals inferred with STRUCTURE for $K = 2$ (C), and Southeast Asia geographical distribution (D). Zooming out the Moluccas geographical distribution (E).

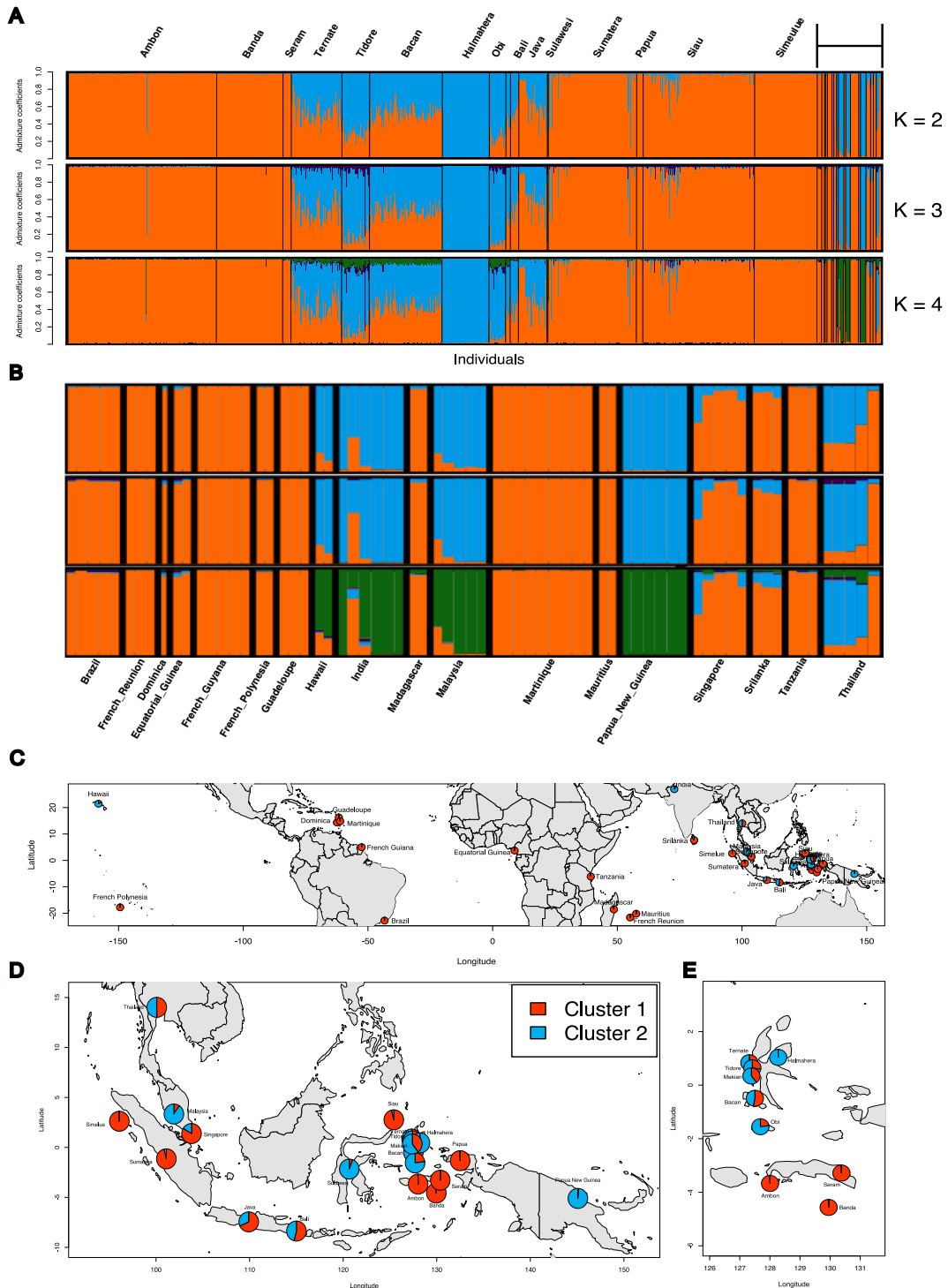


Figure 2. Population structure in all populations using Bayesian clustering inferred in fastSTRUCTURE based on nrDNA markers. Delta K value according to Evanno (2005) method indicating $K = 2$ is the most favorable in both level of analysis (E). Barplot of individual admixture proportions for each individual inferred by fastSTRUCTURE for all populations (A). Geographic projection of $K = 2$ of all populations (B), Southeast Asia (C), and the Moluccas (D).

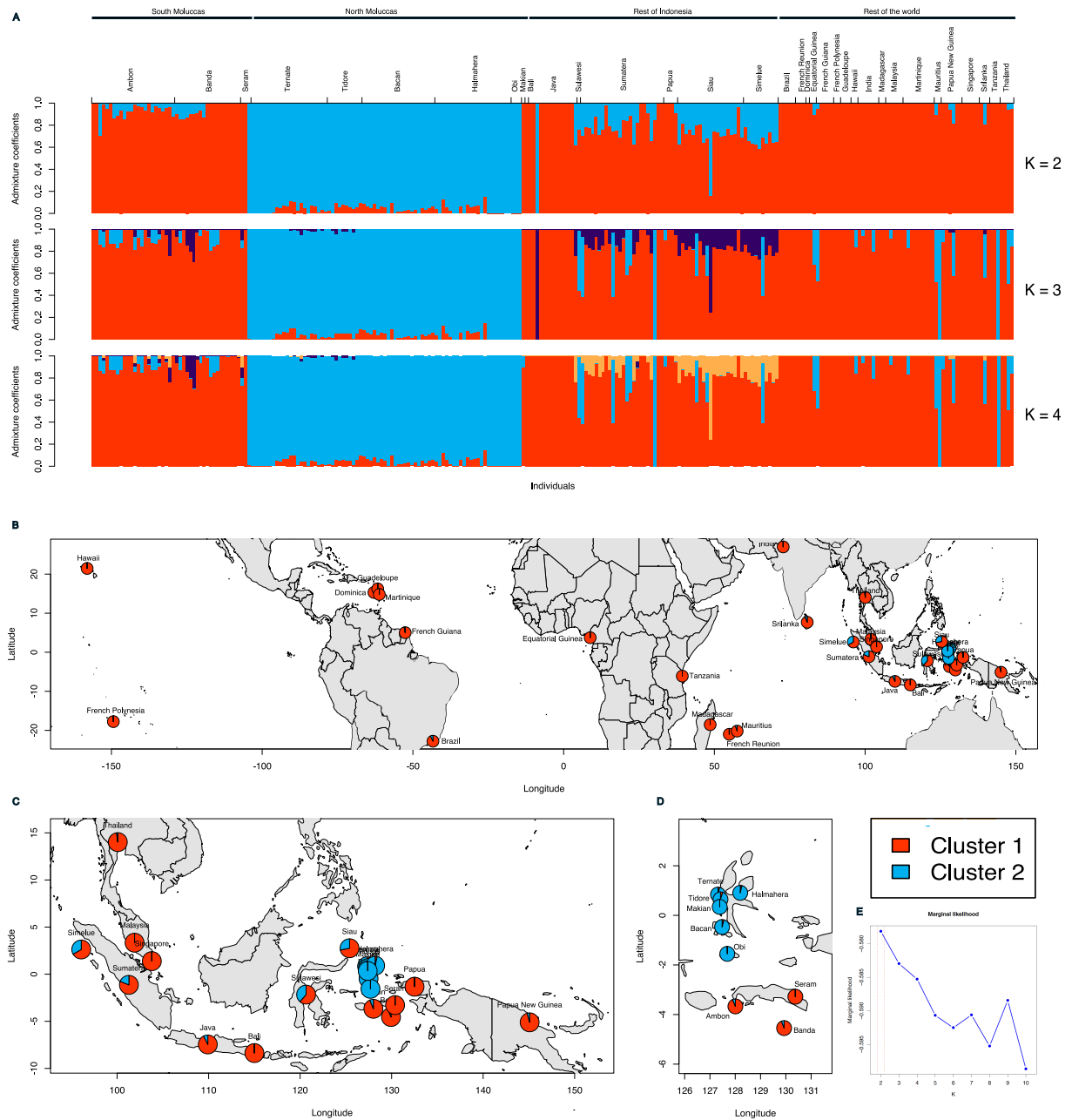
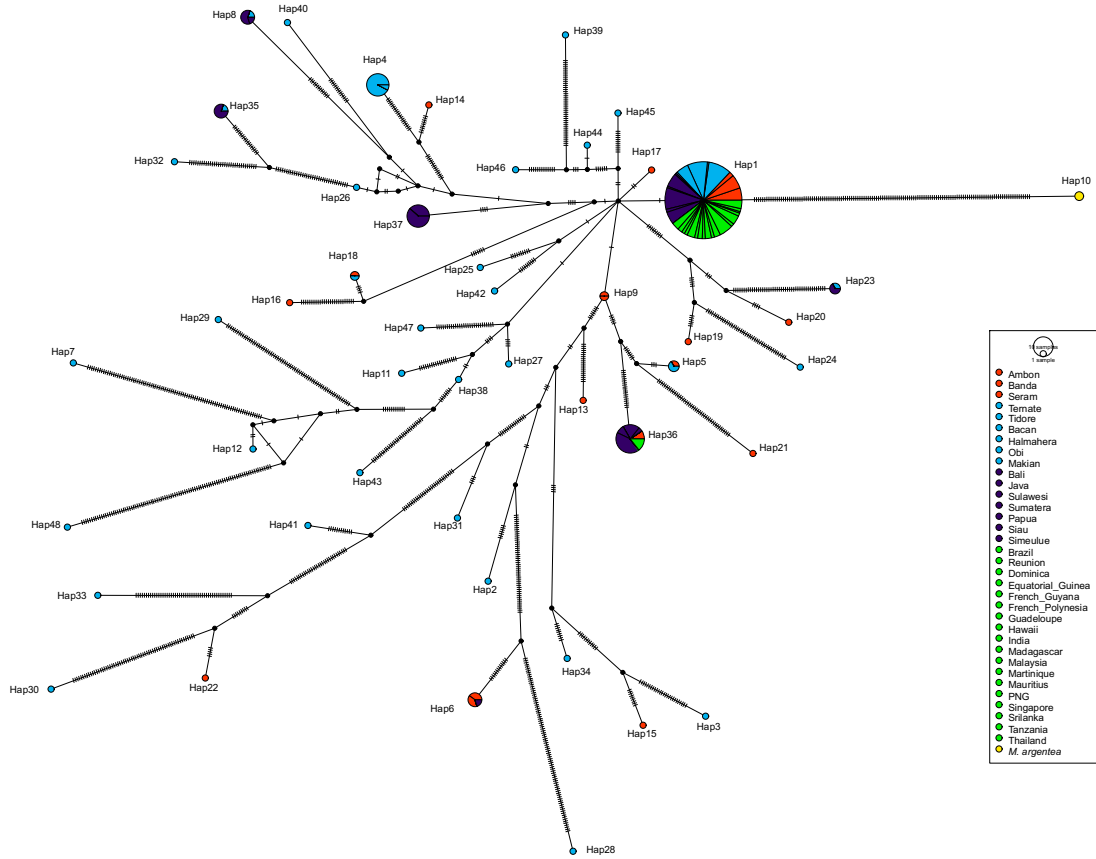


Figure 3. A maximum parsimonious network was constructed based on the cpDNA haplotypes, with each circle representing a unique haplotype, colored according to their populations of origin.



Supplementary Online Materials

Figure S1. Pairwise F_{ST} in all populations inferred by microsatellites. Values are visualized in a heatmap plot (A) and a dendrogram tree (B).

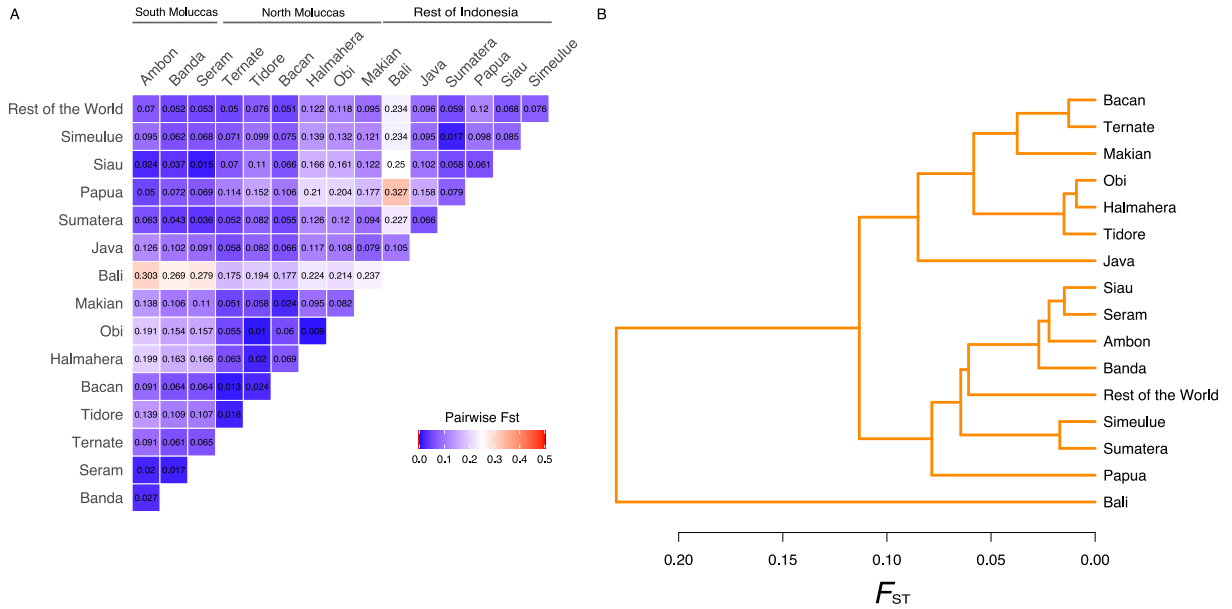


Figure S2. Pairwise F_{ST} in all populations inferred by nrDNA. Values are visualized in a heatmap plot (A) and a dendrogram tree (B).

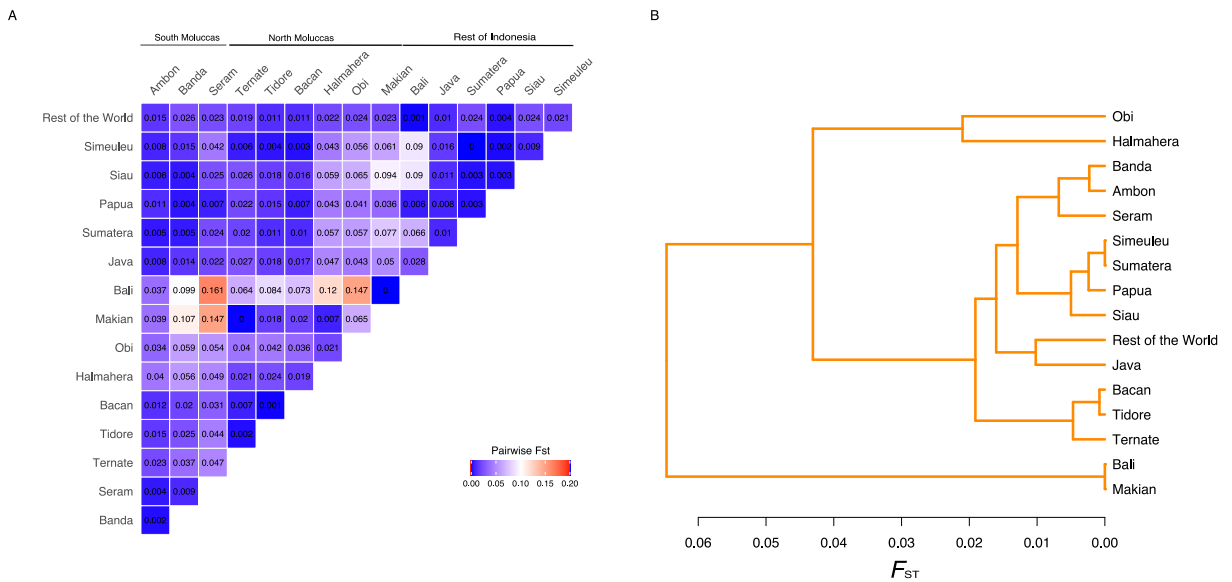
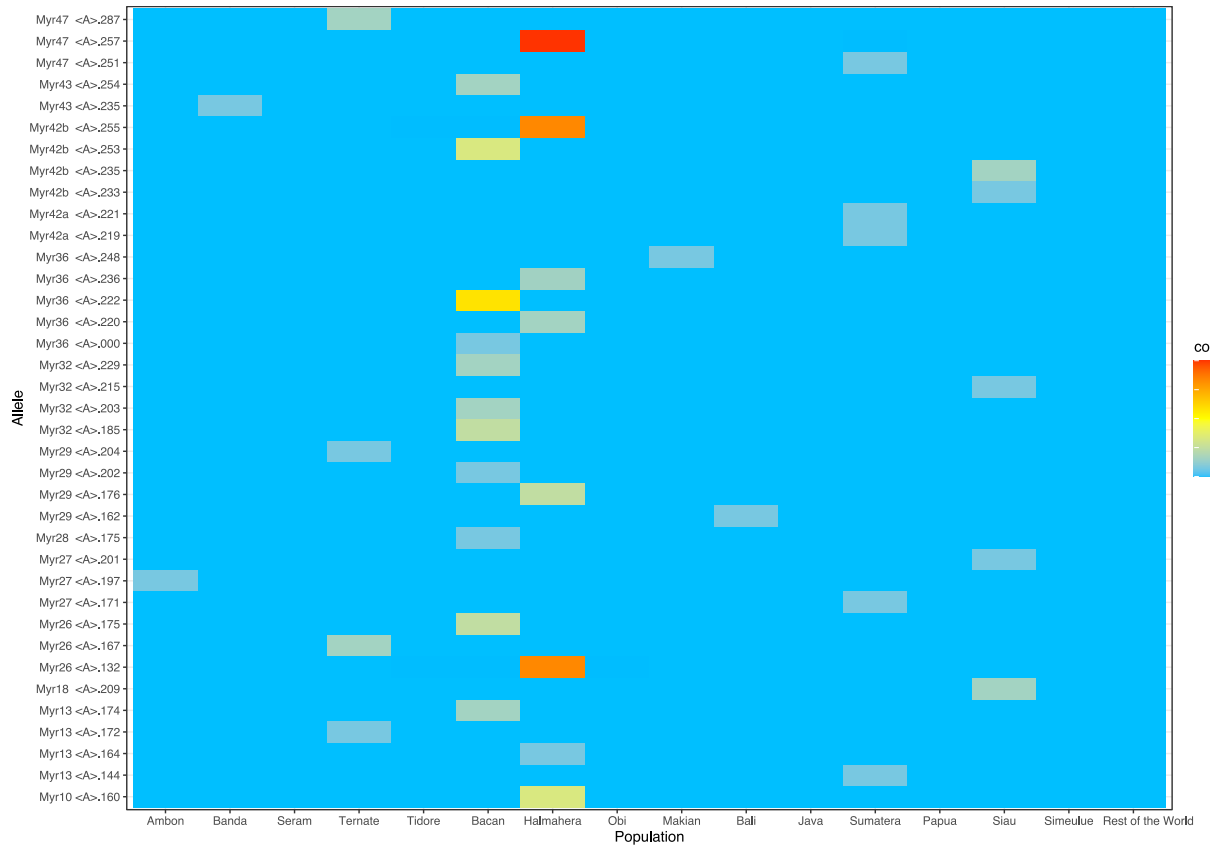


Figure S3. Population-level distribution of private alleles. Blue color represents populations without any private alleles, while the gradient color ranging from light blue to red indicates the individual count of private alleles.

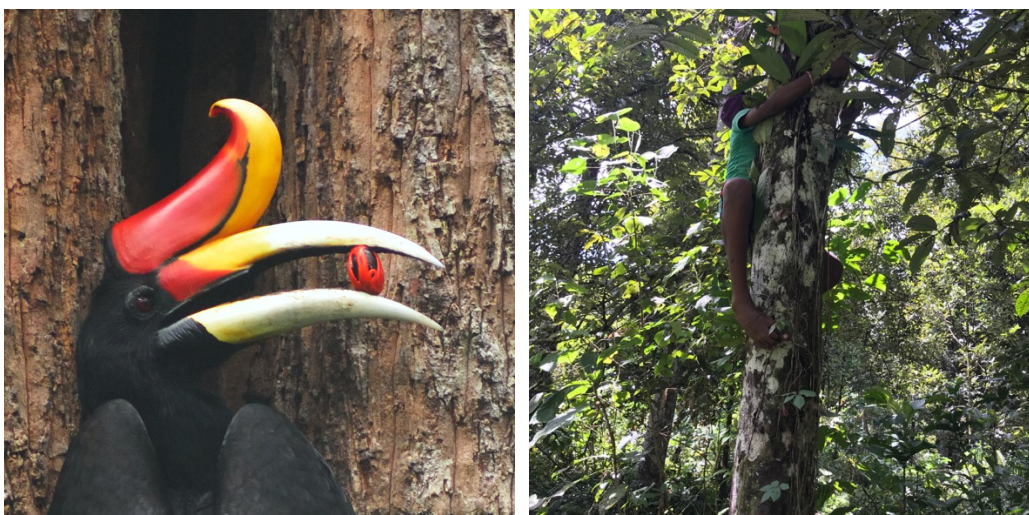


Supplementary Files

Supplementary File S1. The following is a comprehensive inventory of the samples utilized in this study, encompassing species information, geographic coordinates, population details, as well as their application for SSR and cpDNA data analysis. However, due to the longevity of the list and save some paper space for sustainability, we upload them in our GitHub page, which easily accessible in this following link: <https://github.com/jaktykusuma/myristica/tree/master/fragrans>.

CHAPTER 5

Influence of past climate changes and human-mediated introduction on patterns of genetic diversity distribution of the Papuan nutmeg (*Myristica argentea* Warb)



The image on the left showcases the prominent role of **Hornbills** (Bucerotidae) as the primary dispersers of Papuan nutmeg. Within the Fakfak Peninsula, Papua, Indonesia, local communities strictly enforce traditional sanctions against anyone who dares to disrupt the food sources of Hornbills, including Papuan nutmeg. This species thrives abundantly in the forests of Fakfak. In the picture on the right, an individual can be seen collecting leaves and flowers for the purpose of analysis in this study. In certain tribal villages, it is customary to have dozens of Papuan nutmeg trees cultivated in the backyard for local consumption. The propagation materials, specifically the seeds, are obtained from the forests surrounding the villages. The local communities have a diverse array of names for this species, often based on factors such as elevation, surroundings, and even ancestral connections. The image of the Hornbill was captured by Cede Prudente in 2015, while the photograph depicting the Papuan nutmeg forest was taken by Jaky Kusuma in 2019.

Chapter 5

Influence of past climate changes and human-mediated introduction on patterns of genetic diversity distribution of the Papuan nutmeg (*Myristica argentea*) Warb.

(In preparation)

Jakty Kusuma^{1,2}, Nora Scarcelli¹, Victoria Malaval¹, Marie Couderc¹, Jérôme Duminil¹

¹ DIADE, Univ. Montpellier, IRD, CIRAD, Montpellier, France.

² Politeknik Negeri Lampung, Bandar Lampung, Indonesia.

ABSTRACT

Understanding the distribution of genetic diversity of endemic food tree species holds significant importance for conservation planning and the formulation of effective management strategies. Natural and anthropogenic factors can influence species' geographic range and distribution of genetic diversity. Here, we tested the relative influence of barriers to gene flow (rivers, mountains and past climate change) and recent human activities on the evolutionary history of an endemic species from Fakfak district, Papua Island, *Myristica argentea* Warb (Myristicaceae), better known as Papuan nutmeg. *Myristica argentea* is a spice tree species used by local populations since centuries/millennia. The species has been introduced in other locations from the region. Fourteen microsatellite markers (nSSRs), whole chloroplast (cp) genome, and nuclear-targeted sequence genes (nrDNA sequence) were used to characterize the genetic diversity, population structure, and demographic history of the species across its distribution range (native and introduced populations). We found that native populations of the species in Fakfak present higher levels of genetic diversity than introduced populations. We further detected the presence of a moderate genetic structure ($F_{ST} = 0.143$), and the presence of three intraspecific genetic clusters. Additionally, our findings indicate that population fragmentation within the native range of the species in Fakfak, may have been influenced by the presence of riparian networks, mountain ranges and by Pleistocene climatic fluctuations. Our results suggest that all introduced populations were sourced mainly from the native population of Raduria, in the South of Fakfak region. Given the population genetic variation

across the ranges of Papuan nutmeg, management plans should not treat them as single populations, but rather consider the broader genetic diversity within its range.

Key words: Papuan nutmeg, glacial refugia hypothesis, phylogeography, demographic inferences

INTRODUCTION

Islands offer a distinct narrative of species' genetic variation and diversification when compared to their continental counterparts. Despite numerous studies that have assessed evolutionary history at a continental level (Bell et al., 2018; Couvreur et al., 2021; Graham, 2018; Helmstetter et al., 2020; Lu et al., 2018), our understanding of evolutionary dynamics at the island level remains scarce. This knowledge gap is noteworthy considering the pivotal role that islands play in biodiversity, with an enormous contribution to global diversity (Condamine et al., 2017).

The Indonesian archipelago is part of the Malesian biogeographic region and consists of ~17,000 islands (Raes & Van Welzen, 2009; von Rintelen et al., 2017). It is a highly biodiverse region with various climate and unique dynamic geotectonic activities (Joyce et al., 2020), which are drivers of plant diversity. Papua Island (internationally known as New Guinea Island), situated in the eastern part of Indonesia, is renowned for being home to one of the most diverse and abundant flora in the world (Cámara-Leret et al., 2020). An estimated 90% of the tree species found on Papua Island are unique and endemic to the region, showcasing its exceptional biodiversity and ecological significance (Barstow et al., 2022). Thus, Papua Island is home to species rich genera that contain more than 80 species per genus (Cámara-Leret et al., 2020)

The high level of tree species in Papua Island has been attributed to a highly variable topography, a complex geological history and steep environmental gradients (Morley, 2000; Slik et al., 2015), as well as its historical prints of glaciation (Wallace, 1878). It has been argued that during the glacial event (Last

Glacial Maximum), many forest-restricted species persisted in localized regions, referred to as forest refugia, where relatively favorable climatic conditions allowed for their survival (Blatrix et al., 2013; Hewitt, 1996; Pouteau et al., 2015). These forest refugia acted as reservoirs of genetic diversity and, as the climate warmed during the Holocene, these populations expanded and recolonized other areas (Hewitt, 2000).

As a result, the populations that survived in these refugia might have experienced a drastic change in population size as well as a rupture of gene flow among refugia. This is expected to lead to population differentiation, as geographic barriers are commonly recognized as a key mechanism driving population divergence. Mountains, along with rivers, are one of the main topographic factors associated with long-term barriers to gene flow (Khanal et al., 2018; Sánchez-Montes et al., 2018). As the Papua Island topography is highly complex (Baldwin et al., 2012; Toussaint et al., 2021), it is thus interesting to test the forest refugia hypothesis, which remains poorly documented.

Papua Island is home to the highest number of *Myristica* species (de Wilde, 2014), which suggest that this area is the center of origin and diversification of the genus. *Myristica argentea* Warb.—known as Papuan nutmeg, Fakfak nutmeg, or Macassar nutmeg—is one of the most economically and ethnobotanically recognized species from Papua Island. The species is natively distributed in West Papua, Fakfak peninsula (de Wilde, 2014). It is commonly used as a source of nutmeg and mace in the same ways as the Banda nutmeg (*Myristica fragrans*).

There is evidence that the natural distribution of *M. argentea* is confined to a limited geographic range (de Wilde, 1995, 2014), and it is commonly observed that

species with restricted distributions tend to display a low levels of genetic diversity (DeJoode & Wendel, 1992; Frankham, 1996, 1997). The restricted genetic diversity observed in narrowly distributed island taxa, as demonstrated by previous meta-analyses (González et al., 2020), is thought to be influenced by various factors that remain a subject of debate (Frankham, 1997). Some proposed explanations include the effect of limited dispersal abilities (Hindley et al., 2018) and small population sizes (Barret, 1996; Frankham, 1996). However, several examples explained that insular species with endemic status exhibit higher genetic diversity compared with their continental lineages (García-Verdugo et al., 2015; Salmona et al., 2023).

Furthermore, it is also noteworthy that anthropogenic activities can influence the distribution range and the level of genetic diversity of species. For instance, *M. argentea* has been introduced to neighboring islands outside its native range due to its similarity to *M. fragrans*. This led to the expansion of Papuan nutmeg's distribution range from the Fakfak district to nearby islands such as Ambon and Seram in the Moluccas archipelago. Therefore, it is interesting to investigate the levels of genetic diversity within both the restricted native range and the introduced populations of this species.

Using above-mentioned background, our objectives here were: (i) to investigate the influence of topographic (mountains, rivers) and past climate changes on patterns of Papuan nutmeg evolutionary history by characterizing the distribution of genetic diversity in its native range; (ii) to investigate the influence of human-mediated dispersal on species' genetic structure by characterizing the distribution of genetic diversity of introduced populations of Papuan nutmeg.

MATERIALS AND METHODS

Study area, plant materials and geographic sampling

Myristica argentea is native from Fakfak region, Papua Island and primarily grows in forests below 250 meters above sea level (Musaad et al., 2016; Wahyuni & Bermawie, 2020; de Wilde, 2014). The island of Papua itself is a relatively young landmass that has emerged and developed over the last 25 - 2 million years (Evans, 2020), during which it was colonized by numerous species (Oliver et al., 2017; Toussaint et al., 2014). We collected in the field 468 individuals of *M. argentea* (Figure 1 and Supplementary file S1). These samples originated from fifteen populations in Papua Island (thirteen from Fakfak Peninsula, and one from Manokwari and Raja Ampat) and two populations from South Moluccas archipelago (Ambon and Seram Island). The thirteen populations from Fakfak peninsula are located in the region of origin of the species (called ‘native’ populations hereafter), whereas additional populations (Manokwari, Raja Ampat, Seram and Ambon) correspond to human-driven introductions outside of species original distribution (called ‘introduced’ populations hereafter). Interviews with farmers informed us that native populations are composed of a mix of wild individuals, that were kept during land cleaning (the oldest individuals of the population), and planted individuals, originated from seeds of wild individuals from the same location.

Molecular genetic analysis

DNA extraction and nuclear microsatellite markers. For each sample, young fresh leaves were collected and immediately dried in silica gel. DNA was extracted

from dried leaves using the protocol of Mariac *et al.* (2006). Samples were genotyped using 14 nuclear microsatellite markers (nSSRs) following PCR conditions as described in Kusuma *et al.* (2020). Two of the primer pairs (*Myr28* and *Myr42*) amplified two different genomic regions, so we split them into four different markers (namely *Myr28a* and *Myr28b*, *Myr42a* and *Myr42b*). Genotyping was done on an ABI 3500 XL and analyzed using Geneious Prime® 2023.0.4 (2012a), with manual corrections.

Chloroplast DNA marker. A subsample of 43 *M. argentea* individuals (3 to 5 individuals per population) and three individuals of the sister species *M. fragrans*, used as outgroups, were sequenced ($N_{\text{TOTAL}} = 46$; Supplementary File S1). Whole genome library constructions were done following Mariac *et al.* (2014). Then raw data were analyzed following the method explained in Scarcelli *et al.* (2016). Briefly, we cleaned and filtered the reads using DEMULTADAPT (<https://github.com/Maillol/demultadapt>) and CUTADAPT 1.2.1 (Martin, 2011), then we mapped them with BWA (Li & Durbin, 2009) using the *M. argentea* cpDNA genome as reference (Kusuma *et al.*, 2023; GenBank accession: OP866724.1). We called SNPs using VARSCAN ver. 2.3.7 (Koboldt *et al.*, 2012) and filtered them using VCFtools (Danecek *et al.*, 2011). All scripts used are described in Supplementary File S2.

Nuclear targeted-sequence marker. We used the same individuals as for cpDNA analysis ($N = 46$, see Supplementary File S1 for details). Genomic libraries were prepared in ARCAD facilities (Montpellier, France). We then used targeted-capture strategies using a baiting kit of 469 nuclear genes, initially developed for a sister family of Myristicaceae, the Annonaceae (Couvreur *et al.*, 2019). We then

processed the DNA sequences of the 46 individuals with the pipeline SECAPR v.1.14 (Andermann et al., 2018). This method is beneficial for identifying SNPs because it generates a pseudo-reference made from the target regions' sequences, except for the paralog regions, which were identified and taken out automatically. BWA v.0.7.12 (Li & Durbin, 2009) was used for mapping and GATK v.4 (McKenna et al., 2010) was used to remove duplicates and to call SNPs. We filtered the data by keeping only biallelic SNPs with a mapping quality >40%, depth >25, and quality by depth >2 using BCFtools (Li, 2011). Additionally, we removed SNPs with allele frequency <0.01 and excluded monomorphic sites. Scripts used for this analysis are available in <https://github.com/jaktykusuma/myristica/tree/master/argentea>.

Genetic diversity and population structure

Being highly polymorphic, nSSRs are markers of choice to assess the distribution of the genetic diversity within populations. We characterized the levels of genetic diversity within population ($N_{POP} = 17$) using SPAGeDI 1.5 (Hardy & Vekemans, 2002) with the following indices: number of alleles (N_A), observed and expected heterozygosity (H_O and H_E , respectively), and the inbreeding coefficient (F_{IS}). We calculated the rarefied allelic richness (A_R) using R (R Core Team, 2022) package 'hierfstat' (Goudet, 2005). Presence of null alleles (F_{null}) and the F_{IS} corrected for the presence of null alleles were calculated using INEST 2.2 (Chybicki & Burczyk, 2009). Each diversity index was compared among the 26 populations using ANOVA. Levels of genetic diversity was also estimated for each population using the cpDNA dataset. Thus, we determined the number of haplotypes (H),

nucleotide diversity (π), number of segregating sites (S), and number of parsimony-informative sites per population with DnaSP 6.12.03 (Librado & Rozas, 2009).

Using SSR, we characterized the population structure by getting estimates of pairwise F_{ST} between populations using SPAGeDI 1.5 (Hardy & Vekemans, 2002), and by applying a Bayesian clustering approach as implemented in STRUCTURE (Pritchard et al., 2000). In addition, we conducted a population structure analysis using the nrDNA sequence dataset using fastStructure, and used the resulting genetic groups in subsequent demographic analyses. We also visualized the relationships and connections between cpDNA haplotypes within and among populations by reconstructing an haplotype network using the maximum parsimonious tree method (Kannan & Wheeler, 2012) as implemented in PopART (Leigh & Bryant, 2015), with a 95% statistical parsimony criterion.

Inference of population size

In order to test the presence of demographic changes in relation to past climate oscillations, we used the nrDNA sequences data to infer the evolution of native population size through time. The nrDNA sequences are adequate to conduct this analysis given their high potential to discriminate different evolutionary scenarios (Cornuet et al., 2010) and to improve the accuracy of parameters estimations (Smith & Flaxman, 2020). Using the nrDNA sequences, we utilized the Stairway Plot 2 approach (Liu & Fu, 2020), a model-flexible method that employs site-frequency spectra (SFS), to determine changes in effective population size (N_E) over time within each genetic cluster. To ensure the accuracy of our results, we did not apply minor-allele frequency filter, as it could potentially bias the SFS. Since

Stairway Plot relies on SNP counts for N_E inference, the exclusion of SNPs with missing data may introduce count biases. To mitigate this, we calculated the SFS for each intra-specific cluster observed within the native distribution of the species (as revealed by the Bayesian clustering analysis conducted on the nrDNA sequences dataset) following the methodology of Burgarella *et al.* (2018), i.e. by considering the minor-allele frequency at each SNP, multiplied by the mean number of haploid samples. This resulted in a new SFS that incorporated all observed site frequencies while minimizing SNP removal. We utilized two-thirds of the sites and performed 200 bootstraps for robustness. The number of observed sites was determined as the total length of the SECAPR pseudo-reference (see above) for each population, ensuring comprehensive coverage for our analysis. We applied a time calibration by using the Angiosperm wide mutation rate of $5.35e^{-9}$ mutations per site per year (De La Torre *et al.*, 2017) and a generation time of 50 years.

Population history

As an additional test of the influence of past-climate events, we inferred the ancestry and divergence time of the genetic clusters from nrDNA data, defined by fastStructure results, using DIYABC Random Forest 1.1.1 (Collin *et al.*, 2021). The populations assigned to a given genetic cluster according to STRUCTURE results were pooled as a single genetic group to reduce demographic model complexity. We only kept individuals with a cluster assignment probability higher than 0.85.

We tested different scenarios (Supplementary Figure S1) of population divergence using various historical demographic parameters (Supplementary Table S1), without considering any admixture event. Coalescent simulations were

performed for each scenario, with a total of 100,000 simulations. To assess the agreement between simulated and observed data, Principal Component Analysis (PCA) was conducted on the summary statistics from simulated data and the observed data. The study employed Approximate Bayesian Computation (ABC) with random forest (RF) in ‘*abcranger*’ module to perform scenario choice and parameter estimation (Pudlo et al., 2016; Raynal et al., 2019). The best scenarios were selected based on a classification vote, and RFs of 1,000 trees were used to estimate median point estimates and a 95% posterior credible interval for each demographic parameter.

RESULTS

Genetic diversity

In the nSSR dataset, we identified two loci (*Myr12* and *Myr26*) that had more than 40% of missing data and 53 individuals had more than 30% of missing data at different populations (Supplementary Figure S2). The corresponding loci and individuals were removed, resulting in a dataset of 420 individuals on 14 nSSR diploid genotypes. The 14 nSSRs revealed a total of 137 alleles, with a mean number of alleles per population ranging from 2.57 (Raja Ampat) to 6.57 (Mamur) (see Table 1 for detailed values). Using allelic richness as an indicator of genetic diversity, we found that half of the populations within the species’ native range exhibited significantly higher levels when compared to the introduced population (Table 1). Specifically, a significantly higher allelic richness across loci was observed in Perwasak (5.58), Mamur (3.16), Pawada (3.40), Kokas (3.39), Nika (3.36), Ubadari (3.52), and Patimburak (3.31) compared to Ambon (3.09), Raja

Ampat (2.08), Manokwari (2.43), and Seram (3.03) (ANOVA: $P < 0.01$). This finding suggests a greater genetic diversity in most of the native populations when compared to introduced populations. We found low levels of genetic diversity in Raja Ampat and Manokwari, as indicated by the rarefied allelic richness ($AR_{(k=7)}$; Table 1).

Population structure

Pairwise F_{ST} between populations ranged from 0.008 (between Ambon and Seram) to 0.327 (between Raja Ampat and Perwasak; Supplementary Figure S3). Furthermore, the Bayesian clustering analysis for all populations revealed the presence of three distinct genetic clusters ($K = 3$; Figure 2). Based on individual assignment scores (> 0.85), we allocated each population to its corresponding genetic cluster. The first genetic cluster comprised Ambon, Raja Ampat, Manokwari, and Raduria populations. The second genetic cluster consisted of Batuplat and Mabuni-buni populations. The third genetic cluster included Mamur, Kramongmongga, Pawada, Kokas, Nika, Ubadari, and Patimburak populations. Mbahambandara, Purwahabtonngo, and Perwasak populations were considered as admixed populations. Based on these clustering, the native populations were assigned to three different geographically-structured genetic clusters. We found that introduced populations were grouped together with the native population of Raduria. We found a similar structuration pattern with the nrDNA sequence data (Supplementary Figure S4).

We detected 405 SNP, corresponding to twelve different cpDNA haplotypes within our *M. argentea* sampling (Figure 3). All native populations were fixed for

one given haplotype, except for the Batuplat population, which displayed two distinct haplotypes (*Hap4* and *Hap12*). Within native populations, one shared haplotype (*Hap1*) was present in five populations of Fakfak Peninsula (Figure 3A). The distribution of the haplotypes showed that the two haplotypes in Ambon (*Hap4* and *Hap7*) and Seram (*Hap7* and *Hap10*) were shared with the population from its native range, which includes Batuplat (*Hap4*), Raduria (*Hap7*), and Kokas (*Hap10*).

Inference of population size

We observed three intra-specific clusters using the Bayesian clustering analysis conducted on the nrDNA sequences dataset (Figure 4): Cluster 1 included Mbahambandara, Mabuni-buni, and Batuplat populations ($N_{INDV} = 5$), Cluster 2 included Mamur, Kramongmongga, Pawada, Kokas, Nika, Ubadari, and Patimburak populations ($N_{INDV} = 17$) and Cluster 3 included Raduria, Perwasak and Purwahabtonngo populations ($N_{INDV} = 9$). The stairway plot analysis showed that the first two clusters underwent an expansion approximately 500,000 years ago (Figure 4A and 4B). Following this expansion, Cluster 1 exhibited a constant population size, as evidenced by the upper and lower bounds of the 95% confidence interval. Although the statistical support for this observation was not particularly strong ($P > 0.005$), the historical constant size persisted. In contrast, Cluster 2 experienced multiple pulses of population change, with a significant decline starting from approximately ~20,000 to ~7,500 years ago. However, after this decline, the population size remained constant. Regarding Cluster 3 (Figure 4C), a severe population decline was detected around ~100,000 years ago, followed by a

sudden expansion around ~50,000 years ago. Subsequently, the population size stabilized and remained constant.

Population history

The PCA plots of both observed and simulated data were visually inspected (see Supplementary Figure S5). Our findings suggested that the considered models can accurately replicate the observed patterns of genetic diversity. Therefore, we deemed the proposed scenarios and prior combinations sufficient for conducting ABC analysis. Model selection according to DIYABC-RF favors scenario 1 as the best supported scenario (Figure 5), with a posterior probability of 0.783 and prior error of 0.35 and a Bayes factors of 2.889. This scenario suggested that all populations splat at the same time from a single ancestral population. The divergence time was estimated at 201 to 417 generations ago, with median of 276 generations ago. Assuming 50 years of generation time, the split would fall around 10,500 to 20,850 years ago.

DISCUSSION

Influence of topographical barriers and past climate change on population dynamics and patterns of genetic diversity distribution

The genetic diversity within *M. argentea* is not homogeneously distributed throughout its range as demonstrated by all three types of genetic markers (nSSRs, cpDNA, and nrDNA). Our study further supported that certain populations within its native range exhibited higher levels of genetic diversity compared to others.

Three main geographically-structured genetic clusters were observed within the native range of the species. The three genetic clusters corresponded to the following groupings of populations: South of Fakfak (pops 6), North East of Fakfak (pops 2 and 3), and North West of Fakfak (all other populations). This subdivision is further supported by the cpDNA data, and we hypothesized that these three genetic clusters might correspond to the joint action of topographic barriers and past climatic changes.

The region of Fakfak is crossed by the Fakfak Mountains (*Pegunungan Fakfak*) from East to West that represent a strong barrier to gene flow within the distribution of *M. argentea*. We observed a genetic differentiation between the North and South populations of the region, which probably support the role of Fakfak Mountains as a barrier to gene flow. Additionally, within the northern area of the Fakfak region, we have identified genetic variations. This specific region encompasses the populations of Batuplat and Mabuni-buni in the eastern part, as well as Mamur, Kramongmongga, Pawada, Kokas, Nika, Ubadari, and Patimburak in the western part. Notably, in this region, there is a notable riparian network (Kayuni river network) which serves as a geographical midpoint between the two distinct genetic clusters. While the riverine barrier hypothesis has been extensively studied to assess genetic differentiation among populations (Gascon et al., 2000) and confirmed in the context of the Congo river network (Voelker et al., 2013), we did not anticipate the river to act as a barrier in this particular case. This can be attributed to the fact that seed dispersal by Bucerotidae, which is the primary mode of dispersal for Myristicaceae (BBKSDA, 2022), is not solely impeded by the presence of a river. Consequently, we speculate that the Kayuni river network may

have exerted a combined influence on gene flow patterns, alongside past climate fluctuations.

The historical climate fluctuations were characterized by alternating periods of cooler climates, which significantly influenced the distribution and composition of forest refugia (De Frenne et al., 2021). During colder periods, the forest refugia likely contracted and became fragmented, isolating populations and promoting genetic differentiation (Connor, 1986). This situation might reflect the population structure in the two genetic clusters observed at the northern region of Fakfak Peninsula.

The influence of climatic events were also observed on the distinct evolutionary histories in the three genetic clusters within the species' native range. The population inhabiting the northern coast of the Fakfak Peninsula exhibited a pattern of population expansion followed by a period of constant size. In contrast, the population in the southern coast of the peninsula displayed a trend of population decline. This demographic divergence coincides with the contrasting vegetation and topography between the north and south regions of the peninsula (Hill & Gleadow, 1989 ; Segar et al., 2017). The northern part is characterized by relatively flat terrain and dominated by swamp lands, whereas the southern part features hilly landscapes predominantly composed of karst limestone (Balazs, 1968 ; Baldwin et al., 2012 ; Hill & Gleadow, 1989; Toussaint et al., 2021).

We found that the divergence of the three different clusters occurred approximately at the same time between ~10,500 to ~20,850 years ago. This timing of divergence is strikingly much more recent than the timing of demographic changes previously observed. The obtained divergence timings among clusters

correspond to the climatic activities documented during the Late Pleistocene, specifically in Papua Island (Fairbairn et al., 2006). This period was characterized by the occurrence of multiple microclimatic conditions, as evidenced by sedimentation records (Lelono, 2008; Peng et al., 2021). These microclimatic shifts played a crucial role in shaping the vegetation formation in the mountainous range of Papua Island. Over time, the fluctuations in temperature and precipitation patterns influenced the distribution of plant species and impacted the composition of forest types found in the mountains. The dynamic nature of these climatic changes contributed to the rich biodiversity and distinct vegetation patterns observed in the region. Furthermore, the Fakfak Peninsula's location, surrounded by seas on the north and south coasts with different monsoonal conditions (Prentice et al., 2007), may have contributed to the distinct evolutionary histories observed in the different populations of *M. argentea*.

Given the paucity of phylogeographic studies done in the region, we cannot compare our findings with other species. The development of a comparative phylogeographic framework would help investigating if similar patterns of genetic diversity distribution are found in other species, as well as additional studies on the demographic evolution of the species through time and of the timing of divergence of intra-specific clusters.

History of introduction of the species

Our findings provide compelling evidence on the geographic origin of introduced populations in other regions of Papua Island (Manokwari and Raja Ampat) as well as the Moluccas (Ambon and Seram). The nuclear genetic analysis

demonstrates that the introduced populations share the same genetic profile as some native populations, specifically Raduria, which is located in the southern coast of Fakfak.

Slightly different patterns were observed when examining the plastid compartment. In particular, the population in Manokwari did not share any haplotypes with the native range, whereas Ambon exhibited haplotype sharing with Batuplat and Raduria. Similarly, Seram displayed haplotypes shared with Raduria and Kokas. Overall, our results suggest that the Raduria population may have served as a main source of propagation material.

The observed pattern can be attributed to the migration of people from South Fakfak to various other regions of the island (Upton, 2009). The migration patterns discussed are strongly linked to the phenomenon of urbanization (Crosbie Walsh, 1987). The movement of tribal communities from rural areas in Fakfak to different regions of the island is primarily motivated by job prospects in Manokwari. The recent establishment of Manokwari as the capital of West Papua (Ronsumbre & Huwae, 2022), along with the growth of tourism in Raja Ampat (King, 2018), further contributed to this migration trend. It is plausible to speculate that during the migration process, individuals may have carried propagation materials with them to their new locations, enabling them to plant and utilize the species. Furthermore, it is worth noting that this human-mediated dispersal phenomenon extends beyond *M. argentea* and encompasses other forest species as well (Suryawan, 2016).

Finally, the intentional transportation of propagation materials during the migration process indicates a deliberate effort to propagate and utilize various

species in new locations. This finding emphasizes the significance of human activities in shaping the distribution of plant species within the region. Further research in this area may shed light on the specific mechanisms and impacts of human-mediated dispersal on local plant diversity.

CONCLUSION

Our study revealed significant differences in genetic diversity between native and introduced populations of Papuan nutmeg. Specifically, several populations within the species' native range exhibited higher genetic diversity compared to the introduced populations. The observed genetic clustering within the native range was influenced by geographical barriers and past climate changes, which likely contributed to the divergence of the three genetic clusters approximately 10,500 to 20,850 years ago. This divergence was likely facilitated by the presence of multiple microclimatic conditions. Considering these findings, it is imperative to implement sustainable management strategies for the genetic resources of *M. argentea*. One approach is to safeguard the distinct populations or gene pools by establishing specific conservation areas. The conservation efforts by local communities in the Fakfak Peninsula, serve as a valuable example that can be expanded and applied on a larger scale by policymakers. It is important to emphasize that these conservation efforts also contribute to the maintenance of genetic resources for these forest species. By effectively communicating the benefits of conserving genetic diversity, we can further promote and enhance these conservation initiatives.

ACKNOWLEDGEMENT

The authors express their gratitude to the farmers and landowners of Ambon, Seram, Raja Ampat, Manokwari, and Fakfak for generously providing valuable information on the history of Papuan nutmeg and for their assistance in collecting leaf samples. Additionally, we extend our appreciation to our guide and Dina Tri Wahyuni for their valuable support during the field work. We also thank Roxali Bijmoer and Marnel Scherrenberg from Naturalis Biodiversity Center, Leiden, The Netherlands as well as Myriam Gaudeul from National Museum of Natural History, Paris, France for their invaluable help during the specimen collection. We further thank Pierre Mournet from CIRAD, Montpellier, France and Sylvain Santoni, Audrey Weber from INRAe, Montpellier, France for providing assistance during lab analysis. This study was authorized by Indonesian National Research and Innovation Agency (BRIN) with research permit to JD (144/SIP/FRP/E5/Ditk.KI/V/2018) for collecting samples in Ambon, and collecting permit to JK (B-3382/IPH.1/KS.02.04/IX/2019 and B-75/IV/KS.01.04/3/2022). We also greatly appreciate the support from the Indonesia Endowment Fund for Education (LPDP) to JK with grant No. 0003719/AFR/D/BUDI-2018. This project also supported by Agropolis Fondation under the reference ID 1502-503 and 1502-504 through the « *Investissements d'avenir* » program (Labex Agro: ANR-10-LABX-0001-01), under the frame of I-SITE MUSE (ANR-16-IDEX-0006). The authors acknowledge the ISO 9001 certified IRD i-Trop HPC (member of the South Green Platform) at IRD Montpellier for providing HPC resources that have contributed to the research results reported within this paper. URL: <https://bioinfo.ird.fr/>- <http://www.southgreen.fr>.

REFERENCES

- Andermann, T., Cano, Á., Zizka, A., Bacon, C., & Antonelli, A. (2018). SECAPR—a bioinformatics pipeline for the rapid and user-friendly processing of targeted enriched Illumina sequences, from raw reads to alignments. *PeerJ*, 6, e5175. <https://doi.org/10.7717/peerj.5175>
- Balazs, D. (1968). Karst regions in Indonesia. *Karszt-Es Barlangkutatas*, 5, 3–61.
- Baldwin, S. L., Fitzgerald, P. G., & Webb, L. E. (2012). Tectonics of the New Guinea Region. *Annual Review of Earth and Planetary Sciences*, 40(1), 495–520. <https://doi.org/10.1146/annurev-earth-040809-152540>
- Barret, S. C. H. (1996). The reproductive biology and genetics of island plants. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 351(1341), 725–733. <https://doi.org/10.1098/rstb.1996.0067>
- Barstow, M., Jimbo, T., & Davies, K. (2022). Extinction risk to the endemic trees of Papua New Guinea. *PLANTS, PEOPLE, PLANET*. <https://doi.org/10.1002/ppp3.10342>
- BBKSDA, P. B. (2022). *Globally recognized nutmeg from the east (Myristica argentea Warb).* in *Indonesian: Rempah dari timur yang telah mendunia: Pala Papua (Myristica argentea Warb).* Accessed in 20 May 2023 in <https://Bbksda-Papuabarat.Com/Rempah-Dari-Timur-Indonesia-Yang-Telah-Mendunia-Pala-Papua-Myristica-Argentea-Warb/>.
- Bell, N., Griffin, P. C., Hoffmann, A. A., & Miller, A. D. (2018). Spatial patterns of genetic diversity among Australian alpine flora communities revealed by comparative phylogenomics. *Journal of Biogeography*, 45(1), 177–189. <https://doi.org/10.1111/jbi.13120>
- Blatrix, R., McKey, D., & Born, C. (2013). Consequences of past climate change for species engaged in obligatory interactions. *Comptes Rendus Geoscience*, 345(7–8), 306–315. <https://doi.org/10.1016/j.crte.2013.03.006>
- Burgarella, C., Cubry, P., Kane, N. A., Varshney, R. K., Mariac, C., Liu, X., Shi, C., Thudi, M., Couderc, M., Xu, X., Chitikineni, A., Scarcelli, N., Barnaud, A., Rhoné, B., Dupuy, C., François, O., Berthouly-Salazar, C., & Vigouroux, Y. (2018). A western Sahara centre of domestication inferred from pearl millet genomes. *Nature Ecology & Evolution*, 2(9), 1377–1380. <https://doi.org/10.1038/s41559-018-0643-y>
- Cámara-Leret, R., Frodin, D. G., Adema, F., Anderson, C., Appelhans, M. S., Argent, G., Arias Guerrero, S., Ashton, P., Baker, W. J., Barfod, A. S., Barrington, D., Borosova, R., Bramley, G. L. C., Briggs, M., Buerki, S., Cahen, D., Callmander, M. W., Cheek, M., Chen, C.-W., ... van Welzen, P.

- C. (2020). New Guinea has the world's richest island flora. *Nature*, 584(7822), 579–583. <https://doi.org/10.1038/s41586-020-2549-5>
- Chybicki, I. J., & Buczynski, 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>
- Condamine, F. L., Leslie, A. B., & Antonelli, A. (2017). Ancient islands acted as refugia and pumps for conifer diversity. *Cladistics*, 33(1), 69–92. <https://doi.org/10.1111/cla.12155>
- Connor, E. F. (1986). The role of Pleistocene forest refugia in the evolution and biogeography of tropical biotas. *Trends in Ecology & Evolution*, 1(6), 165–168. [https://doi.org/10.1016/0169-5347\(86\)90047-9](https://doi.org/10.1016/0169-5347(86)90047-9)
- Cornuet, J.-M., Ravigné, V., & Estoup, A. (2010). Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1.0). *BMC Bioinformatics*, 11(1), 401. <https://doi.org/10.1186/1471-2105-11-401>
- Couvreur, T. L. P., Dauby, G., Blach-Overgaard, A., Deblauwe, V., Dessein, S., Droissart, V., Hardy, O. J., Harris, D. J., Janssens, S. B., Ley, A. C., Mackinder, B. A., Sonké, B., Sosef, M. S. M., Stévant, T., Svenning, J., Wieringa, J. J., Faye, A., Missou, A. D., Tolley, K. A., ... Sepulchre, P. (2021). Tectonics, climate and the diversification of the tropical African terrestrial flora and fauna. *Biological Reviews*, 96(1), 16–51. <https://doi.org/10.1111/brv.12644>
- Couvreur, T. L. P., Helmstetter, A. J., Koenen, E. J. M., Bethune, K., Brandão, R. D., Little, S. A., Sauquet, H., & Erkens, R. H. J. (2019). Phylogenomics of the Major Tropical Plant Family Annonaceae Using Targeted Enrichment of Nuclear Genes. *Frontiers in Plant Science*, 9. <https://doi.org/10.3389/fpls.2018.01941>
- 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>
- De Frenne, P., Lenoir, J., Luoto, M., Scheffers, B. R., Zellweger, F., Aalto, J., Ashcroft, M. B., Christiansen, D. M., Decocq, G., De Pauw, K., Govaert, S., Greiser, C., Gril, E., Hampe, A., Jucker, T., Klimes, D. H., Koelemeijer, I. A.,

- Lembrechts, J. J., Marrec, R., ... Hylander, K. (2021). Forest microclimates and climate change: Importance, drivers and future research agenda. *Global Change Biology*, 27(11), 2279–2297. <https://doi.org/10.1111/gcb.15569>
- De La Torre, A. R., Li, Z., Van de Peer, Y., & Ingvarsson, P. K. (2017). Contrasting Rates of Molecular Evolution and Patterns of Selection among Gymnosperms and Flowering Plants. *Molecular Biology and Evolution*, 34(6), 1363–1377. <https://doi.org/10.1093/molbev/msx069>
- de Wilde, W. J. J. O. (1995). Census of Myristica (Myristicaceae) in New Guinea anno 1994. *Blumea: Biodiversity, Evolution and Biogeography of Plants*, 40(2), 237–344.
- de Wilde, W. J. J. O. (2014). *Flora Malesiana*. Pensoft Publishers. <https://doi.org/10.3897/ab.e1141>
- DeJooe, D. R., & Wendel, J. F. (1992). Genetic diversity and origin of the Hawaiian islands cotton, *Gossypium tomentosum*. *American Journal of Botany*, 79(11), 1311–1319. <https://doi.org/10.1002/j.1537-2197.1992.tb13735.x>
- Evans, D. L. (2020). A Geophysical and Climatological Assessment of New Guinea — Implications for the Origins of *Saccharum*. *BioRxiv*, 2020.06.20.162842. <https://doi.org/10.1101/2020.06.20.162842>
- Fairbairn, A. S., Hope, G. S., & Summerhayes, G. R. (2006). Pleistocene Occupation of New Guinea's Highland and Subalpine Environments. *World Archaeology*, 38(3), 371–386. <http://www.jstor.org/stable/40026639>
- Frankham, R. (1996). Relationship of Genetic Variation to Population Size in Wildlife. *Conservation Biology*, 10(6), 1500–1508. <https://doi.org/10.1046/j.1523-1739.1996.10061500.x>
- Frankham, R. (1997). Do island populations have less genetic variation than mainland populations? *Heredity*, 78(3), 311–327. <https://doi.org/10.1038/hdy.1997.46>
- García-Verdugo, C., Sajeve, M., La Mantia, T., Harrouni, C., Msanda, F., & Caujapé-Castells, J. (2015). Do island plant populations really have lower genetic variation than mainland populations? Effects of selection and distribution range on genetic diversity estimates. *Molecular Ecology*, 24(4), 726–741. <https://doi.org/10.1111/mec.13060>
- Gascon, C., Malcolm, J. R., Patton, J. L., da Silva, M. N. F., Bogart, J. P., Lougheed, S. C., Peres, C. A., Neckel, S., & Boag, P. T. (2000). Riverine barriers and the geographic distribution of Amazonian species. *Proceedings of the National Academy of Sciences*, 97(25), 13672–13677. <https://doi.org/10.1073/pnas.230136397>

- González, A. V., Gómez-Silva, V., Ramírez, M. J., & Fontúrbel, F. E. (2020). Meta-analysis of the differential effects of habitat fragmentation and degradation on plant genetic diversity. *Conservation Biology*, 34(3), 711–720. <https://doi.org/10.1111/cobi.13422>
- 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>
- Graham, A. (2018). The role of land bridges, ancient environments, and migrations in the assembly of the North American flora. *Journal of Systematics and Evolution*, 56(5), 405–429. <https://doi.org/10.1111/jse.12302>
- 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>
- Helmstetter, A. J., Béthune, K., Kamdem, N. G., Sonké, B., & Couvreur, T. L. P. (2020). Individualistic evolutionary responses of Central African rain forest plants to Pleistocene climatic fluctuations. *Proceedings of the National Academy of Sciences*, 117(51), 32509–32518. <https://doi.org/10.1073/pnas.2001018117>
- Hewitt, G. (1996). Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, 58(3), 247–276. <https://doi.org/10.1006/bijl.1996.0035>
- Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, 405(6789), 907–913. <https://doi.org/10.1038/35016000>
- Hill, K. C., & Gleadow, A. J. W. (1989). Uplift and thermal history of the Papuan Fold Belt, Papua New Guinea: Apatite fission track analysis. *Australian Journal of Earth Sciences*, 36(4), 515–539. <https://doi.org/10.1080/08120098908729507>
- Hindley, J. A., Graham, B. A., Pulgarin-R., P. C., & Burg, T. M. (2018). The influence of latitude, geographic distance, and habitat discontinuities on genetic variation in a high latitude montane species. *Scientific Reports*, 8(1), 11846. <https://doi.org/10.1038/s41598-018-29982-7>
- Joyce, E., Thiele, K., Slik, F., & Crayn, D. (2020). Checklist of the vascular flora of the Sunda-Sahul Convergence Zone. *Biodiversity Data Journal*, 8. <https://doi.org/10.3897/BDJ.8.e51094>
- 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>
- Khanal, L., Chalise, M. K., Wan, T., & Jiang, X. (2018). Riverine barrier effects on population genetic structure of the Hanuman langur (*Semnopithecus entellus*) in the Nepal Himalaya. *BMC Evolutionary Biology*, 18(1), 159. <https://doi.org/10.1186/s12862-018-1280-4>
- King, C. (2018). Tourism in Raja Ampat: New Chances and Challenges. *Independent Study Project (ISP) Collection*, 2618.
- 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>
- 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>
- Leigh, J. W., & Bryant, D. (2015). popart: full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, 6(9), 1110–1116. <https://doi.org/10.1111/2041-210X.12410>
- Lelono, E. B. (2008). Pleistocene palynology of the Waipona Basin, Papua. *Scientific Contributions of Oil and Gas*, 2(31).
- Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, 27(21), 2987–2993. <https://doi.org/10.1093/bioinformatics/btr509>
- 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>
- 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, X., & Fu, Y.-X. (2020). Stairway Plot 2: demographic history inference with folded SNP frequency spectra. *Genome Biology*, 21(1), 280. <https://doi.org/10.1186/s13059-020-02196-9>
- Lu, L.-M., Mao, L.-F., Yang, T., Ye, J.-F., Liu, B., Li, H.-L., Sun, M., Miller, J. T., Mathews, S., Hu, H.-H., Niu, Y.-T., Peng, D.-X., Chen, Y.-H., Smith, S. A., Chen, M., Xiang, K.-L., Le, C.-T., Dang, V.-C., Lu, A.-M., ... Chen, Z.-D. (2018). Evolutionary history of the angiosperm flora of China. *Nature*, 554(7691), 234–238. <https://doi.org/10.1038/nature25485>
- Mariac, C., Luong, V., Kapran, I., Mamadou, A., Sagnard, F., Deu, M., Chantereau, J., Gerard, B., Ndjeunga, J., Bezançon, 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., Kougbéadjou, 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>
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, 20(9), 1297–1303. <https://doi.org/10.1101/gr.107524.110>
- Morley, R. J. (2000). *Origin and Evolution of Tropical Rain Forests* (1st ed., Vol. 1). Wiley. <https://doi.org/10.2307/3515620>
- Oliver, P. M., Iannella, A., Richards, S. J., & Lee, M. S. Y. (2017). Mountain colonisation, miniaturisation and ecological evolution in a radiation of direct-developing New Guinea Frogs (Choerophryne, Microhylidae). *PeerJ*, 5, e3077. <https://doi.org/10.7717/peerj.3077>
- Peng, N., Dang, H., Wu, J., Aiello, I. W., & Jian, Z. (2021). Tectonic and climatic controls on the Plio-Pleistocene evolution of sediment discharge from Papua New Guinea. *Marine Geology*, 441, 106627. <https://doi.org/10.1016/j.margeo.2021.106627>

- Pouteau, R., Trueba, S., Feild, T. S., & Isnard, S. (2015). New Caledonia: a Pleistocene refugium for rain forest lineages of relict angiosperms. *Journal of Biogeography*, 42(11), 2062–2077. <https://doi.org/10.1111/jbi.12581>
- Prentice, M. L., Hope, G. S., & others. (2007). Climate of Papua. *The Ecology of Papua*, 1, 479–494.
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, 155(2), 945 LP – 959. <http://www.genetics.org/content/155/2/945.abstract>
- Pudlo, P., Marin, J.-M., Estoup, A., Cornuet, J.-M., Gautier, M., & Robert, C. P. (2016). Reliable ABC model choice via random forests. *Bioinformatics*, 32(6), 859–866. <https://doi.org/10.1093/bioinformatics/btv684>
- R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- Raes, N., & Van Welzen, P. C. (2009). The demarcation and internal division of Flora Malesiana: 1857--present. *Blumea-Biodiversity, Evolution and Biogeography of Plants*, 54(1–2), 6–8.
- Raynal, L., Marin, J.-M., Pudlo, P., Ribatet, M., Robert, C. P., & Estoup, A. (2019). ABC random forests for Bayesian parameter inference. *Bioinformatics*, 35(10), 1720–1728. <https://doi.org/10.1093/bioinformatics/bty867>
- Ronsumbre, A., & Huwae, M. A. (2022). Identity, power, myth contestation and land control in Manokwari, West Papua- Indonesia. *World Journal of Advanced Research and Reviews*, 14(1), 223–228. <https://doi.org/10.30574/wjarr.2022.14.1.0318>
- Salmona, J., Dresen, A., Ranaivoson, A. E., Manzi, S., Le Pors, B., Hong-Wa, C., Razanatsoa, J., Andriaholinirina, N. V., Rasoloharijaona, S., Vavitsara, M., & Besnard, G. (2023). How ancient forest fragmentation and riparian connectivity generate high levels of genetic diversity in a microendemic Malagasy tree. *Molecular Ecology*, 32(2), 299–315. <https://doi.org/10.1111/mec.16759>
- Sánchez-Montes, G., Wang, J., Ariño, A. H., & Martínez-Solano, Í. (2018). Mountains as barriers to gene flow in amphibians: Quantifying the differential effect of a major mountain ridge on the genetic structure of four sympatric species with different life history traits. *Journal of Biogeography*, 45(2), 318–331. <https://doi.org/10.1111/jbi.13132>
- 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 NGS 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>
- Segar, S. T., Volf, M., Zima Jnr, J., Isua, B., Sisol, M., Sam, L., Sam, K., Souto-Vilarós, D., & Novotny, V. (2017). Speciation in a keystone plant genus is driven by elevation: a case study in New Guinean *Ficus*. *Journal of Evolutionary Biology*, 30(3), 512–523. <https://doi.org/10.1111/jeb.13020>
- Slik, J. W. F., Arroyo-Rodríguez, V., Aiba, S.-I., Alvarez-Loayza, P., Alves, L. F., Ashton, P., Balvanera, P., Bastian, M. L., Bellingham, P. J., van den Berg, E., Bernacci, L., da Conceição Bispo, P., Blanc, L., Böhning-Gaese, K., Boeckx, P., Bongers, F., Boyle, B., Bradford, M., Brearley, F. Q., ... Venticinque, E. M. (2015). An estimate of the number of tropical tree species. *Proceedings of the National Academy of Sciences*, 112(24), 7472–7477. <https://doi.org/10.1073/pnas.1423147112>
- Smith, C. C. R., & Flaxman, S. M. (2020). Leveraging whole genome sequencing data for demographic inference with approximate Bayesian computation. *Molecular Ecology Resources*, 20(1), 125–139. <https://doi.org/10.1111/1755-0998.13092>
- Suryawan, I. N. (2016). Teknikalisasi pemetaan wilayah adat: ketahanan sosial budaya komunitas etnis Sumuri di Kabupaten Teluk Bintuni, Papua. *Islam Realitas: Journal of Islamic & Social Studies*, 2(2), 142. https://doi.org/10.30983/islam_realitas.v2i2.185
- Toussaint, E. F. A., Hall, R., Monaghan, M. T., Sagata, K., Ibalim, S., Shaverdo, H. V., Vogler, A. P., Pons, J., & Balke, M. (2014). The towering orogeny of New Guinea as a trigger for arthropod megadiversity. *Nature Communications*, 5(1), 4001. <https://doi.org/10.1038/ncomms5001>
- Toussaint, E. F. A., White, L. T., Shaverdo, H., Lam, A., Surbakti, S., Panjaitan, R., Sumoked, B., von Rintelen, T., Sagata, K., & Balke, M. (2021). New Guinean orogenic dynamics and biota evolution revealed using a custom geospatial analysis pipeline. *BMC Ecology and Evolution*, 21(1), 51. <https://doi.org/10.1186/s12862-021-01764-2>
- Upton, S. I. (2009). *The impact of migration on the people of Papua, Indonesia: A historical demographic analysis*. Universitas New South Wales.
- Voelker, G., Marks, B. D., Kahindo, C., A'genonga, U., Bapeamoni, F., Duffie, L. E., Huntley, J. W., Mulotwa, E., Rosenbaum, S. A., & Light, J. E. (2013). River barriers and cryptic biodiversity in an evolutionary museum. *Ecology and Evolution*, 3(3), 536–545. <https://doi.org/10.1002/ece3.482>
- von Rintelen, K., Arida, E., & Häuser, C. (2017). A review of biodiversity-related issues and challenges in megadiverse Indonesia and other Southeast Asian

countries. *Research Ideas and Outcomes*, 3, e20860.
<https://doi.org/10.3897/rio.3.e20860>

Wallace, A. R. (1878). *Tropical nature, and other essays* . London, Macmillan and co, 1878. <https://www.biodiversitylibrary.org/item/16104>

Tables

Table 1. Genetic diversity of *Myristica argentea* inferred by 14 nSSRs in all populations

Population	nSSRS						cpDNA				
	N_{INDV}	N_A	$AR_{(k=7)}$	H_E	H_O	F_{IS}	N_{INDV}	H	π	S	P
Ambon	7	4.79	3.09 ^{bcd}	0.512	0.576	0.316	5	2	0.153	144	50
Raja Ampat	8	2.57	2.08 ^f	0.424	0.357	0.119*	-	-	-	-	-
Manokwari	63	5.07	2.43 ^{def}	0.415	0.479	0.142	2	2	0.115	65	0
Mbahambandara	16	4.71	2.93 ^{de}	0.500	0.557	0.230	1	1			
Batuplat	19	4.79	2.94 ^{de}	0.419	0.541	0.133	2	2			
Mabuni-buni	39	5.29	2.88 ^{de}	0.533	0.604	0.239	2	1			
Mamur	58	6.57	3.16 ^{bcd}	0.505	0.582	0.127	3	1			
Kramongmongga	18	4.79	3.01 ^{bcd}	0.456	0.490	0.163	3	1			
Raduria	31	4.43	2.52 ^e	0.670	0.647	0.086	3	1			
Pawada	24	6.00	3.40 ^{ab}	0.629	0.630	0.014	3	1	0.259	306	187
Kokas	25	6.14	3.39 ^{abc}	0.661	0.646	0.023	2	1			
Nika	20	5.43	3.36 ^{abc}	0.663	0.624	0.004	2	1			
Ubadari	7	4.57	3.52 ^{ab}	0.626	0.629	0.015	2	1			
Patimburak	27	6.07	3.31 ^{abc}	0.567	0.544	0.024	2	1			
Purwahabtongggo	12	4.36	3.00 ^{bcd}	0.514	0.443	0.009	4	1			
Perwasak	7	2.93	5.58 ^a	0.561	0.665	0.048	2	1			
Seram	39	4.86	3.03 ^{bcd}	0.412	0.476	0.153	4	2	0.163	121	0

Number of individuals analyzed (N_{INDV}), mean number of allele per populations (N_A), rarefied allelic richness for $k=7$ gene copies ($AR_{(k=7)}$) for all populations, observed heterozygosity (H_O), expected heterozygosity (H_E), inbreeding coefficient (F_{IS}) calculated overall loci on each population, where * indicates values significantly different from zero with $P < 0.01$. Values with different letters are significantly different. For cpDNA, we calculated number of haplotype (H), nucleotide diversity (π), number of segregating sites (S), number of parsimony-informative sites (P) obtained by cpDNA loci. To account for the limited informative data at the population level, we calculated the overall values of π (nucleotide diversity), S (number of segregating sites), and P (number of parsimony-informative sites) for the Fakfak Peninsula. This approach was necessary due to the prevalence of shared haplotypes among most populations in the peninsula, resulting in a lack of meaningful information when calculating these parameters on a per-population basis. Grey-colored rows represent populations of the species in its core distribution.

Figures

Figure 1. The inner caption within Indonesian archipelago is pointing to (A) sampling locations of *Myristica argentea* populations in Indonesia; (B) sampling locations of *M. argentea* subpopulations in Fakfak Peninsula population. Individuals are indicated by a color according to their population.

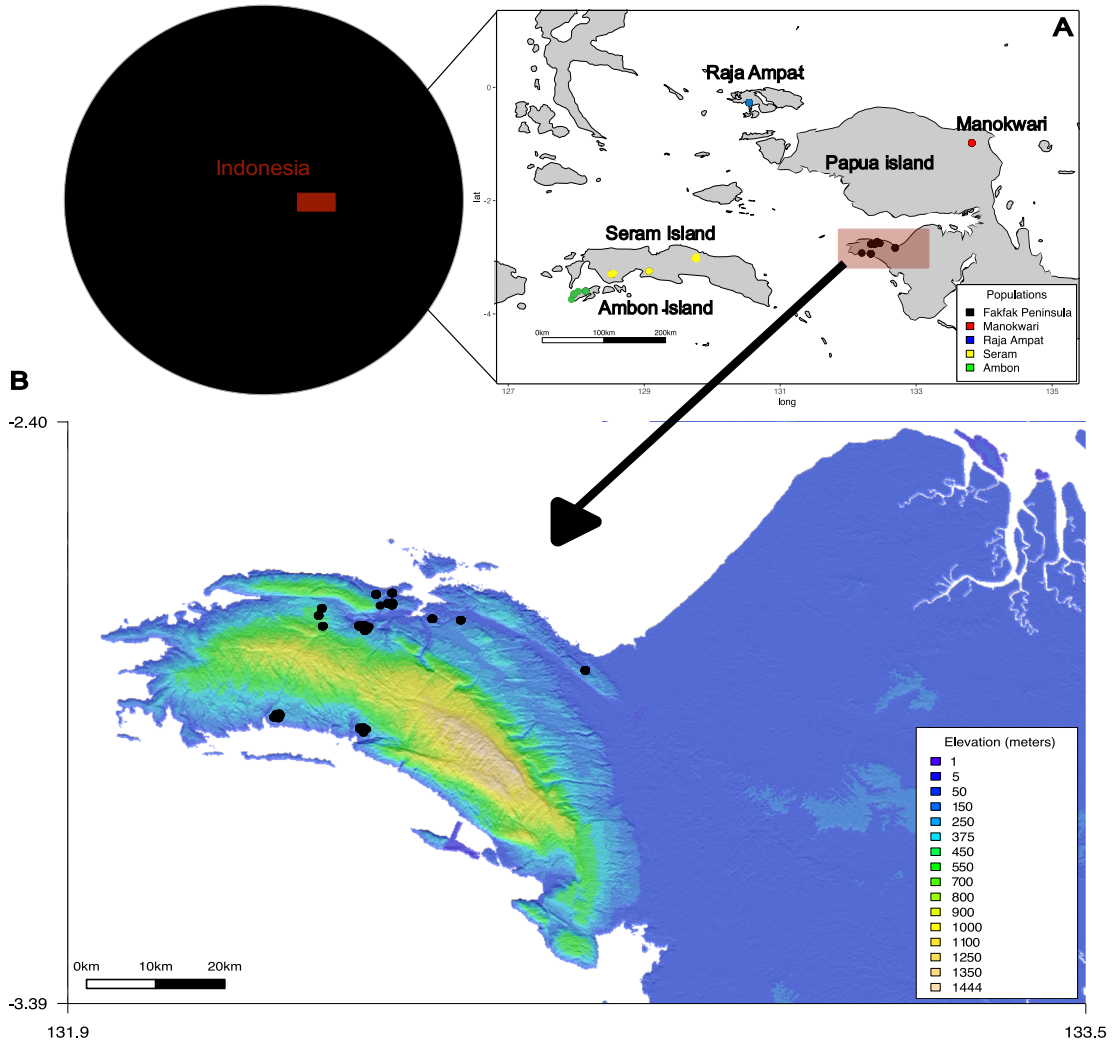


Figure 2. Population structure based on nSSRs in all seventeen populations using STRUCTURE. Geographical distribution of individuals for (A) $K = 4$ in all seventeen populations, and (B) in the species' core distribution, Fakfak Peninsula. Delta K value (C), and barplot of individual admixture proportions for each individual inferred by STRUCTURE for $K = 2$ to $K = 5$ (D).

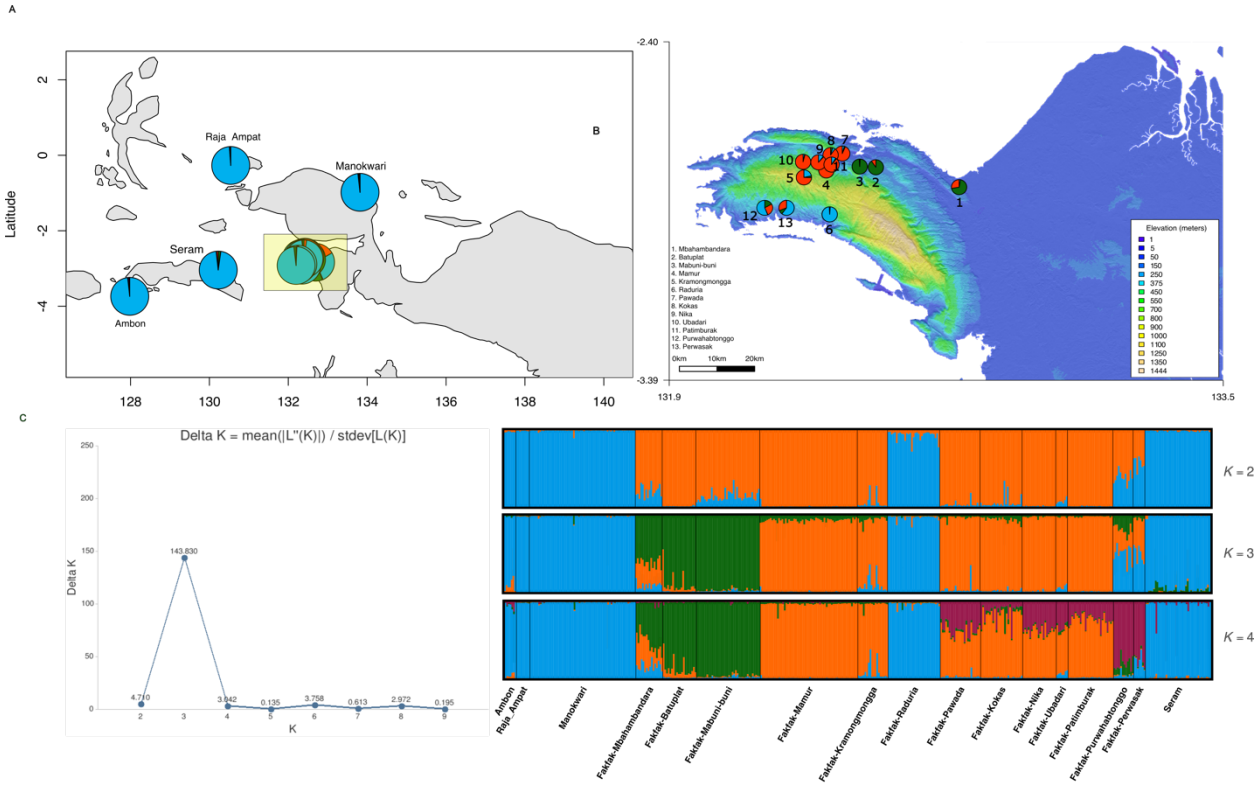


Figure 3. Maximum parsimonious network based on the cpDNA haplotypes. Each circle represents a unique haplotype according to their populations. The haplotypes were color-coded based on their geographical origin. Purple represents Ambon and Seram, pink represents Manokwari, and yellow represents the outgroup (*M. fragrans*). The color coding in Fakfak Peninsula corresponds to structure analysis. Cluster 1 (dark green) includes Mbahambandara, Mabuni-buni, and Batuplat. Cluster 2 (orange) consists of Mamur, Kramongmongga, Pawada, Kokas, Nika, Ubadari, and Patimburak. Cluster 3 (light blue) is composed of Raduria, Perwasak, and Purwahabtonggo. The legend provides the assignment of haplotypes in each population.

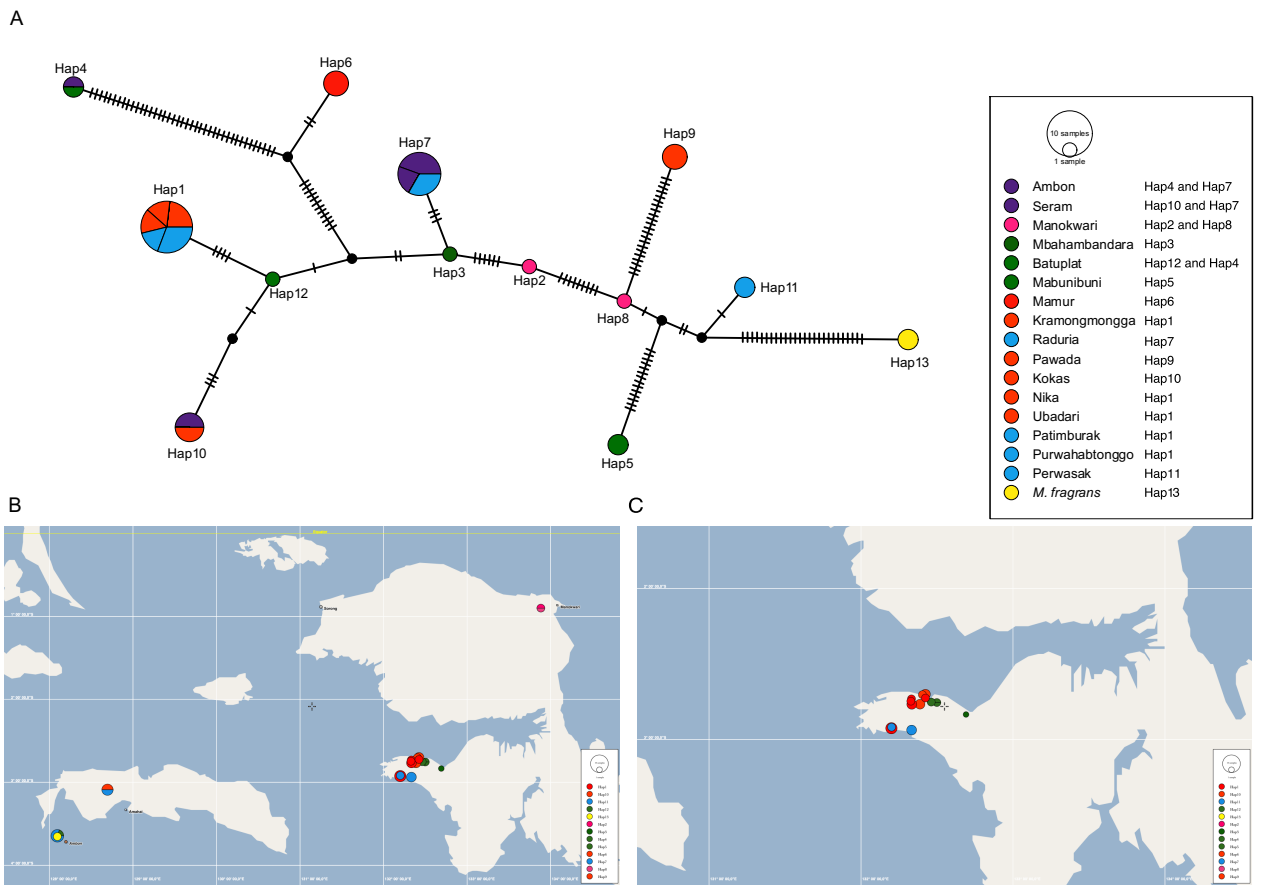


Figure 4. Inference of population histories produced by Stairway Plot on the nrDNA sequence in Fakkak subpopulations. (A) Cluster 1 (red dots in the map), including Mbahambandara, Mabuni-buni, and Batuplat; (B) Cluster 2 (blue dots), including Mamur, Kramongmongga, Pawada, Kokas, Nika, Ubadari, and Patimburak; and (C) Cluster 3 (green dots), including Raduria, Perwasak and Purwahabtonngo. Median values of N_e is indicated in solid colored line, while dashed lines correspond to the upper and lower bounds of the 95% confidence intervals. Original plots generated by Stairway Plot and confidence interval values in table format are available in Dryad (<https://doi.org/10.5061/dryad.f7m0cfz1w>). (D) Geographical distribution on each cluster.

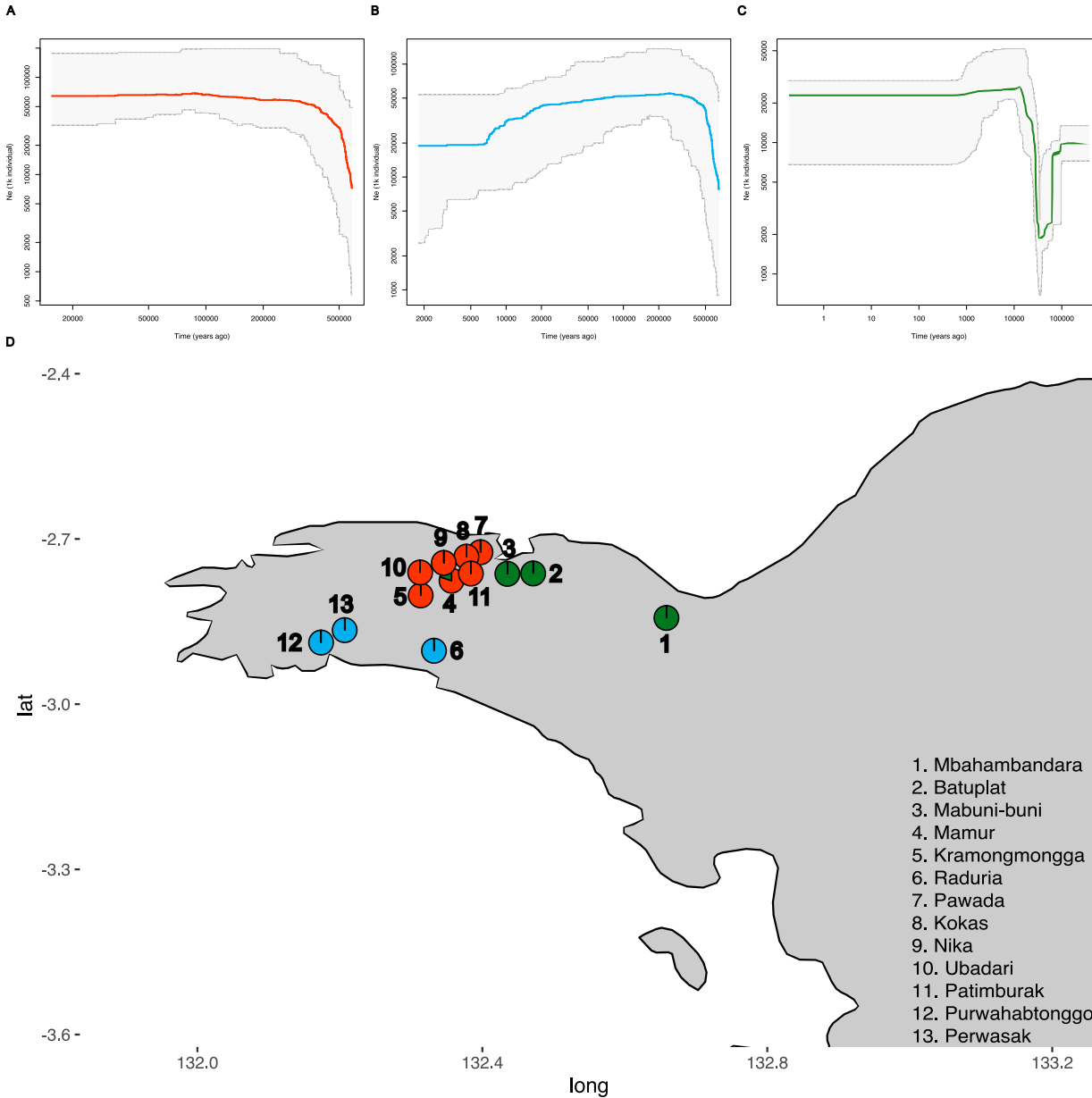
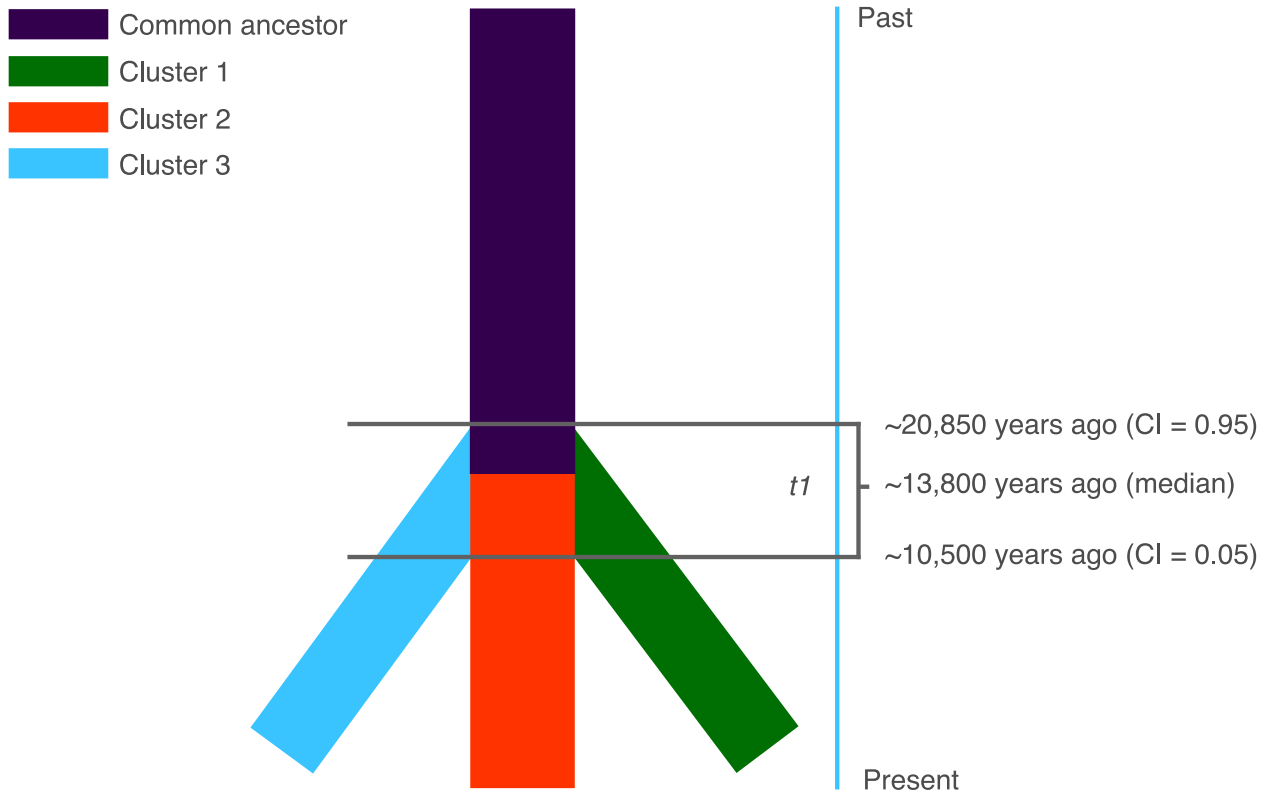


Figure 5. Most favored scenarios (Scenario 1) according to ABC-RF at the genepool level in the species' core distribution. We defined three populations for this analysis: Cluster 1 includes Mbahambandara, Mabuni-buni, and Batuplat; Cluster 2 consists of Mamur, Kramongmongga, Pawada, Kokas, Nika, Ubadari, and Patimburak; and Cluster 3 is composed of Raduria, Perwasak and Purwahabtonngo. This scenario describes all three Cluster split at the same time during $t1$.



Supplementary Online Materials

Table S1. Description of the historical parameter priors on the demographic inference of *M. fragrans* in ABC analysis.

Parameter type	Parameter name	Prior distributions
Population size [†]	Ne	Log_U [10 - 50,000,000)
	N1	Log_U [10 - 1,000,000]
	N2	Log_U [10 - 25,000,000]
	N3	Log_U [10 - 5,000,000]
Divergence time [‡]	<i>t</i> 1	U [10 - 1000]
	<i>t</i> 2	U [10 - 10000]
Conditions	<i>t</i> 2> <i>t</i> 1	

[†]In number of reproductively effective diploid individuals. [‡]In number of generations, assuming a generation time of 50 years. Distributions were considered either uniform (U) or Log-uniform (Log_U). Reference tables were constructed for 100,000 data sets per scenario.

Figure S1. Graphic representation of inferred population history of *M. argentea*. The three distinct genepool of *M. argentea*'s core distribution was based on STRUCTURE analysis. The aim is to identify the ancestral population genetic cluster that served as the basis for the genetic clusters outside of its native range. We defined three populations for this analysis: Pop 1 includes Mbahambandara, Mabuni-buni, and Batuplat; Pop 2 consists of Mamur, Kramongmongga, Pawada, Kokas, Nika, Ubadari, and Patimburak; and Pop 3 is composed of Raduria, Perwasak and Purwahabtonggo. We proposed four scenarios to infer demographic history using ABC analysis. In Scenario 1, all three Pops split at the same time at t_1 . Scenario 2 suggests that Pop 2 and Pop 3 diverged at t_2 and Pop 1 originated from Pop 3 at t_1 . In Scenario 3, Pop 1 and Pop 2 diverged at t_2 and originated from Pop 3 at t_1 . Finally, in Scenario 4, Pop 1 and Pop 3 diverged at t_2 and Pop 2 originated from Pop 1 at t_1 .

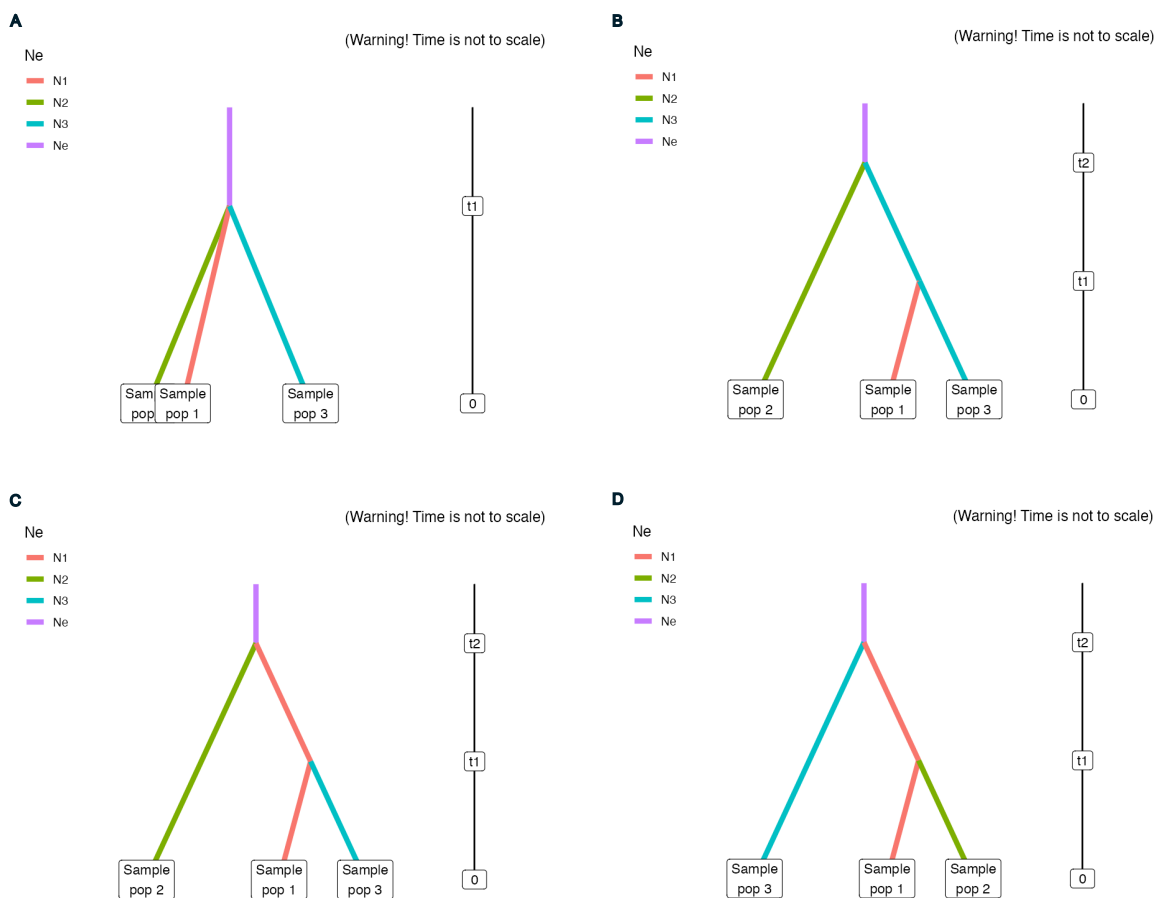


Figure S2. Percentage of missing data of 473 *M. argentea* individuals in 5 different populations as inferred using 16 nSSRs. Masohi and Sawai are located in a same island as Seram population, we therefore grouped them together.

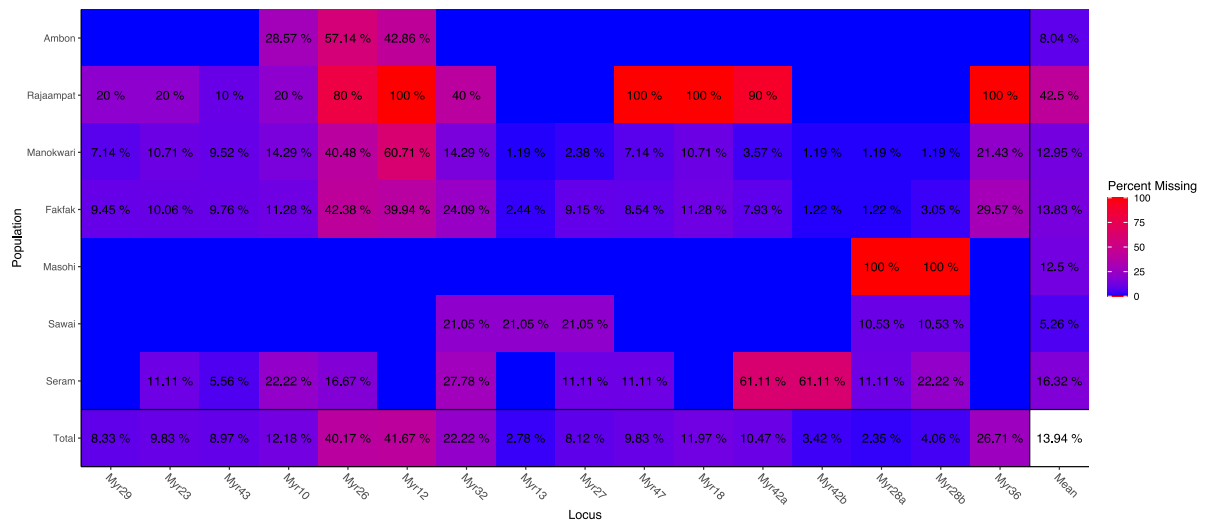


Figure S3. Pairwise F_{ST} in all populations inferred by SSR. Values are visualized in a heatmap plot (A) and a dendrogram tree (B) in all thirteen populations.

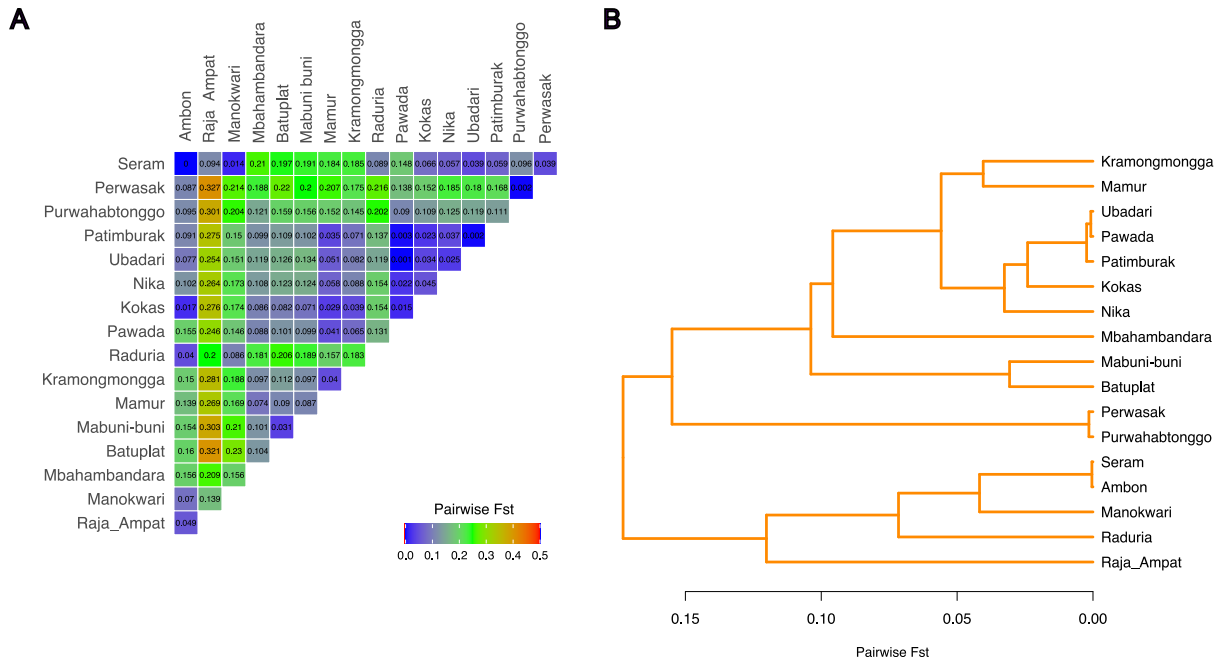
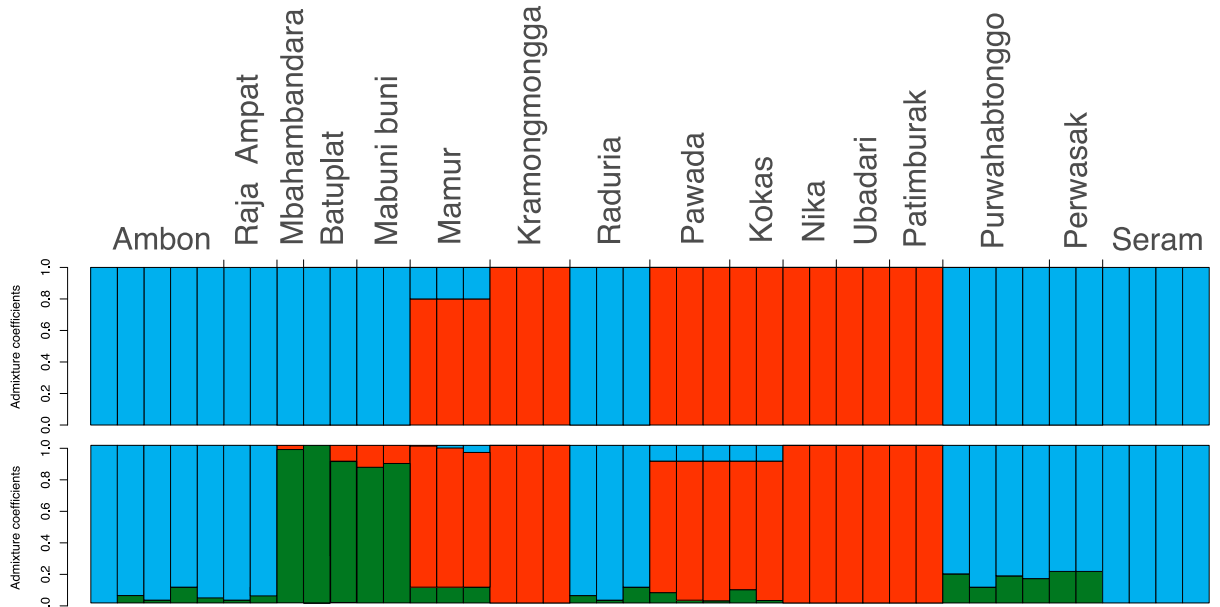


Figure S4. Population structure in all populations using fastSTRUCTURE. Barplots of individual admixture proportions, for $K = 2$ and $K = 3$ (A). Geographic projection of $K = 2$ in the species' core distribution, Fakfak Peninsula, Papua Island, Indonesia (B).

A



B

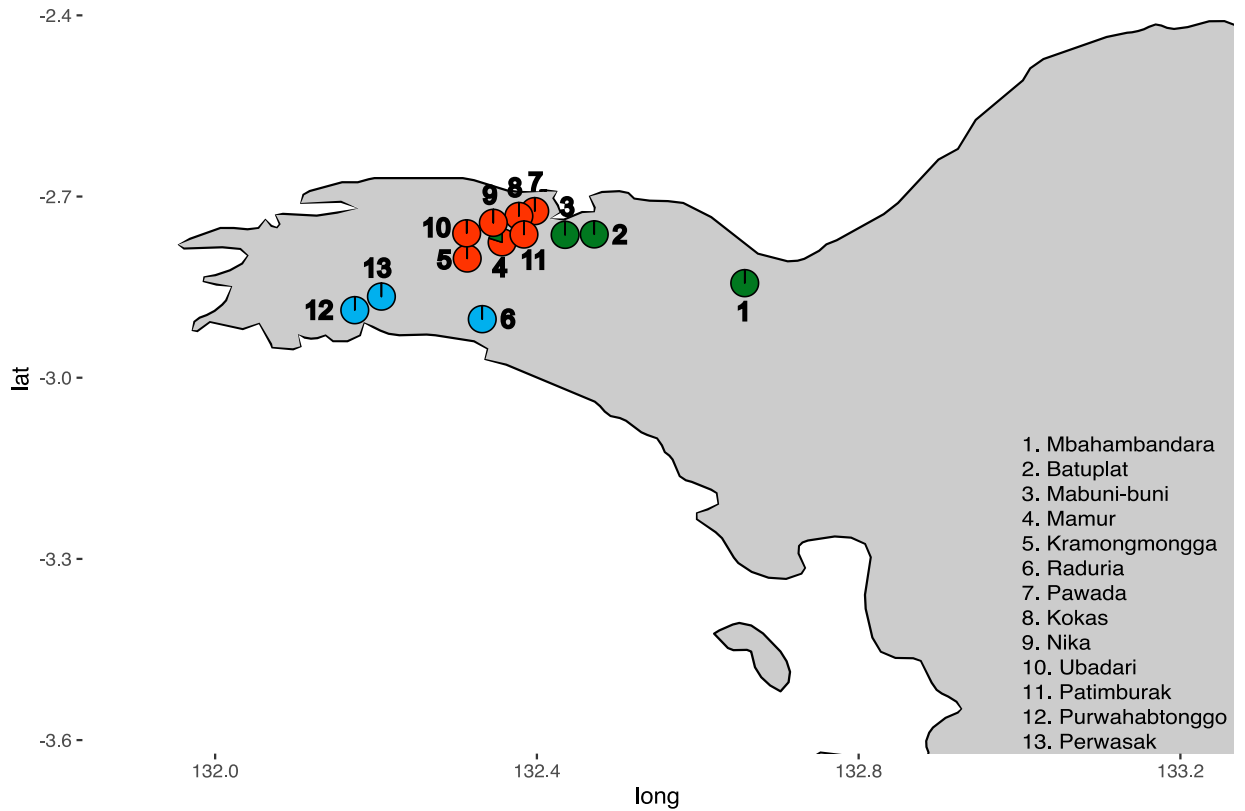
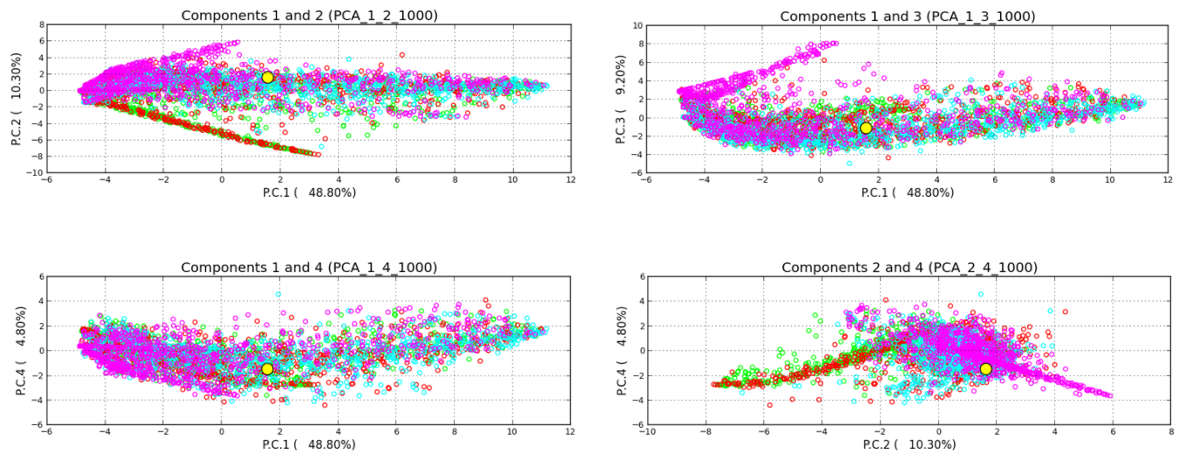


Figure S5. Principal Component Analysis (PCA) of the prior-scenario combinations and the observed dataset using DNA sequence. Scenario 1, 2, 3, and 4 priors are indicated with green, blue, purple, and red dots respectively, and observed dataset is indicated by a yellow dot.



Supplementary Files

Supplementary File S1. The following is a comprehensive inventory of the samples utilized in this study, encompassing species information, geographic coordinates, population details, as well as their application for SSR, cpDNA, and nrDNA data analysis. However, due to the longevity of the list and save some paper space for sustainability, we upload them in our GitHub page, which easily accessible in this following link: <https://github.com/jaktykusuma/myristica/tree/master/argentea>.

Supplementary File S2. This script serves the purpose of filtering fastq files based on desired quality parameters. We have made it available on our GitHub page for convenient access. You can find the script easily by following this link: <https://github.com/jaktykusuma/myristica/tree/master/scripts>.

6. Discussion

6.1 Chapter synthesis: main results

Prior to the commencement of this study, it was recognized that achieving a comprehensive understanding of the evolutionary history of *M. fragrans* and *M. argentea* would be challenging due to the scarcity of specific genetic markers available in these species. A thorough bibliographic search conducted in 2019 revealed only two available genetic resources: an incomplete cpDNA genome sequence of *M. fragrans* and SSR from *M. malabarica* (see Hemmilä et al., 2010; Tallei & Kolondam, 2015). Yet, the authors did not provide the SSR transferability from *M. malabarica* to other *Myristica* species. These limitations underscored the necessity for further investigations and the development of more reliable and tailored genetic markers to reach this thesis objectives.

Consequently, there was a particular motivation to develop new genomic resources, as detailed in **Chapter 1** and **Chapter 2** of this study. The nuclear microsatellite markers we developed were polymorphic and their transferability to other *Myristica* species (*M. fatua*, *M. argentea*, and *M. crassipes*) and to a different genus in the Myristicaceae family (*Horsfieldia palauensis*) was successful. Furthermore, this study achieved the *de novo* assembly and annotation of the chloroplast DNA (cpDNA) genomes of two *Myristica* species (*M. argentea* and *M. fatua*) with coverages of 41.2x and 31.8x, respectively. This accomplishment is significant as annotated cpDNA genome sequences are essential for understanding evolutionary history of species using maternally-inherited molecular markers. Moreover, the resulting phylogenetic reconstruction based on the entire cpDNA genome confirmed a monophyletic relationship of *Myristica* species as previously

reported (Sauquet et al., 2003), providing valuable insights into the evolution and classification of the genus.

Using newly-developed SSRs and a whole cpDNA approach, the evolutionary history of *M. fragrans* in its native range (the Moluccas archipelago) was studied (**Chapter 3**). The *M. fragrans* populations in the Moluccas exhibited a subdivision into two distinct intra-specific genetic clusters, one located in the South Moluccas and the other in the North Moluccas. The North Moluccas cluster displayed significantly greater genetic diversity compared to the South Moluccas. Bayesian demographic inference provided compelling support for an origin of the species in the South Moluccas, a colonization of the North Moluccas during the Middle to Late Pleistocene, and a relatively recent demographic decline (during the Late Pleistocene, around 20,000 to 15,000 years ago) in the South Moluccas.

In **Chapter 4**, we expanded the sampling range available in chapter 3 in order to cover the worldwide distribution of *M. fragrans* (both native and introduced populations) and to provide a comprehensive understanding of the spice's worldwide propagation history. Besides the use of nSSRs and cpDNA whole genomes, we also used targeted-nuclear sequences to strengthen our inferences of the evolutionary history of the species. By analyzing data from various sources, including herbarium specimens, the study presents a detailed account of the global distribution of *M. fragrans*. We demonstrated that introduced nutmeg populations were mainly assigned to the South Moluccas cluster, suggesting that the species introduction initiated by Pierre Poivre to all parts of the world were originated from the South Moluccas, particularly from Banda islands.

Also, additional analyses done on Chapter 4 data set confirmed and clarified results obtained in chapter 3 on the origin of the species (these new results were not included in Chapter 4 as we focused on the history of introduction of the species). Based on a more complete sampling and additional genetic data, we were able to more accurately date the past demographic events of *M. fragrans* intra-specific clusters within the Moluccas. We showed that the North Moluccas cluster, including Halmahera, Makian, and Obi islands, consistently exhibited higher genetic diversity compared to the South Moluccas cluster, which included the Seram island. Notably, our research revealed that the Halmahera population played a significant role as the main contributor to the North Moluccas cluster, while the populations in Ternate and Bacan displayed admixture from both the North Moluccas and South Moluccas clusters. Bayesian analysis on the population history of the two genetic clusters detected a migration from ~22,750 to 25,600 years ago, followed by a bottleneck event in the South Moluccas cluster around ~7,400 – 10,550 years ago. The Bayesian demographic inferences were also consistent with results of the previous chapter, i.e. the North experienced a significant expansion whereas the South experienced population decline. Using nrDNA sequences, we detected a pattern of continuous decline in the South, started ~10,000 until ~5,000 years ago. In the North, we found an expansion during ~90,000 to ~50,000 years ago.

Finally, in **Chapter 5**, we demonstrated that the Papuan nutmeg, *M. argentea*, displayed clear genetic discontinuities in the very narrow distribution range of the Fakfak peninsula, Papua Island. Bayesian clustering analyses revealed the presence of three distinct genetic clusters within the native range of the species. Introduced

populations probably mainly derived from the native population of Raduria. The populations in Fakfak Peninsula shared globally present similar levels of genetic diversity, while generally presented a higher genetic diversity levels than their introduced counterpart.

6.2 On the origin and genetic differentiation of Indonesian nutmegs: where and how?

Before we started this thesis, the hypothesis of the origin of *M. fragrans* and *M. argentea* was based solely on historical data and was never tested through a genetic approach. In the case of *M. fragrans*, there was a general consensus that the species origin was Banda Islands (Beekman, 2011; Fish, 2012; Friedrich, 2019; Hostetler, 2005; Liao, 2015), leading to its designation as Banda nutmeg. However, historical texts mentioning the center origin of this species in the Banda Islands lacked specific island references, as the Moluccas archipelago was often cited as the origin without specifying a particular island (Brixius, 2015, 2018). Through our analysis using Approximate Bayesian Computation (ABC), we found evidences suggesting that the gene pool of *M. fragrans* in the South Moluccas region represents the ancestral population. This implies that the South Moluccas can be considered as the geographical origin of *M. fragrans*. Similarly, for *M. argentea*, it has been traditionally believed to originate from the Fakfak Peninsula, earning it the names Fakfak nutmeg or Papuan nutmeg.

The origin of insular species has been hypothesized to involve colonization from their continental counterparts. Ali (2017) proposed several frameworks to explain how ancestors of these species colonized the islands, including: (1) recent

land bridges, (2) the presence of ice sheets, (3) deep-time vicariance, and (4) water dispersal. Among these, water dispersal appears to be more likely for the Moluccas archipelago, as it emerged independently from the mainland and remained unconnected even during glacial periods (Lohman et al., 2011). This notion is supported by the possibility of sea-mediated seed dispersal in nutmeg (Quigley et al., 2020). Regardless of the geological histories and age of the island, long-distance dispersal appears to be the primary factor driving subsequent divergence, rather than geological vicariance (Christenhusz & Chase, 2013; Dupin et al., 2017; Queiroz, 2005).

Long-distance seed dispersal plays a crucial role in the colonization of islands. A comprehensive review conducted by Cain et al. (2000) suggests that plant species with diverse dispersal modes, even those lacking apparent means of long-distance dispersal, were able to reach the Krakatau archipelago in the Sunda Strait, Indonesia, within approximately 50 years. It is also reported by Thornton (1996) that these species likely traveled distances of 12-40 km from Java, Sumatra, and various smaller islands to reach Krakatau. Islands separated by hundreds of kilometers from source populations are typically colonized by plant species dispersed by water or animals.

At the intraspecific level, the divergence of gene pools in *M. fragrans* suggests that the colonization mechanism from South to North was likely hydrochory (dispersal by water). The timing of this divergence aligns with past sea fluctuations in the Moluccas Sea around ~150,000 to ~100,000 years ago, during the formation of Sangihe arc (Hall et al., 1991; Morrice et al., 1983). Therefore, this could have facilitated long-distance dispersal from the South to the North. Similar

mechanisms have been reported in *Plantago* species (Iwanycki et al., 2019) and the Canaries (Fernández-Palacios et al., 2021).

Based on the aforementioned genetic differentiation speculations, it can be inferred that nutmeg was already present in the North Moluccas region prior to the arrival of Europeans. Therefore, it raises an intriguing question as to why the Europeans specifically targeted the Banda Islands during the era of the Spice Trade if *M. fragrans* was also present in the North Moluccas. The British and Dutch, in particular, engaged in significant transactions, such as trading of the island of Run (located east of the Banda Islands), in exchange of Manhattan (now part of the United States of America) (Geerken, 2021). It was recorded in Garcia de Orta's text that the tree were grown in North Moluccas, but bears no fruit (Orta, 1563), and it was recorded in the Arthur Basillo de Sá's text that the locals made no attempt to cultivate it (Sá, 1506). This historical context invites further explorations into the factors that made the Banda Islands a focal point for spice trade and why alternative sources of *M. fragrans* in the North Moluccas were not as extensively utilized or emphasized.

In contrast, our ABC analysis revealed a more straightforward divergence pattern for *M. argentea*. The species native distribution in the Fakfak Peninsula indicate a split from a common ancestor, likely influenced by climatic fluctuations during the Last Glacial Maximum (LGM) in West Papua. The LGM period triggered significant climatic changes that led to the emergence of new lineages in numerous species confined to limited ranges (Helmstetter et al., 2020; Knowles & Richards, 2005; Piñeiro et al., 2017; Vanden Abeele et al., 2021). It is well-documented that many forest-restricted species persisted in localized areas, which

provided relatively favorable climatic conditions for their survival (Blatrix et al., 2013; Hewitt, 1996; Pouteau et al., 2015). These refugia served as genetic reservoirs and, as the climate gradually warmed during the Holocene, populations expanded and recolonized other regions (Hewitt, 2000). Consequently, the populations that survived within these refugia likely create new lineages, genetically different from one refugium to the other.

Both for *M. fragrans* and *M. argentea* we observed geographically-structured intra-specific clusters within their native range. Regarding *M. fragrans*, it is probable that following the colonization event from the South Moluccas to the North Moluccas, the barrier between these two biogeographical regions (*sensu* Rutgrink et al., 2018) became closed due to the gradual rise in sea levels and warming temperatures. The reason why plants could not keep migrating may be because the changing climate made their seeds less able to spread effectively (Veresoglou & Halley, 2018). Restricted dispersal between the two regions established the fundamental basis for the observed genetic structure observed, facilitated by the changing environment. Moreover, the distinct floral composition observed between the South and North Moluccas (Rutgrink et al., 2018) could potentially mirror the observed genetic discontinuities at the population level in *M. fragrans*, indicating the role of the sea as a geographical barrier. However, in the case of the North Moluccas islands (excluding Tidore and Halmahera), a genetic admixture from both the South Moluccas and North Moluccas genetic clusters has been observed in certain individuals. It has been speculated that the initial gene pool of these islands was similar to that of Tidore and Halmahera, but human activities probably played a role in shaping the genetic makeup. Actually, as the South

Moluccas became the primary nutmeg producer at the national level (Dumatubun et al., 2020; Kakerissa & Hahury, 2021), seed materials were probably exported to neighboring islands in the North despite the presence of local material (Assagaf et al., 2021). This would have led to genetic mixing between the two genetic clusters, resulting in an admixed population. In contrast, introduced populations outside of the native range of *M. fragrans* tell a different story. These populations can be seen as colonizing new geographic ranges with no existing population. It is observed that introduced populations are genetically similar to their source populations, but they have not had sufficient time to generate new lineages, particularly considering that they are tree species due to their long generation time.

In the case of *M. argentea*, the diverse and fragmented landscape of the Fakfak Peninsula has facilitated distinct differentiation among stable populations residing in forest refugia. Similar patterns of population structure have been observed in micro-endemic species in both tropical and temperate regions (Guerrina et al., 2022; Rico et al., 2021; Salmona et al., 2023; Souza et al., 2022). The topography of the Fakfak Peninsula is characterized by a mountain range resulting from tectonic uplift, specifically the Seram Fold in the Western bound of the West Papua (Memmo et al., 2013; White et al., 2019). This ongoing geological process has shaped the micro-climate within the peninsula, further accentuated by the glaciation period that led to the formation of multiple riverine networks.

6.3 On the genetic diversity of Indonesian nutmegs

The concept of the center of origin in relation to high levels of genetic diversity is a fundamental principle in evolutionary study. The center of origin

refers to the geographic region where a particular species is believed to have originated and where its genetic diversity is expected to be the highest. This concept is based on the assumption that the longer a species has been present in a specific area, the more opportunities it has had to accumulate genetic variation through mutation. Several studies on tree species, such as *Prunus* (Liu et al., 2019; Zhang et al., 2021) and *Persea* (Sandoval-Castro et al., 2021), have provided empirical evidences supporting the correlation between the center of origin and high levels of genetic diversity.

Theoretically, we could expect that species presenting a restricted distribution might be more sensible to habitat disturbance. A disturbance of their habitat would reduce their population size and exacerbate the reduction of genetic diversity through drift. To the contrary, tree species that have a broad geographic range, such as *Populus*, *Betula*, *Alnus*, and *Fraxinus*, may have a greater genetic diversity (see review in Verbylaitė et al., 2023). These species are able to move across large distances and interact with other populations, leading to increased gene flow and a wider range of genetic variation. However, the *M. fragrans* populations in the North Moluccas and the *M. argentea* populations in the Fakfak Peninsula present a high amount of genetic diversity. This situation requires further explanations, rather than simply attributing it to the species' inherent genetic diversity. (Isabel et al., 2020) There is an ongoing debate among several reports regarding the various mechanisms responsible for generating and maintaining high genetic diversity (Isabel et al., 2020; Petit & Hampe, 2006; Scotti et al., 2016). One possible explanation for this phenomenon is that the tree species has a stable population size over time, which helps to minimize the effects of genetic drift. (Salmona et al.,

2023). Moreover, it is important to consider the possibility of introgression from closely related species, given the abundance of *Myristica* species on Papua Island (de Wilde, 1995, 2014).

Understanding the differences in genetic diversity levels between oceanic islands and continental settings is an intriguing aspect. It is widely acknowledged that continental populations generally exhibit higher genetic diversity compared to island populations (Algar & Losos, 2011; Fernández-Mazuecos & Vargas, 2011; Franks, 2010; Laenen et al., 2011). However, the findings from this study, as well as other investigations on model organisms (Fernández-Mazuecos & Vargas, 2011; Nagata et al., 2021; Salmona et al., 2023), do not always conform to this pattern. In the context of higher genetic diversity in continental settings, Hutsemekers et al. (2011) proposed that spatial heterogeneity and habitat diversity play pivotal roles in driving the levels of genetic diversity. Interestingly, this proposition can also be applicable to island settings such as the North Moluccas and Papua Island, despite their limited land area but notable variation in spatial heterogeneity (Dalsgaard et al., 2014; Siemaszko, 2015; Udy et al., 2021). These islands encompass numerous vegetation systems (Gamoga et al., 2021; Irmayanti et al., 2022), which can be interpreted as a reflection of their rich spatial heterogeneity and habitat diversity, as highlighted by Hutsemekers et al. (2011).

The human influences also have a significant impact on the distribution of genetic diversity. In the case of *M. fragrans*, the assumption that the center of origin of a species carries higher levels of genetic diversity does not hold true in the South Moluccas, due to the observed low levels of genetic diversity. It is probable that the extensive tree extirpation during the 16th century may have contributed to this

phenomenon (Geerken, 2021; Lape, 2000). In the case of *M. argentea*, its native range in Fakfak Peninsula exhibited high level genetic diversity, suggesting a stability of their habitat since the LGM and a relatively low impact of human management practices of the species. High levels of genetic diversity were also detected in few introduced populations of *M. argentea*. These populations were probably founded from several different native populations from Fakfak Peninsula, as supported by our cpDNA results.

Understanding the patterns of genetic diversity distribution in cultivated tree species and identifying their centers of origin is not only valuable for conservation efforts but also for breeding programs, as it enables the identification of genetically diverse populations that can contribute to the development of improved varieties with enhanced resilience to the changing environments.

6.4 On the population dynamics of nutmegs

The climate fluctuations during the LGM proved to be a driven factor of population size changes (Lu et al., 2020; Tsuda et al., 2015). Population size changes, such as declines or expansions, have significant impacts on the evolutionary history of a population. For example, during the LGM, many tree species experienced a significant reduction in their distribution range and population size (see examples by Leroy et al., 2020; Piñeiro et al., 2017). The evidence of historical expansions or declines may therefore indicate broader global shifts in forest vegetation patterns, encompassing not just alterations in forest cover but also transformations in forest vegetation types.

In this study, a population expansion was observed in *M. fragrans* population in the North Moluccas. The expansion of *M. fragrans* populations could be attributed to favorable ecological conditions. This includes the availability of suitable habitats and resources, as the coverage land area of North Moluccas is larger than South Moluccas. Furthermore, the historical stability of *M. argentea* populations since the LGM suggests the persistence of favorable habitat conditions on Papua Island since this period.

A different story was observed in the population of *M. fragrans* in the South Moluccas, where a recent bottleneck was identified, possibly linked to human activities. During the Spice Trade Age, the Dutch East India Company (VOC) played a significant role in the spice trade, particularly in the Banda Islands. Recognizing the profitability of the spice, the VOC sought to gain exclusive control over its production and distribution. To achieve this monopoly, the VOC adopted a policy of forcibly removing competing European traders from the Banda Islands and compelling the local population to sell their nutmeg harvest solely to the company. The VOC's control over the spice trade in the region was further solidified through the signing of exclusive trade agreements with local rulers. Furthermore, the VOC implemented a policy of "extirpation" or "extermination," which involved the systematic destruction of nutmeg trees on the Banda Islands themselves. By reducing the supply of nutmeg, the VOC aimed to maintain high prices and deter any potential competitors from establishing their own plantations.

The population size decline resulting from human-induced factors can have severe consequences, including the potential for extinction of distinct populations. These factors observed in the population of *M. fragrans*, raise concerns about the

long-term persistence and survival of this distinct population if human impacts are not mitigated. It has been suggested that maintaining a minimum effective population size (N_e) of around 500 is crucial for ensuring long-term evolutionary potential, while a N_e of approximately 50 is necessary for short-term viability (Franklin, 1980). However, for insular species like *M. fragrans* and *M. argentea*, achieving such N_e values may be challenging, given their narrow native range. When a population undergoes a size reduction, the combined effects of genetic drift and inbreeding can further diminish its fitness, potentially reinforcing the decline of the population size (Gilpin, 1986). In the short, these forces can even push a population towards extinction. Therefore, it is imperative to take proactive measures aiming at maintaining or increasing the population size of these species within their core distribution areas, exceeding their current levels.

While it is possible to identify some changes in the historical populations size, it is important to exercise caution when interpreting the exact dates due to the simplified nature of the models used. Additionally, the unavailability of information regarding the mutation rate of the focal species poses another challenge. Lastly, the consideration of generation time, which refers to the reproductive lifespan of the species, further adds to the complexities associated with accurately examining population size changes. Nonetheless, the consistency of findings regarding population size changes across unrelated species with similar habitat requirements enhances the reliability of the estimates obtained in this study.

It is also important to note that the inference of population size in the nuclear compartment used in this study has a fundamental limitation. Terhorst and Song (2015) highlighted a crucial limitation in inferring population size using the sample

frequency spectrum (SFS). They demonstrated that the accuracy of population size estimation is constrained by a lower bound, which deteriorates exponentially as the number of independent segregating sites utilized in the analysis increases. This lower bound is more unfavorable compared to the convergence rates observed in numerous classical estimation problems in statistics. It is important to note that this limitation specifically pertains to populations that have experienced a bottleneck. When utilizing microsatellite data, it is crucial to consider the potential limitation of biased parameter estimation caused by the existence of discrete demes (Leblois et al., 2014).

6.5 Limitations on the regional understanding of the importance of genetic diversity

From the commencement of the bibliographical search, it has become apparent that there was a scarcity of sources available for developing the context of floral diversity in the Malay Archipelago (*sensu* Wallace, 1886). The floral diversity in the Moluccas and Papua regions is a fascinating area of study due to the high levels of species richness (Crowther et al., 2015), offering valuable insights into biodiversity estimates. Despite its significance, there has been a limited number of research articles addressing the importance of genetic diversity and its conservation implications in this particular region. Furthermore, there is a notable geographical bias in ecological and evolutionary studies, with only 24% of the 60,000 peer-reviewed studies focused on the tropical region, while the majority pertains to the temperate regions (Culumber et al., 2019). Due to the limited scope

of studies and research findings in the area, our understanding of how to effectively preserve diversity and develop optimal conservation strategies is also constrained.

It is widely acknowledged that the Malay Archipelago is an integral part of globally recognized biodiversity hotspots (Mittermeier et al., 1999; Myers et al., 2000). These hotspots of biodiversity are threatened by anthropogenic activities, as evidenced by alarmingly high rates of deforestation (Figure 1; Hughes, 2018; Voigt et al., 2021), and the increase of forest cover change (Hansen et al., 2013). These conditions suggest that multiple factors must be considered to safeguarding the biodiversity in the Malay Archipelago: (1) by increasing the number of research activities in the region related with biodiversity measures; (2) by informing the policy makers of multiple research findings on the loss of genetic diversity; (3) by suggesting *in situ* conservation strategies to maintain genetic variation, with the aim to conserve distinct populations, as suggested by Minter et al. (2021). In relation with this issue, it has been suggested in this thesis that the excessive extirpation of *M. fragrans* trees by the VOC had caused a loss of genetic diversity. Furthermore, preserving genetic diversity plays a crucial role in mitigating the loss of biodiversity (Reed & Frankham, 2003) and facilitating species' adaptation to dynamic environments (Wernberg et al., 2018).

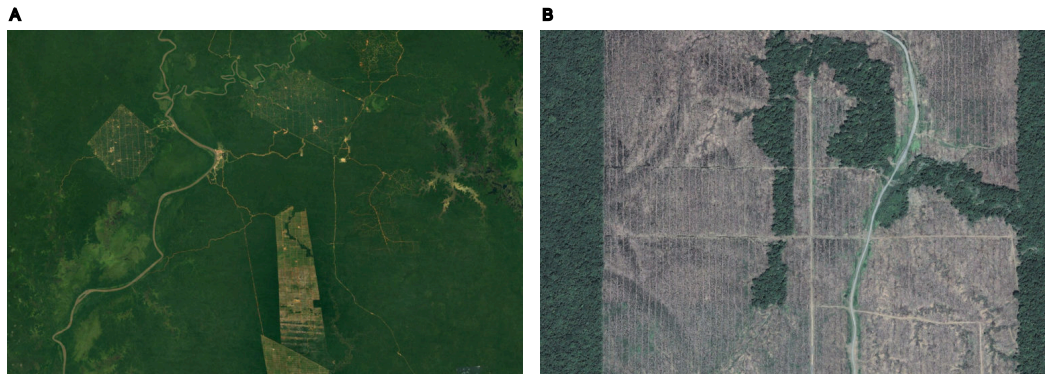


Figure 1. In Papua Island, recent deforestation for oil palm plantations (A) and the remaining deforestation from non-productive palm oil activities without subsequent reforestation efforts (B) have become prominent issues. Photo credit: NASA Landsat.

However, the scarcity of biodiversity and evolutionary research as well as the research-oriented policies in the region could be attributed to the administrative challenges associated with conducting research in Indonesia. In relation with this issue, the research effort is burdened by the persistent and never-ending bureaucratic procedure involved in acquiring genetic material (this includes Indonesia and Malaysia), despite the initial aim of preventing numerous cases of biopiracy by foreign countries (Maydrawati, 2020).

The National Research and Innovation Agency (known as BRIN in the Indonesian language), has established a streamlined procedure for obtaining research permits, which is relatively easy and quick. However, following this initial step, researchers are faced with a complex process of obtaining multiple permits from various government bodies, particularly in the field of ecology and evolution. These include permits for conducting research, for entering protected areas, obtaining permission from the National Police, for collecting samples, for exporting samples from *in situ* to *ex situ* areas, and for exporting samples from Indonesia to other countries. This lengthy and intricate process can span up to a year,

discouraging many researchers and leading them to reconsider and abandon their plans to conduct research in Indonesia.

Following the implementation of the Nagoya Protocol in Indonesia, legislation pertaining to the access of genetic resources was established in 2013. As part of the national commitment to the Nagoya Protocol, an Access and Benefit Sharing (ABS) National Focal Point (NFP) has been designated. However, it has become apparent that the NFP position is influenced by political considerations rather than conservation ideology, which regrettably introduces additional complexities when attempting to gather information. The list of contact details provided on the website (<https://absch.cbd.int/en/countries/ID>) proves to be futile, and it is discouraging to note that the webpage has not been updated since 2019.

It is fascinating to observe the stark contrast in the efforts made by local communities to conserve the species at the population level. In the North Moluccas, specifically on the islands of Ternate and Tidore, the forests of *M. fragrans* are protected by the locals. Their motivation stems not from the economic value of nutmeg, but rather from the belief that nutmeg trees stabilize the ground against tectonic activities caused by volcanoes. A similar scenario unfolds in the *M. argentea* forests of the Fakfak Peninsula, where the local communities impose traditional sanctions on those who harm Papuan nutmeg trees. This protective measure is taken because Papuan nutmeg trees serve as a valuable resource and also ensure the survival of Hornbills during their migration. These examples are also applicable to other tree species within the forest, extending beyond the scope of nutmeg trees.

Finally, effective conservation necessitates collective endeavors and collaborative partnerships among policymakers, researchers, and local communities. The preservation of intra-specific genetic diversity is essential in enhancing the resilience of populations against climate variations and human-induced pressures. Although *M. fragrans* and *M. argentea* are not presently classified as endangered, their conservation status is categorized as DD (data deficient) for *M. fragrans* (World Conservation Monitoring Centre, 1998), and VU (vulnerable) for *M. argentea* (Chadburn, 2020). Consequently, these insular species are deemed vulnerable to extinction due to their limited geographical distribution. Hence, it is imperative to reinforce policies that prioritize *in situ* conservation efforts with a focus on genetic diversity management.

6.6 On the conservation and sustainable management: perspectives and the future

The study conducted on *M. fragrans* and *M. argentea* provided valuable insights into the evolutionary history of these two species and its relationship with the environment and anthropogenic activities. These perspectives can inform conservation strategies, such as identifying priority areas for protection and implementing measures to mitigate the impacts of human activities on their genetic integrity. This study provides a deeper exploration and understanding of the dynamics and conservation needs of these insular species in the face of ongoing environmental challenges. Thus, the findings of this study have several perspectives and future inputs.

First, the study highlights the significance of utilizing genetic markers from various compartments (nuclear and plastid) to facilitate a precise and comprehensive genetic analyses of plant species. The development of species-specific microsatellite markers addresses previous limitations and offers enhanced resolution and accuracy in discerning migration rates and relatedness among individuals (Selkoe & Toonen, 2006). This is particularly valuable for non-model species, where the utilization of species-specific markers proves to be the optimal choice (Glenn & Schable, 2005). In addition, another valuable aspect is the complete chloroplast genome sequences of *M. fatua* and *M. argentea*. The availability of these chloroplast genomes opens up opportunities for researchers in various fields, including plant phylogeny, molecular ecology, comparative genomics, population genetics, and evolutionary studies of *Myristica* species.

Second, our findings on the distribution of genetic diversity in these focal species shed light on the contrasting levels of genetic diversity in the introduced populations, and among their native populations. As an example, the impact of human activities on the levels of genetic diversity of *M. fragrans* was demonstrated by the contrasting experiences of two distinct populations of *M. fragrans*. One population remains unaffected by human intervention, while the other has experienced historical extirpation caused by human activities. The perspectives gained from intra-specific genetic diversity analysis will be instrumental in understanding how different populations may respond to human activities, and will provide valuable guidance for developing targeted conservation strategies within the species' core distribution. Equipped with this information, we can establish a robust and integrated framework for *in situ* genetic conservation, ensuring the long-

term viability and resilience of these species in their native range. As previously recommended by Minter et al. (2021), Gene Conservation Units (GCU) for tree species should be addressed to a distinct or local population.

Third, in this thesis, we observed a shared gene pool between the introduced population and one of the populations in the native range of both focal species. From a conservation standpoint, agricultural intensifications can result in genetic homogenization (see Archidona-Yuste et al., 2021; Vellend et al., 2007), potentially eradicating crucial distinct gene pools that may be essential for the species' adaptation to a changing environment. To safeguard these distinct populations or gene pools, it is necessary to establish specific conservation areas, similar to the approach taken by local communities in preserving the *M. argentea* population in the Fakfak Peninsula. Their exceptional efforts in safeguarding and maintaining genetic diversity at the gene pool level serve as an example that can be expanded and applied at a larger geographical scale by policymakers. It is worth noting that this conservation effort is unconscious (as they actually conserve the trees for the ecosystem services they provide), and that it will be worth communicating that these conservation efforts are also beneficial to maintain genetic resources of these forest species.

Fourth, the demographic inference analysis provides valuable insights into the geographic origin of the species and illustrates how populations have responded to anthropogenic activities. The evidence of bottlenecked populations in the South Moluccas, the center of origin of *M. fragrans*, highlights the significant impact of human influence in the past. It is suggested that detecting recent changes in population size is crucial for designing effective conservation management plans.

While it is important to enhance the power of demographic inference through addition of comprehensive genetic resources, we also emphasize the need for a well-distributed sample collection across all islands to ensure a comprehensive understanding of the species' demographic history.

Lastly, the collection of samples in Indonesia presents significant challenges. In addition to the administrative procedures mentioned earlier, the unique island composition poses difficulties in accessing various locations. Moving from one island to another entails substantial costs and consumes considerable time. Furthermore, the unpredictable nature of the monsoon season can disrupt flight and ferry schedules, further complicating fieldwork. Many of these islands are particularly vulnerable to the impacts of climate change, including rising sea levels that threaten the very existence of these islands and the diverse life they support. However, the collection of materials, regardless of the remoteness of the location, is crucial. Numerous findings indicate that endemic species within a specific population may serve as a testament to their own evolutionary history (Barstow et al., 2022; Kier et al., 2009; Salmons et al., 2023; Stuessy et al., 2014). As a result, it is possible that we are still underestimating the full biological and genetic complexity of the species, particularly in the case of insular species like *M. fragrans* and *M. argentea*. This underscores the significance of not only conservation efforts, but also the preservation of these species at the population level, with their specimens being sampled or stored in the herbarium.

It would be unfortunate if we were unable to preserve threatened populations, as they could hold valuable genetic information crucial for conservation efforts. An examination of the global database of herbarium sampling (GBIF.org, 2023a,

2023b), unveils an atypical geographical distribution for *M. fragrans* and *M. argentea*, indicating that the species does not fully cover the entire geographical range of their native distributions. By addressing these challenges, we will have the opportunity to strategically select distinct populations for genetic monitoring, ensuring representation from key areas of conservation priority, cultural importance, ecosystem services, and wild harvesting. This approach would enable a more comprehensive understanding of the species' genetic diversity and facilitate targeted conservation efforts that are aligned with the specific needs of these populations.

References

- Algar, A. C., & Losos, J. B. (2011). Evolutionary assembly of island faunas reverses the classic island-mainland richness difference in *Anolis* lizards. *Journal of Biogeography*, 38(6), 1125–1137. <https://doi.org/10.1111/j.1365-2699.2010.02466.x>
- Ali, J. R. (2017). Islands as biological substrates: classification of the biological assemblage components and the physical island types. *Journal of Biogeography*, 44(5), 984–994. <https://doi.org/10.1111/jbi.12872>
- Archidona-Yuste, A., Wiegand, T., Eisenhauer, N., Cantalapiedra-Navarrete, C., Palomares-Rius, J. E., & Castillo, P. (2021). Agriculture causes homogenization of plant-feeding nematode communities at the regional scale. *Journal of Applied Ecology*, 58(12), 2881–2891. <https://doi.org/10.1111/1365-2664.14025>
- Assagaf, M., Musyavak, A., Wahab, A., Cahyaningrum, H., & Amiruddin, A. (2021). The competitiveness analysis of Nutmeg farming as spices in Ternate, North Moluccas. *IOP Conference Series: Earth and Environmental Science*, 807(4), 042077. <https://doi.org/10.1088/1755-1315/807/4/042077>
- Barstow, M., Jimbo, T., & Davies, K. (2022). Extinction risk to the endemic trees of Papua New Guinea. *PLANTS, PEOPLE, PLANET*. <https://doi.org/10.1002/ppp3.10342>
- Beekman, E. M. (2011). *The Ambonese Herbal by Rumphius* (1st ed., Vol. 2). Yale University Press.
- Blatrix, R., McKey, D., & Born, C. (2013). Consequences of past climate change for species engaged in obligatory interactions. *Comptes Rendus Geoscience*, 345(7–8), 306–315. <https://doi.org/10.1016/j.crte.2013.03.006>
- Brixius, D. (2015). A Pepper acquiring Nutmeg: Pierre Poivre, The French Spice Quest and the Role of Mediators in Southeast Asia, 1740s to 1770s. *Journal of the Western Society for French History*, 43(1).
- 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>
- Cain, M. L., Milligan, B. G., & Strand, A. E. (2000). Long-Distance Seed Dispersal in Plant Populations. *American Journal of Botany*, 87(9), 1217–1227. <https://doi.org/10.2307/2656714>
- Chadburn, H. (2020). *Myristica argentea*. The IUCN Red List of Threatened Species 2020: E.T144674963A144774444.

<https://dx.doi.org/10.2305/IUCN.UK.2020-2.RLTS.T144674963A144774444.en>

- Christenhusz, M. J. M., & Chase, M. W. (2013). Biogeographical patterns of plants in the Neotropics - dispersal rather than plate tectonics is most explanatory. *Botanical Journal of the Linnean Society*, 171(1), 277–286. <https://doi.org/10.1111/j.1095-8339.2012.01301.x>
- Crowther, T. W., Glick, H. B., Covey, K. R., Bettigole, C., Maynard, D. S., Thomas, S. M., Smith, J. R., Hintler, G., Duguid, M. C., Amatulli, G., Tuanmu, M.-N., Jetz, W., Salas, C., Stam, C., Piotta, D., Tavani, R., Green, S., Bruce, G., Williams, S. J., ... Bradford, M. A. (2015). Mapping tree density at a global scale. *Nature*, 525(7568), 201–205. <https://doi.org/10.1038/nature14967>
- Culumber, Z. W., Anaya-Rojas, J. M., Booker, W. W., Hooks, A. P., Lange, E. C., Puer, B., Ramírez-Bullón, N., & Travis, J. (2019). Widespread Biases in Ecological and Evolutionary Studies. *BioScience*, 69(8), 631–640. <https://doi.org/10.1093/biosci/biz063>
- Dalsgaard, B., Carstensen, D. W., Fjeldså, J., Maruyama, P. K., Rahbek, C., Sandel, B., Sonne, J., Svenning, J., Wang, Z., & Sutherland, W. J. (2014). Determinants of bird species richness, endemism, and island network roles in Wallacea and the West Indies: is geography sufficient or does current and historical climate matter? *Ecology and Evolution*, 4(20), 4019–4031. <https://doi.org/10.1002/ece3.1276>
- de Wilde, W. J. J. O. (1995). Census of *Myristica* (Myristicaceae) in New Guinea anno 1994. *Blumea: Biodiversity, Evolution and Biogeography of Plants*, 40(2), 237–344.
- de Wilde, W. J. J. O. (2014). *Flora Malesiana*. Pensoft Publishers. <https://doi.org/10.3897/ab.e1141>
- Dumatubun, E., Pattinama, M., & Timisela, N. (2020). Strategi pengembangan komoditas biji pala di Ambon. *Agrilan : Jurnal Agribisnis Kepulauan*, 8, 190. <https://doi.org/10.30598/agrilan.v8i2.1046>
- Dupin, J., Matzke, N. J., Särkinen, T., Knapp, S., Olmstead, R. G., Bohs, L., & Smith, S. D. (2017). Bayesian estimation of the global biogeographical history of the Solanaceae. *Journal of Biogeography*, 44(4), 887–899. <https://doi.org/10.1111/jbi.12898>
- Fernández-Mazuecos, M., & Vargas, P. (2011). Genetically Depauperate in the Continent but Rich in Oceanic Islands: *Cistus monspeliensis* (Cistaceae) in the Canary Islands. *PLoS ONE*, 6(2), e17172. <https://doi.org/10.1371/journal.pone.0017172>

- Fernández-Palacios, J. M., Otto, R., Borregaard, M. K., Kreft, H., Price, J. P., Steinbauer, M. J., Weigelt, P., & Whittaker, R. J. (2021). Evolutionary winners are ecological losers among oceanic island plants. *Journal of Biogeography*, 48(9), 2186–2198. <https://doi.org/10.1111/jbi.14143>
- Fish, S. (2012). *The manila-acapulco galleons : the treasure ships of the pacific : with an annotated list of the transpacific galleons 1565-1815*. AuthorHouse.
- Franklin, I. R. (1980). *Evolutionary change in small populations*.
- Franks, S. J. (2010). Genetics, Evolution, and Conservation of Island Plants. *Journal of Plant Biology*, 53(1), 1–9. <https://doi.org/10.1007/s12374-009-9086-y>
- 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>
- Gamoga, G., Turia, R., Abe, H., Haraguchi, M., & Iuda, O. (2021). The Forest Extent in 2015 and the Drivers of Forest Change Between 2000 and 2015 in Papua New Guinea. *Case Studies in the Environment*, 5(1). <https://doi.org/10.1525/cse.2021.1442018>
- GBIF.org. (2023a). *Myristica argentea* occurrence. <https://doi.org/10.15468/DI.Uue8w9>.
- GBIF.org. (2023b, May 15). *Myristica fragrans* occurrence. GBIF Occurrence Download <https://doi.org/10.15468/DI.Uue8w9>.
- Geerken, H. H. (2021). *The Gold of the Bandas: The History of the Nutmeg: The forgotten islands that once made world history*. BoD--Books on Demand.
- Gilpin, M. (1986). *Minimum viable populations : Processes of species extinction*.
- Glenn, T. C., & Schable, N. A. (2005). *Isolating Microsatellite DNA Loci* (pp. 202–222). [https://doi.org/10.1016/S0076-6879\(05\)95013-1](https://doi.org/10.1016/S0076-6879(05)95013-1)
- Guerrina, M., Theodoridis, S., Minuto, L., Conti, E., & Casazza, G. (2022). First evidence of post-glacial contraction of Alpine endemics: Insights from *Berardia subacaulis* in the European Alps. *Journal of Biogeography*, 49(1), 79–93. <https://doi.org/10.1111/jbi.14282>
- Hall, R., Nichols, G., Ballantyne, P., Charlton, T., & Ali, J. (1991). The character and significance of basement rocks of the southern Molucca Sea region. *Journal of Southeast Asian Earth Sciences*, 6(3–4), 249–258. [https://doi.org/10.1016/0743-9547\(91\)90071-5](https://doi.org/10.1016/0743-9547(91)90071-5)

- 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., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O., & Townshend, J. R. G. (2013). High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science*, 342(6160), 850–853. <https://doi.org/10.1126/science.1244693>
- Helmstetter, A. J., Béthune, K., Kamdem, N. G., Sonké, B., & Couvreur, T. L. P. (2020). Individualistic evolutionary responses of Central African rain forest plants to Pleistocene climatic fluctuations. *Proceedings of the National Academy of Sciences*, 117(51), 32509–32518. <https://doi.org/10.1073/pnas.2001018117>
- Hemmilä, S., Kumara, M., Ravikanth, G., Gustafsson, S., Vasudeva, R., Ganeshiah, K. N., Uma Shaanker, R., & Lascoux, M. (2010). Development of eleven microsatellite markers in the red-listed tree species *Myristica malabarica*. *Conservation Genetics Resources*, 2(S1), 305–307. <https://doi.org/10.1007/s12686-010-9212-7>
- Hewitt, G. (1996). Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, 58(3), 247–276. <https://doi.org/10.1006/bijl.1996.0035>
- Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, 405(6789), 907–913. <https://doi.org/10.1038/35016000>
- Hostetler, L. (2005). *Qing colonial enterprise : ethnography and cartography in early modern china*. University of Chicago Press.
- Hughes, A. C. (2018). Have Indo-Malaysian forests reached the end of the road? *Biological Conservation*, 223, 129–137. <https://doi.org/10.1016/j.biocon.2018.04.029>
- 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>
- Irmayanti, L., Ashari, R., Peniwidiyanti, Umanailo, F., Rangkuti, A. B., Fatrawana, A., Nurhikmah, Nurdin, A. S., & Nur, M. (2022). Flora composition and diversity in Mount Sibela Educational Forest, South Halmahera, North Maluku. *IOP Conference Series: Earth and Environmental Science*, 959(1), 012015. <https://doi.org/10.1088/1755-1315/959/1/012015>
- 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. <https://doi.org/10.1111/eva.12902>

- Iwanycki, N., Verstraete, B., Hassemer, G., Dunbar-Co, S., Hoggard, R., Meudt, H. M., & Rønsted, N. (2019). Ancestral range reconstruction of remote oceanic island species of *Plantago* (Plantaginaceae) reveals differing scales and modes of dispersal. *Journal of Biogeography*, 46(4), 706–722. <https://doi.org/10.1111/jbi.13525>
- Kakerissa, A. L., & Hahury, H. D. (2021). Pala Booi : Prospek pembangunan industri rumah tangga berbasis sumberdaya lokal. *ALE Proceeding*, 1, 122–128. <https://doi.org/10.30598/ale.1.2018.122-128>
- Kier, G., Kreft, H., Lee, T. M., Jetz, W., Ibsch, P. L., Nowicki, C., Mutke, J., & Barthlott, W. (2009). A global assessment of endemism and species richness across island and mainland regions. *Proceedings of the National Academy of Sciences*, 106(23), 9322–9327. <https://doi.org/10.1073/pnas.0810306106>
- Knowles, L. L., & Richards, C. L. (2005). Importance of genetic drift during Pleistocene divergence as revealed by analyses of genomic variation. *Molecular Ecology*, 14(13), 4023–4032. <https://doi.org/10.1111/j.1365-294X.2005.02711.x>
- Laenen, B., Désamoré, A., Devos, N., Shaw, A. J., González-Mancebo, J. M., Carine, M. A., & Vanderpoorten, A. (2011). Macaronesia: a source of hidden genetic diversity for post-glacial recolonization of western Europe in the leafy liverwort *Radula lindenbergiana*. *Journal of Biogeography*, 38(4), 631–639. <https://doi.org/10.1111/j.1365-2699.2010.02440.x>
- 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., Néron, 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>
- Leroy, S. A. G., Arpe, K., Mikolajewicz, U., & Wu, J. (2020). Climate simulations and pollen data reveal the distribution and connectivity of temperate tree populations in eastern Asia during the Last Glacial Maximum. *Climate of the Past*, 16(6), 2039–2054. <https://doi.org/10.5194/cp-16-2039-2020>
- 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
- Liu, S., Cornille, A., Decroocq, S., Tricon, D., Chague, A., Eyquard, J., Liu, W., Giraud, T., & Decroocq, V. (2019). The complex evolutionary history of apricots: Species divergence, gene flow and multiple domestication events. *Molecular Ecology*, 28(24), 5299–5314. <https://doi.org/10.1111/mec.15296>

- Lohman, D., de Bruyn, M., Page, T., von Rintelen, K., Hall, R., Ng, P., Shih, H.-T., Carvalho, G., & Rintelen, T. (2011). Biogeography of the Indo-Australian Archipelago. *Annual Review of Ecology Evolution and Systematics*, 42. <https://doi.org/10.1146/annurev-ecolsys-102710-145001>
- Lu, R.-S., Chen, Y., Tamaki, I., Sakaguchi, S., Ding, Y.-Q., Takahashi, D., Li, P., Isaji, Y., Chen, J., & Qiu, Y.-X. (2020). Pre-Quaternary diversification and glacial demographic expansions of *Cardiocrinum* (Liliaceae) in temperate forest biomes of Sino-Japanese Floristic Region. *Molecular Phylogenetics and Evolution*, 143, 106693. <https://doi.org/10.1016/j.ympev.2019.106693>
- Maydrawati, T. R. (2020). *The role of the Prior Informed Consent principle of the Nagoya Protocol in preventing biopiracy in the context of biodiversity protection* [Universitas Gadjah Mada]. <http://etd.repository.ugm.ac.id/penelitian/detail/183963>
- Memmo, V., Bertoni, C., Masini, M., Alvarez, J., Imran, Z., Echanove, A., & Orange, D. (2013). *Deposition and deformation in the recent Biak Basin (Papua Province, Eastern Indonesia)*. <https://doi.org/10.13140/2.1.1123.7120>
- Minter, M., O'Brien, D., Cottrell, J., Ennos, R., Hill, J. K., & Hall, J. (2021). Exploring the potential for 'Gene Conservation Units' to conserve genetic diversity in wild populations. *Ecological Solutions and Evidence*, 2(2). <https://doi.org/10.1002/2688-8319.12061>
- Mittermeier, R., Mittermeier, C. G., & Myers, N. (1999). *Hotspots: Earth's Biologically Richest and Most Endangered Terrestrial Ecoregions* (R. Mittermeier, Ed.; 1st ed., Vol. 1). Conservation International.
- Morrice, M. G., Jezek, P. A., Gill, J. B., Whitford, D. J., & Monoarfa, M. (1983). An introduction to the Sangihe arc: Volcanism accompanying arc—Arc collision in the Molucca Sea, Indonesia. *Journal of Volcanology and Geothermal Research*, 19(1–2), 135–165. [https://doi.org/10.1016/0377-0273\(83\)90129-4](https://doi.org/10.1016/0377-0273(83)90129-4)
- Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, 403(6772), 853–858. <https://doi.org/10.1038/35002501>
- Nagata, N., Toda, M., Ohbayashi, T., Hayashi, M., & Sota, T. (2021). Phylogeography of cicadas on continental and oceanic islands in the northwestern Pacific region. *Journal of Biogeography*, 48(12), 3060–3071. <https://doi.org/10.1111/jbi.14262>
- Orta, G. da. (1563). *Coloquios dos simples e drogas da India* (Vol. 1). Maxtor.

- Petit, R. J., & Hampe, A. (2006). Some Evolutionary Consequences of Being a Tree. *Annual Review of Ecology, Evolution, and Systematics*, 37(1), 187–214. <https://doi.org/10.1146/annurev.ecolsys.37.091305.110215>
- Piñeiro, R., Dauby, G., Kaymak, E., & Hardy, O. J. (2017). Pleistocene population expansions of shade-tolerant trees indicate fragmentation of the African rainforest during the Ice Ages. *Proceedings of the Royal Society B: Biological Sciences*, 284(1866), 20171800. <https://doi.org/10.1098/rspb.2017.1800>
- Pouteau, R., Trueba, S., Feild, T. S., & Isnard, S. (2015). New Caledonia: a Pleistocene refugium for rain forest lineages of relict angiosperms. *Journal of Biogeography*, 42(11), 2062–2077. <https://doi.org/10.1111/jbi.12581>
- Queiroz, A. (2005). The resurrection of oceanic dispersal in historical biogeography. *Trends in Ecology & Evolution*, 20(2), 68–73. <https://doi.org/10.1016/j.tree.2004.11.006>
- Quigley, D., MacNamara, L., & Gainey, P. (2020). First records of stranded Nutmegs *Myristica fragrans* Houttuyn, 1774 (Magnoliales: Myristicaceae) on the Irish Coast and a review of North Atlantic records. *Bulletin of the Irish Biogeographical Society*, 44(1), 178–188.
- Reed, D. H., & Frankham, R. (2003). Correlation between Fitness and Genetic Diversity. *Conservation Biology*, 17(1), 230–237. <https://doi.org/10.1046/j.1523-1739.2003.01236.x>
- Rico, Y., León-Tapia, M. Á., Zurita-Solís, M., Rodríguez-Gómez, F., & Vásquez-Morales, S. G. (2021). Influence of Pleistocene climatic oscillations on the phylogeography and demographic history of endemic vulnerable trees (section Magnolia) of the Tropical Montane Cloud Forest in Mexico. *PeerJ*, 9, e12181. <https://doi.org/10.7717/peerj.12181>
- 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>
- Sá, A. (1506). *Documentação para a historia das missões do padroado portugues do Oriente : Insulíndia* (Vol. 4).
- Salmona, J., Dresen, A., Ranaivoson, A. E., Manzi, S., Le Pors, B., Hong-Wa, C., Razanatsoa, J., Andriaholinirina, N. V., Rasoloharijaona, S., Vavitsara, M., & Besnard, G. (2023). How ancient forest fragmentation and riparian connectivity generate high levels of genetic diversity in a microendemic Malagasy tree. *Molecular Ecology*, 32(2), 299–315. <https://doi.org/10.1111/mec.16759>
- Sandoval-Castro, E., Peraza-Magallanes, A. Y., Dodd, R. S., Ashworth, V. E. T. M., Cruz-Mendivil, A., & Calderón-Vázquez, C. L. (2021). Exploring genetic diversity of lowland avocado (*Persea americana* Mill.) as a genetic reservoir

- for breeding. *Genetic Resources and Crop Evolution*, 68(7), 2757–2766. <https://doi.org/10.1007/s10722-021-01238-w>
- 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>
- 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. <https://doi.org/10.1007/s13595-015-0471-z>
- Selkoe, K. A., & Toonen, R. J. (2006). Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters*, 9(5), 615–629. <https://doi.org/10.1111/j.1461-0248.2006.00889.x>
- Siemaszko, M. (2015). *Altitudinal gradient and bird species distribution in the limestone forest of the Manusela National Park, Seram, Moluccas, Indonesia*.
- Souza, A., Giudicelli, G. C., Teixeira, M. C., Turchetto, C., Bonatto, S. L., & Freitas, L. B. (2022). Genetic diversity in micro-endemic plants from highland grasslands in southern Brazil. *Botanical Journal of the Linnean Society*, 199(1), 235–251. <https://doi.org/10.1093/botlinnean/boab057>
- Stuessy, T. F., Takayama, K., López-Sepúlveda, P., & Crawford, D. J. (2014). Interpretation of patterns of genetic variation in endemic plant species of oceanic islands. *Botanical Journal of the Linnean Society*, 174(3), 276–288. <https://doi.org/10.1111/boj.12088>
- Tallei, T. E., & Kolondam, B. J. (2015). DNA Barcoding of Sangihe Nutmeg (*Myristica fragrans*) using matK Gene. *HAYATI Journal of Biosciences*, 22(1), 41–47. <https://doi.org/https://doi.org/10.4308/hjb.22.1.41>
- Terhorst, J., & Song, Y. S. (2015). Fundamental limits on the accuracy of demographic inference based on the sample frequency spectrum. *Proceedings of the National Academy of Sciences*, 112(25), 7677–7682. <https://doi.org/10.1073/pnas.1503717112>
- Thornton, I. W. (1996). *Krakatau: the destruction and reassembly of an island ecosystem* (1st ed.). Harvard University Press.
- Tsuda, Y., Nakao, K., Ide, Y., & Tsumura, Y. (2015). The population demography of *Betula maximowicziana*, a cool-temperate tree species in Japan, in relation to the last glacial period: its admixture-like genetic structure is the result of simple population splitting not admixing. *Molecular Ecology*, 24(7), 1403–1418. <https://doi.org/10.1111/mec.13123>

- Udy, K., Fritsch, M., Meyer, K. M., Grass, I., Hanß, S., Hartig, F., Kneib, T., Kreft, H., Kukunda, C. B., Pe'er, G., Reininghaus, H., Tietjen, B., Tschardt, T., Waveren, C., & Wiegand, K. (2021). Environmental heterogeneity predicts global species richness patterns better than area. *Global Ecology and Biogeography*, 30(4), 842–851. <https://doi.org/10.1111/geb.13261>
- Vanden Abeele, S., Janssens, S. B., Piñeiro, R., & Hardy, O. J. (2021). Evidence of past forest fragmentation in the Congo Basin from the phylogeography of a shade-tolerant tree with limited seed dispersal: *Scorodophloeus zenkeri* (Fabaceae, Detarioideae). *BMC Ecology and Evolution*, 21(1), 50. <https://doi.org/10.1186/s12862-021-01781-1>
- Vellend, M., Verheyen, K., Flinn, K., Jacquemyn, H., Kolb, A., Van Calster, H., Peterk, G., Graae, B. J., Bellemare, J., Honnay, O., Brunet, J., Wulf, M., Gerhardt, F., & Hermy, M. (2007). Homogenization of forest plant communities and weakening of species-environment relationships via agricultural land use. *Journal of Ecology*, 95(3), 565–573. <https://doi.org/10.1111/j.1365-2745.2007.01233.x>
- Verbylaitė, R., Pliūra, A., Lygis, V., Suchockas, V., Jankauskienė, J., & Labokas, J. (2023). Genetic Diversity of Five Broadleaved Tree Species and Its Spatial Distribution in Self-Regenerating Stands. *Forests*, 14(2), 281. <https://doi.org/10.3390/f14020281>
- Veresoglou, S. D., & Halley, J. M. (2018). Seed mass predicts migration lag of European trees. *Annals of Forest Science*, 75(3), 86. <https://doi.org/10.1007/s13595-018-0766-y>
- Voigt, M., Supriatna, J., Deere, N. J., Kastanya, A., Mitchell, S. L., Rosa, I. M. D., Santika, T., Siregar, R., Tasirin, J. S., Widyanto, A., Winarni, N. L., Zakaria, Z., Mumbunan, S., Davies, Z. G., & Struebig, M. J. (2021). Emerging threats from deforestation and forest fragmentation in the Wallacea centre of endemism. *Environmental Research Letters*, 16(9), 094048. <https://doi.org/10.1088/1748-9326/ac15cd>
- Wallace, A. R. (1886). *The Malay Archipelago : the land of the orang-utan and the bird of paradise ; a narrative of travel, with studies of man and nature*. Macmillan. <https://www.biodiversitylibrary.org/item/133732>
- Wernberg, T., Coleman, M. A., Bennett, S., Thomsen, M. S., Tuya, F., & Kelaher, B. P. (2018). Genetic diversity and kelp forest vulnerability to climatic stress. *Scientific Reports*, 8(1), 1851. <https://doi.org/10.1038/s41598-018-20009-9>
- White, L., Webb, M., Jost, B., Gold, D., Gunawan, I., & Tiranda, H. (2019). *New insights into the geological evolution of West Papua from recent field and laboratory studies*.

World Conservation Monitoring Centre. (1998). *Myristica fragrans*. The IUCN Red List of Threatened Species 1998: E.T33986A9820569. <https://dx.doi.org/10.2305/IUCN.UK.1998.RLTS.T33986A9820569.en>

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>

Appendixes

Appendix 1. Published article in Molecular Biology reports (Chapter 1)

Appendix 2. Published article in Mitochondrial DNA Part B (Chapter 2)

Appendix 3. Poster presentation in 4th World Agroforestry Conference, France 2019.



Microsatellite markers development for Indonesian nutmeg (*Myristica fragrans* Houtt.) and transferability to other *Myristicaceae* spp.

Jakty Kusuma^{1,2} · Nora Scarcelli² · Marie Couderc² · Cédric Mariac² · Leila Zekraoui² · Jérôme Duminil²

Received: 2 April 2020 / Accepted: 14 May 2020
© Springer Nature B.V. 2020

Abstract

Myristica fragrans (Myristicaceae) is a tropical evergreen tree that yields the two famous spices: nutmeg and mace. Despite its socio-economic importance, the spatial distribution of its genetic diversity is barely documented. In this aim, 48 nuclear microsatellite markers were isolated of which 14 were polymorphic in *M. fragrans*. Number of alleles per locus ranged from 2 to 6. The level of observed heterozygosity ranged from 0.038 to 0.929 across loci. Transferability of these microsatellites in other *Myristica* species (*M. fatua*, *M. argentea*, and *M. crassipes*) and *Myristicaceae* species (*Horsfieldia palauensis*) was tested and successful. These new microsatellites will be useful for future investigation on genetic diversity and population structure of *M. fragrans* and phylogenetically-related species.

Keywords *Myristica* · Nutmeg · Microsatellites · Fruit tree species · Spices

Introduction

Myristica fragrans Houtt, belongs to the Myristicaceae family. This dioecious and entomophilous [1] tropical tree is considered as one of the most important food tree species to Indonesian people. It is utilized as spices, medicine [2] and raw material for perfume industries [3]. The species

is most probably native from Banda islands in the Moluccas (the so-called ‘spice islands’) in Indonesia [4]. Given its economic value, the species has been widely propagated and cultivated across tropical regions, particularly in India, Malaysia, Sri Lanka and the Caribbean islands [5]. Despite its socio-economic importance, its conservation status is not known (IUCN red list status: ‘data deficient’) [6]. Moreover, the spatial distribution of its genetic diversity remains poorly known, which impedes the development of conservation and sustainable management and use strategies. In order to achieve such objectives, polymorphic molecular markers need to be available. In this study, 14 polymorphic nuclear microsatellite markers were developed and characterized. Additionally, cross-amplification of these markers was tested in *M. fatua*, *M. argentea*, *M. crassipes* and another Myristicaceae species, *Horsfieldia palauensis*.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11033-020-05535-y>) contains supplementary material, which is available to authorized users.

✉ Jakty Kusuma
kusuma@polinela.ac.id
Nora Scarcelli
nora.scarcelli@ird.fr
Marie Couderc
marie.couderc@ird.fr
Cédric Mariac
cedric.mariac@ird.fr
Leila Zekraoui
leila.zekraoui@ird.fr
Jérôme Duminil
jerome.duminil@ird.fr

¹ Department of Plant Science, Politeknik Negeri Lampung, Lampung, Indonesia

² DIADE, IRD, Univ Montpellier, Montpellier, France

Methods and results

Plant material and extraction

A total of 75 individuals of *M. fragrans* were sampled in three different populations in the Moluccas, Indonesia (three different islands: Ambon, Banda Besar, Banda Neira; details in Online Appendix 1). In order to test the transferability of these markers to other *Myristica* species and another

Myristicaceae genus, we further sampled individuals of *M. fatua* (five individuals) and *M. argentea* (five individuals) in the field, as well as *M. crassipes* (two individuals) and *Horsfieldia palauensis* (two individuals) from the New York Botanical Garden (NY) and the Herbar de l'Université de Montpellier (MPU) (Online Appendix 1). All DNA were extracted following the protocol of Mariac et al. [7].

Microsatellite development

A genomic library was prepared following Mariac et al. [8] on one *M. fragrans* sample (JK 131, Online Appendix 1), and then paired-end sequenced using Illumina MiSeq v2 reagents and 2 × 250 bp at the CIRAD Genotyping Platform (Montpellier, France). Script used for following activities are available on Online Appendix 2. Illumina adapters were removed using CUTADAPT 1.8 [9] with parameters as follow: overlap length = 7, minimum length = 35 and quality = 20. Reads under a quality threshold of 30 were discarded using a PERL script (https://github.com/SouthGreenPlatform/arcad-hts/blob/master/scripts/arcad_hts_2_Filter_Fastq_On_Mean_Quality.pl). R1 and R2 reads were then paired using a PERL script (https://github.com/SouthGreenPlatform/arcad-hts/blob/master/scripts/arcad_hts_3_synchronized_paired_fastq.pl). We obtained 634,906 paired reads that were then joined using FLASH v1.2.11 [10] and produced a total of 182,869 joined reads. Microsatellite repeat units were identified and primers were designed using default setting in QDD software v3.1.2 [11]. From the 2374 microsatellite identified, 48 loci were selected according to the following criteria: (1) perfect (i.e., not compound) microsatellite di- or trinucleotide repeat units; (2) repeats from 10 to 20; (3) primers situated at least 20 bp from the microsatellite region; (4) PCR product length between 90 and 300 bp.

Microsatellite amplification test

The 48 selected primer pairs were tested using three samples of *M. fragrans* (samples JK 131, JK 161, and JK 216; Online Appendix 1). A labeled tail (with either 6-FAM, NED, VIC or PET dye) was added to the 5' end of the forward primer of each locus in order to allow co-amplification of the loci in multiplex (Table 1). PCR reactions were done in a total volume of 10 µL as follows: 2 µL of 5 × Colorless GoTaq® buffer (Promega, France) with final concentration at 1 × (1.5 mM MgCl₂), 0.2 µL of dNTP (Promega, France) at 10 mM, 0.05 µL of GoTaq® polymerase (Promega, France) at 5 U/µL, 0.5 µL of each primer at 10 mM, 2 µL DNA template (of ca. 15–20 ng/µL), and water qs 10 µL. PCR conditions were as follows: initial denaturation at 95 °C (5 min), 30 cycles of denaturation at 95 °C (23 s), annealing at 57 °C (1 min 30 s), elongation at 72 °C (1 min), and a final extension step at 72 °C (10 min). The quality of the PCR products

was checked using a QIAxcel Advanced system (QIAGEN). Only the 40 loci showing consistent amplification in all three samples of *M. fragrans* were kept for the following analyses.

The polymorphism of the 40 loci was first tested on six *M. fragrans* individuals, i.e. on two individuals per population (three populations corresponding to three different islands; see details in Online Appendix 1). PCR amplifications were done in four multiplexes (details not shown) in a total volume of 10 µL for each multiplex: 0.15 µL of fluorescent labeled primers [12] at 10 µM, 0.15 µL of the reverse and 0.10 µL of the forward at 10 µM for both primers, 7.5 µL Type-it Microsatellite PCR Master Mix (Qiagen), 1 µL DNA template and water qs 10 µL. PCR conditions were identical to the amplification test (see above). The 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). Peaks on each locus were analyzed using Geneious Prime® 2019.2.3 [13]. Loci that were difficult to read as well as monomorphic loci were excluded. In total, 14 primer pairs were kept and were combined in two multiplex reactions (the list of loci grouped in each multiplex is provided in Table 2) using Multiplex Manager 1.0 software [14].

Microsatellite diversity analysis

An analysis of genetic diversity was then performed on a larger set of samples of the three *M. fragrans* populations (22 to 29 samples per population; Online Appendix 1). The 14 primer pairs were amplified in two multiplexes (details provided in Table 2) in a final reaction volume of 15 µL, consisted of 0.15 µL of fluorescent labeled primers at 10 µM, 0.15 µL of the reverse and 0.10 µL of the forward at 10 µM for both microsatellite primers, 7.5 µL Type-it Multiplex PCR Master Mix (Qiagen), 1 µL DNA template and water qs 10 µL. PCR conditions were identical to the amplification test (see above).

Genotypes were analyzed locus per locus using the automatic procedure implemented in Geneious Prime® 2019.2.3 [13] and manually corrected. Locus Myr42 showed two duplicated loci and was divided into two different loci (Myr42a and Myr42b) to be included in the following analyses. The genetic diversity was evaluated in each population by estimating the number of alleles per locus (*A*), expected and observed heterozygosity (*H_E* and *H_O*, respectively), fixation index *F_{IS}*, and the deviations from Hardy–Weinberg equilibrium by using the program SPAGeDi 1.5 [15], whereas frequencies of null alleles (*r*) were calculated using INEST 2.2 [16]. Tests for genotypic linkage disequilibrium were calculated population per population using Genepop 4.2 [17].

Table 1 Characteristics of 14 microsatellite loci developed in *Myristica fragrans*

Locus	Primer sequences (5'–3') ^a	Labelled primer	T _m (°C)	Allele size range (bp)	Genebank accession
Myr10	F: <u>TAGGAGTGCAGCAAGC</u> ATTACTCCCGCACCAC TAAA R: AGCAGAACCTACAAGAAGCAA	NED	57	132–164	MT175707
Myr12	F: CTAGTTATTGCTCAGCGGTAGAGTGACGGGTTGT ATGCA R: AGGAATGACGATCGAGGATAGA	PET	57	155–161	MT175708
Myr13	F: <u>TGTA</u> AAACGACGGCCAGTCCTTTCACATGGGTT GAAATGC R: TCCATCAAGTCCGCAGAACG	6-FAM	57	146–164	MT175709
Myr18	F: <u>TAGGAGTGCAGCAAGC</u> ATTACGCCGATTGGACC TCATT R: ACTCGTCTGGATGACTACAGC	NED	57	173–191	MT1757010
Myr23	F: <u>CACTGCTTAGAGCGATGC</u> GCCTCTGAAACCCG TTGTA R: TGATGAGAGCTACTGGCCTTG	VIC	57	153–175	MT1757011
Myr26	F: <u>TAGGAGTGCAGCAAGC</u> ATAGTAAGTGTATCCAG CCGCC R: CGACCTTTCGTCTCCTGGTA	NED	57	173–185	MT1757012
Myr27	F: <u>CACTGCTTAGAGCGATGC</u> AGTGGGTTGGCGATT TAGGT R: TGGCCAAGGAGATACCATTCT	VIC	57	175–197	MT1757013
Myr28	F: <u>CTAGTTATTGCTCAGCGGT</u> TCCAAGATCCACAGA CCAGC R: CAAAGCGAACGAACCCGAG	PET	57	169–203	MT1757014
Myr29	F: <u>TGTA</u> AAACGACGGCCAGTAGTTGCAATTGAGC TCCAGA R: ATGAAGCACACGCACATTGT	6-FAM	57	172–198	MT1757015
Myr32	F: <u>CTAGTTATTGCTCAGCGGT</u> TGGCCATTAACACAT CACTCCA R: GCATCCAAAACAGAGCCTCAG	PET	57	187–211	MT1757016
Myr36	F: <u>CTAGTTATTGCTCAGCGGT</u> TGGGCAAACCTTTCA AGGACA R: CTAGGGCATCTGCTGGTGAG	PET	57	218–242	MT1757017
Myr42 ^b	F: <u>TAGGAGTGCAGCAAGC</u> ATCAACTGTGGGATCTG CCAGA R: GGCCGCTTGACTGTCAACTT	NED	57	223–237 241–251	MT1757018
Myr43	F: <u>CACTGCTTAGAGCGATGC</u> CACATGCAGGTAGTC CTCGT R: ACGCTAGAGATGGTTGCCAG	VIC	57	214–242	MT1757019
Myr47	F: <u>CACTGCTTAGAGCGATGC</u> GGATCACGATATAAA TGGCAGC R: GGGTATAAACGGGAAAGGGC	VIC	57	247–269	MT1757020

^aThe universal linkers attached to the forward primers are underlined (TGTA³AAACGACGGCCAGT, [19]; TAGGAGTGCAGCAAGCAT, CAC TGCTTAGAGCGATGC and CTAGTTATTGCTCAGCGGT Culley et al. [12]);

^bDuplicated locus

For each locus, the number of alleles ranged between two to six in all populations investigated (Table 2). Observed heterozygosity of the 14 loci on Ambon, Banda Besar, and Banda Neira population varied between 0.038 to 0.769, 0.227 to 0.810, and 0.333 to 0.929, respectively. The mean observed heterozygosity value per population was 0.487 in Ambon, 0.582 in Banda Besar, and 0.595 in Banda Neira (Table 2). Some loci showed an excess of homozygosity (F_{IS} significantly greater than zero), but this was mainly due to

the presence of null alleles. One out of the three populations was not in Hardy–Weinberg equilibrium (Ambon; Table 1). *Myristica fragrans* is a dioecious cultivated species, and males are frequently cut down to reach an approximate ratio of ten females for one male. Such management practices can have detrimental effects on population panmixia. Genotypic linkage disequilibria were found in seven pairs of loci in the Ambon population, seven pairs in the Banda Besar population, and in five pairs in the Banda Neira population.

Table 2 Results of initial screening of 14 microsatellite loci developed in *Myristica fragrans*

Multiplex	Locus	Ambon (N=26)					Banda Besar (N=22)					Banda Naira (N=27)					Total				
		A	H _o	H _e	F _{IS}	r	A	H _o	H _e	F _{IS}	r	A	H _o	H _e	F _{IS}	r	A	H _o	H _e	F _{IS}	r
Multiplex-1	Myr10	3	0.308	0.347	0.117	0.01 ± 0.03	2	0.273	0.304	0.106	0.02 ± 0.06	4	0.444	0.533	0.169	0.06 ± 0.06	4	0.240	0.284	0.107	
	Myr12	2	0.269	0.237	-0.136	0.00 ± 0.00	2	0.591	0.459	-0.294	0.00 ± 0.00	2	0.370	0.452	0.185	0.05 ± 0.06	2	0.277	0.273	-0.016	
	Myr23	5	0.577	0.671	0.144	0.06 ± 0.05	5	0.773	0.765	-0.010	0.00 ± 0.00	6	0.741	0.795	0.071	0.00 ± 0.01	6	0.481	0.517	0.107	
	Myr26	4	0.654	0.578	-0.133	0.00 ± 0.00	4	0.273	0.324	0.163	0.05 ± 0.06	4	0.593	0.584	-0.015	0.00 ± 0.00	4	0.361	0.364	0.009	
	Myr29	4	0.577	0.696	0.174	0.04 ± 0.06	5	0.864	0.745	-0.163	0.00 ± 0.00	5	0.929	0.719	-0.295**	0.00 ± 0.00	6	0.546	0.507	-0.077	
	Myr32	4	0.462	0.676	0.322*	0.11 ± 0.05	3	0.455	0.512	0.116	0.02 ± 0.06	3	0.778	0.544	-0.441*	0.00 ± 0.00	5	0.397	0.413	0.038	
Multiplex-2	Myr43	4	0.615	0.670	0.084	0.03 ± 0.05	6	0.773	0.746	-0.036	0.00 ± 0.00	6	0.778	0.735	-0.059	0.00 ± 0.00	8	0.501	0.497	-0.005	
	Myr13	4	0.615	0.684	0.102	0.05 ± 0.06	4	0.636	0.737	0.140	0.00 ± 0.00	4	0.704	0.712	0.012	0.00 ± 0.00	4	0.453	0.490	0.076	
	Myr18	2	0.038	0.038	-0.000	0.00 ± 0.00	2	0.386	0.371	0.008	0.28 ± 0.10	3	0.333	0.344	0.034	0.25 ± 0.09	3	0.161	0.174	0.079	
	Myr27	3	0.731	0.584	-0.257	0.00 ± 0.00	4	0.636	0.627	-0.014	0.00 ± 0.00	3	0.741	0.677	-0.096	0.00 ± 0.00	4	0.490	0.442	-0.109	
	Myr28	5	0.769	0.677	-0.139	0.00 ± 0.00	6	0.810	0.768	-0.054	0.08 ± 0.07	6	0.630	0.690	0.090	0.04 ± 0.04	7	0.506	0.517	0.021	
	Myr36	3	0.231	0.548	0.584**	0.20 ± 0.05	4	0.227	0.700	0.681**	0.27 ± 0.04	4	0.481	0.724	0.340**	0.13 ± 0.05	4	0.222	0.471	0.368**	
	Myr42a	3	0.385	0.561	0.320	0.09 ± 0.05	3	0.727	0.576	-0.270	0.00 ± 0.00	3	0.481	0.552	0.131	0.04 ± 0.06	3	0.161	0.392	0.079	
	Myr42b	3	0.500	0.518	0.036	0.00 ± 0.00	3	0.636	0.586	-0.087	0.00 ± 0.00	3	0.556	0.634	0.127	0.05 ± 0.05	3	0.361	0.402	0.033	
	Myr47	4	0.577	0.595	0.031	0.02 ± 0.04	4	0.628	0.606	-0.127	0.00 ± 0.00	3	0.370	0.469	0.215	0.08 ± 0.06	4	0.370	0.383	0.034	
	Multilocus average		3.53	0.487	0.539	0.098*	0.04 ± 0.02	3.80	0.582	0.589	0.013	0.04 ± 0.02	3.93	0.595	0.611	0.027	0.04 ± 0.03	4.47	0.384	0.408	0.061

N number of individuals sampled, A number of alleles, H_o observed heterozygosity, H_e expected heterozygosity, F_{IS} inbreeding coefficient, r frequency of null alleles
Significant deviation from Hardy-Weinberg equilibrium at *p < 0.05, **p < 0.01

Table 3 Alleles size ranges observed in cross-amplification of microsatellite loci isolated from *Myristica fragrans* and tested in four additional Myristicaceae species

Locus	<i>M. fatua</i> (N=5)	<i>M. argentea</i> (N=5)	<i>M. crassipes</i> (N=2)	<i>H. palauensis</i> (N=2)
Myr10	126–154 bp	132–162 bp	134–152 bp	128 bp
Myr12	155–161 bp	153–161 bp	155–161 bp	155 bp
Myr23	153–173 bp	165–189 bp	149–171 bp	175 bp
Myr26	177–185 bp	177–185 bp	181–191 bp	181–183 bp
Myr29	182–198 bp	178–194 bp	172–182 bp	174–190 bp
Myr32	193–209 bp	195–245 bp	193–199 bp	–
Myr43	214–240 bp	214–246 bp	218–224 bp	216–236 bp
Myr13	146–160 bp	146–160 bp	146–152 bp	148–150 bp
Myr18	173–191 bp	175–183 bp	173–183 bp	173 bp
Myr27	179–191 bp	179–197 bp	–	185–205 bp
Myr28	169–203 bp	191–207 bp	–	–
Myr36	202–238 bp	228–242 bp	–	–
Myr42a	223–235 bp	223–233 bp	217–235 bp	205–223 bp
Myr42b	241–245 bp	235–245 bp	–	245–251 bp
Myr47	247–267 bp	263–275 bp	269–285 bp	251–273 bp

N number of individuals tested; – no amplification

However, only one locus pair remained significant (Online Appendix 3), once applied a sequential Holm–Bonferroni correction for multiple comparisons [18].

Transferability of the markers in *M. fatua*, *M. argentea*, *M. crassipes*, and *H. palauensis*

Cross-amplification of the markers was tested in three other *Myristica* species (*M. fatua*, *M. argentea*, and *M. crassipes*) and another Myristicaceae species (*Horsfieldia palauensis*), using the same PCR conditions as before.

Fourteen individuals from five different species was successfully genotyped (Table 3). All of these markers were found polymorphic in species *M. fatua* and *M. argentea*. Four and three markers did not amplify respectively in *M. crassipes* and *H. palauensis*, although the number of samples were low. Moreover, four loci were monomorphic in *H. palauensis*. These scores of transferability suggest that *M. fatua* and *M. argentea* are more phylogenetically related to *M. fragrans* than *M. crassipes*.

Conclusion

Fourteen polymorphic microsatellite markers were developed for an economically important food tree species, *Myristica fragrans*. These microsatellite markers will be useful for assessing species' genetic diversity and population structure as well in phylogenetically-related species given the high transferability of these markers.

Acknowledgements The authors would like to thank Gregory M. Plunkett, Sean Thackurdeen and Michael J. Balick from New York Botanical Garden (NY) for having provided samples of *H. palauensis*. The authors also thank Herbar de l'Université de Montpellier (MPU) through Jérôme Munzinger for having provided samples of *M. crassipes*. Authors also greatly appreciate the support from the Indonesia Endowment Fund for Education (LPDP). This project was supported by Agropolis Fondation under the reference ID 1502-503 and 1502-504 through the « Investissements d'avenir » programme (Labex Agro: ANR-10-LABX-0001-01), under the frame of I-SITE MUSE (ANR-16-IDEX-0006), Indonesian Endowment Fund For Education (Grant No. 0003719/AFR/D/BUDI-2018).

Author Contributions All authors contributed to the study. Material preparation and data collection were performed by JK and JD. JK, MC, LZ, and CM undertook the experiment. JK, JD and NS analyzed the data. The first draft of the manuscript was written by JK and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Data availability DNA sequences of microsatellites described in this paper were deposited in GenBank (Table 1).

Code availability Scripts used in this paper were described in Online Appendix 2.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethics approval The authors declare no ethics approval needed.

References

- Sharma MV, Armstrong JE (2013) Pollination of *Myristica* and other nutmegs in natural populations. *Trop Conserv Sci* 6(5):595–607
- Arumugam G, Purushotham B, Swamy MK (2019) *Myristica fragrans* Houtt.: botanical, pharmacological, and toxicological aspects. In: Swamy MK, Akhtar MS (eds) *Natural bio-active compounds*. Springer, Singapore, pp 81–106
- Periasamy G, Karim A, Gibrelibanos M, Gebremedhin G, Gilani AH (2016) Nutmeg (*Myristica fragrans* Houtt.) oils. In: Preedy VR (ed) *Essential oils in food preservation, flavor and safety*. Academic Press, New York, pp 607–616
- Rumphius GE (1741–1750). *Herbarium Amboinense ... pars prima (-sexta)*. In: Burman J (ed) Amsterdam: Franciscum Changuion, Joannem Catuffe, Hermannum Uytwerf.
- Swetha VP, Parvathy VA, Sheeja TE, Sasikumar B (2017) Authentication of *Myristica fragrans* Houtt. using DNA barcoding. *Food Control* 73:1010–1015
- World Conservation Monitoring Centre. 1998. *Myristica fragrans*. *The IUCN Red List of Threatened Species* 1998: e.T33986A9820569. <https://doi.org/10.2305/IUCN.UK.1998.RLTS.T33986A9820569.en>. Accessed 29 Apr 2020
- Mariac C, Luong V, Kapran I et al (2006) Diversity of wild and cultivated pearl millet accessions (*Pennisetum glaucum* [L.] R.Br.) in Niger assessed by microsatellite markers. *Theor Appl Genet* 114(1):49–58
- Mariac C, Scarcelli N, Pouzadou J et al (2014) Cost-effective enrichment hybridization capture of chloroplast genomes at deep

- multiplexing levels for population genetics and phylogeography studies. *Mol Ecol Resour* 14(6):1103–1113
9. Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* 17(1):10
 10. Magoc T, Salzberg SL (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 27(21):2957–2963
 11. Megléc E, Costedoat C, Dubut V et al (2010) QDD: a user-friendly program to select microsatellite markers and design primers from large sequencing projects. *Bioinformatics* 26(3):403–404
 12. Culley TM, Weller SG, Sakai AK, Putnam KA (2008) Characterization of microsatellite loci in the Hawaiian endemic shrub *Schiedea adamantis* (Caryophyllaceae) and amplification in related species and genera. *Mol Ecol Resour* 8(5):1081–1084
 13. Kearse M et al (2012) Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28(12):1647–1649
 14. Holleley CE, Geerts PG (2009) Multiplex Manager 1.0: a cross-platform computer program that plans and optimizes multiplex PCR. *Biotechniques* 46(7):511–517
 15. Hardy OJ, Vekemans X (2002) Spagedi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol Ecol Notes* 2(4):618–620
 16. Chybicki IJ, Burczyk J (2009) Simultaneous estimation of null alleles and inbreeding coefficients. *J Hered* 100(1):106–113
 17. Rousset F (2008) Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Mol Ecol Resources* 8:103–106
 18. Rice WR (1989) Analyzing tables of statistical tests. *Evolution* 43(1):223–225
 19. Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. *Nat Biotechnol* 18:233–234

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Mitochondrial DNA Part B Resources

ISSN: (Print) (Online) Journal homepage: <https://www.tandfonline.com/loi/tmdn20>

Complete chloroplast genome of two nutmeg species, *Myristica argentea* Warb. 1891 and *Myristica fatua* Houtt. 1774 (Myristicaceae)

Jakty Kusuma, Marie Couderc, Nora Scarcelli & Jérôme Duminil

To cite this article: Jakty Kusuma, Marie Couderc, Nora Scarcelli & Jérôme Duminil (2023) Complete chloroplast genome of two nutmeg species, *Myristica argentea* Warb. 1891 and *Myristica fatua* Houtt. 1774 (Myristicaceae), *Mitochondrial DNA Part B*, 8:7, 751-755, DOI: [10.1080/23802359.2023.2233154](https://doi.org/10.1080/23802359.2023.2233154)

To link to this article: <https://doi.org/10.1080/23802359.2023.2233154>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 19 Jul 2023.



[Submit your article to this journal](#)







[View related articles](#)



[View Crossmark data](#)

Complete chloroplast genome of two nutmeg species, *Myristica argentea* Warb. 1891 and *Myristica fatua* Houtt. 1774 (Myristicaceae)

Jakty Kusuma^{a,b} , Marie Couderc^a , Nora Scarcelli^a  and Jérôme Duminil^a 

^aDIADÉ, University of Montpellier, IRD, CIRAD, Montpellier, France; ^bPoliteknik Negeri Lampung, Lampung, Indonesia

ABSTRACT

Myristica argentea Warb. 1891 and *M. fatua* Houtt. 1774 are two South-East Asian food tree species. They are harvested from the wild or cultivated for local uses as a condiment (nutmeg and mace), medicine, and source of wood. In this study, we reconstructed the complete chloroplast (cp) genomes of these two species from whole genome sequencing data using the Illumina NovaSeq platform. The genome sizes of *M. argentea* and *M. fatua* were respectively 155,871 base pairs (bp) and 155,898 bp, including 126 genes and an overall GC content of 39.20% in both species. Our study provides useful resources for future evolutionary research and diversity analysis of *Myristica* species.

ARTICLE HISTORY

Received 10 January 2023
Accepted 29 June 2023

KEYWORDS

Chloroplast; *Myristica*; phylogeography; phylogenetic; nutmeg

Introduction

The *Myristica* genus contains 172 species that are mainly distributed in Tropical Southeast Asia, the Malay Archipelago

(Malesia) and the western Pacific (de Wilde 2014; Govaerts et al. 2021). *Myristica argentea* Warb. 1891, known as Papuan nutmeg or Fakfak nutmeg, is endemic to Papua island,

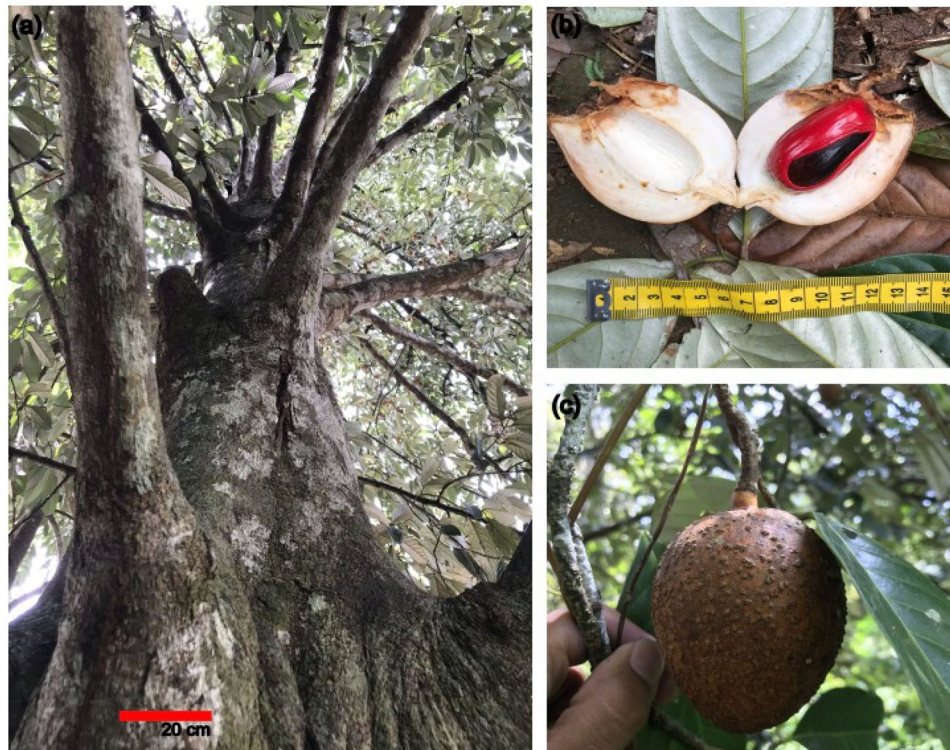





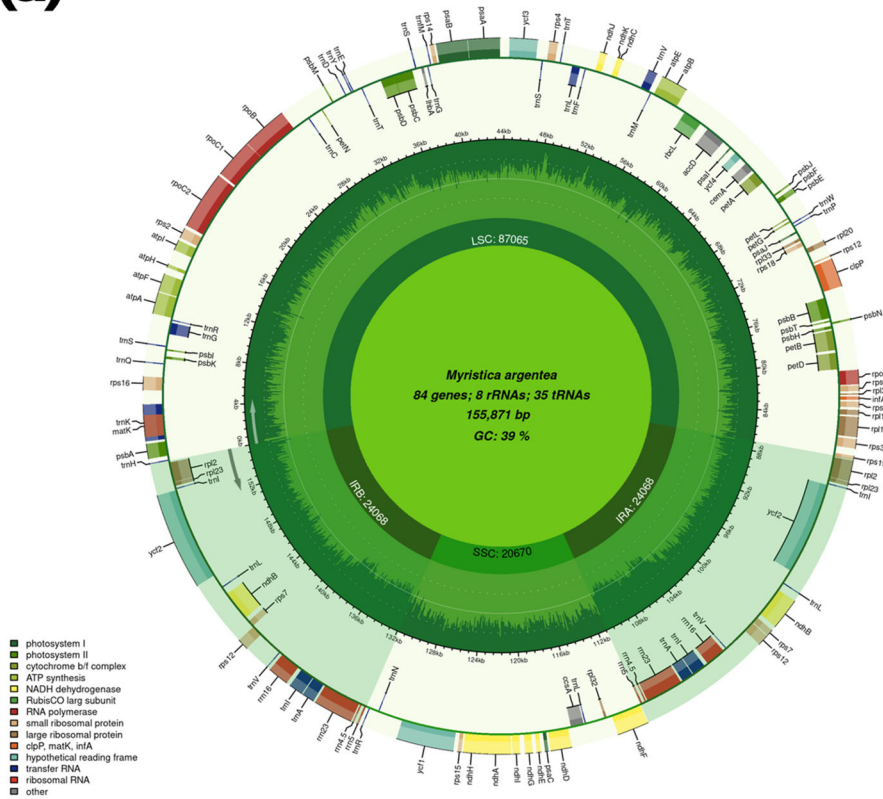
Figure 1. *Myristica argentea* grown in the forest of West Papua, Indonesia. (A) Tree trunk. (B) Mature fruit, with a description as follows: fruits solitary, ±ellipsoid-fusiform, 6.5–7.5 by 3.5–4.5 cm, apex narrowed, base somewhat contracted into a pseudostalk, 3–8 mm long, early glabrescent, brown, usually set with conspicuous coarse pale pustules or lenticels, similar as reported by de Wilde (2014). (C) Longitudinal cut of the fruit showing red mace and nut (C). Pictures by Jakty Kusuma, Ambon, Indonesia, 2018. For the figure of *Myristica fatua*, picture is available in <https://sweetgum.nybg.org/images3/3009/671/03509312.jpg>.

CONTACT Jakty Kusuma  jakty.kusuma@ird.fr  DIADÉ, University of Montpellier, IRD, CIRAD, Montpellier, France; Politeknik Negeri Lampung, Lampung, Indonesia

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/23802359.2023.2233154>.

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

(a)



(b)

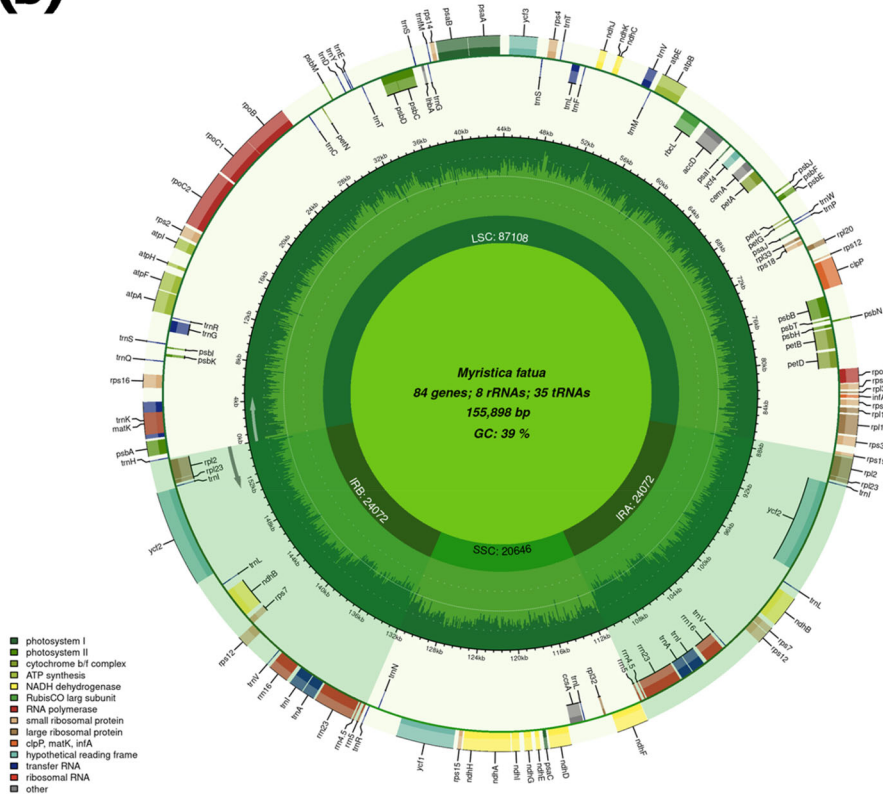


Figure 2. Annotation of the *M. argentea* (a) and *M. fatua* (b) chloroplast genomes inferred by CPGAVAS2. The outer circle represents the genes annotated, classified by color based on their function. The length of genome is presented in the inner circle. Large single-copy (LSC), small single-copy (SSC), and inverted repeat regions (IRA and IRB) are marked.

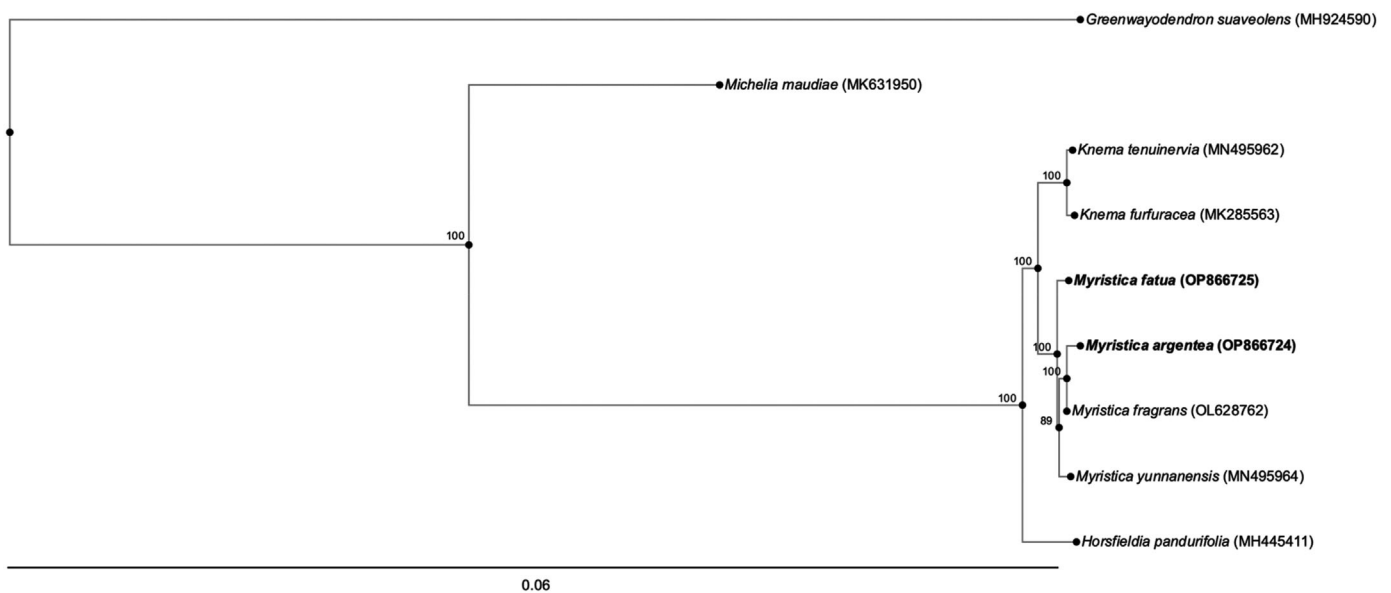


Figure 3. The best maximum-likelihood phylogenetic tree reconstructed using RAxML ver. 7.2.7 based on complete chloroplast genomes. The following sequences were used: *M. fragrans* (OL628762 (Cai et al. 2021)), *M. yunnanensis* (MN495964 (Mao et al. 2019)), *Knema tenuinervia* (MN495962 (Cai et al. 2021)), *Knema furfuracea* (MK285563 (Yang et al. 2019)), *Horsfieldia pandurifolia* (MH445411 (Zhang et al. 2019)), as well as *Michelia maudiae* (MK631950 (Wang et al. 2019)) and *Greenwayodendron suaveolens* (MH924590 (Migliore et al. 2018)) as outgroups. Numbers on branches indicate bootstrap values.

Indonesia, and is named after the region where the species occurs. The species is utilized for spices and medicinal purposes, by using the seed and its aril. In contrast with the relatively narrow distribution of *M. argentea*, *Myristica fatua* Houtt. 1774 presents a large distribution, from the Malesian Archipelago to the Pacific islands (Lim 2012). Their distribution in these archipelagos makes them compelling subjects for investigating the biogeography of island plants in the region. Since their evolutionary history remains largely undocumented, the reconstruction of their chloroplast (cp) genome will help conducting phylogeographic studies.

Materials

Leaf samples of *M. fatua* originated from Tafea, Vanuatu (19° 35' 52.8" S, 169° 20' 45.6" E) were collected from New York Botanical Garden herbaria. A specimen was deposited at New York Botanical Garden (<https://sweetgum.nybg.org/science/vh/specimen-details/?irn=3167800>; Gregory M. Plunkett, gplunkett@nybg.org) under the voucher number 03509312. Further, leaf samples of *M. argentea* (Figure 1) from a silica-dried sample were collected in Ambon, Indonesia (3° 44' 38.112" S, 127° 55' 47.568" E). A specimen of *M. argentea* was deposited at UMR DIADE, IRD, France (https://dataverse.ird.fr/dataverse/umr_diade; jakty.kusuma@ird.fr) under the number JK0155.

Methods

Genomic DNA was extracted using the MATAB protocol following Mariac et al. (2006). We constructed genomic libraries following protocol described in Mariac et al. (2014). Libraries generated from real-time PCR were then sequenced on an Illumina NovaSeq (Illumina, San Diego, CA). Paired-end reads were filtered as described in Scarcelli et al. (2016), and aligned

to verified *M. fragrans* whole cp genome (OL628762). Cleaned reads were then assembled using the *de novo* assembler GetOrganelle ver.1.7.6.1 (Jin et al. 2020) (Supplementary Methods S1), and annotated using CPGAVAS2 (<http://www.herbalgenomics.org/cpgavas2/>) (Shi et al. 2019). We visualized the circular genome using CHLOROPlot (<https://irscope.shinyapps.io/ChloroPlot/>) (Zheng et al. 2020). All genomic information was deposited in GenBank (see Data availability statement).

We then constructed a phylogenetic tree including the complete cp genome sequences of these two species as well as seven available complete cp genome sequences of closely related species of the family Myristicaceae (*M. fragrans* OL628762 (Cai et al. 2021); *M. yunnanensis* MN495964 (Mao et al. 2019); *Knema tenuinervia* MN495962 (Cai et al. 2021); *Knema furfuracea* MK285563 (Yang et al. 2019); and *Horsfieldia amygdalina* MK285561 (Zhang et al. 2019)), a species from Magnoliaceae (*Michelia maudiae* MK631950 (Wang et al. 2019)), and from Annonaceae, *Greenwayodendron suaveolens* MH924590 (Migliore et al. 2018) as an outgroup. MAFFT (<https://mafft.cbrc.jp/alignment/server/index.html>) ver. 7 (Kato et al. 2019) and RAxML ver. 7.2.7 (<https://cme.h-its.org/exelixis/php/countSource727.php>) (Stamatakis 2014) were used to align the whole cpDNA sequences and infer a phylogenetic tree (Supplementary Methods S2).

Results

The full-length cp genome of *M. argentea* is 155,871 bp, while *M. fatua* is 155,898 bp and they exhibit respective assembly coverages of 41.2× and 31.8× (Figure 2 and Supplementary Figure S1). The cpDNA genomes include a large single-copy (LSC) region (87,065 bp and 87,108 bp), a small single-copy (SSC) region (20,670 bp and 20,646 bp), and a pair of inverted repeat (IR) region (24,068 bp and 24,072 bp) in *M. argentea*

and *M. fatua*, respectively. For both species, we identified 126 genes corresponding to 83 protein-coding genes, 35 tRNA, and eight rRNA, less genes in comparison with *M. fragrans* (146 genes) (de Oliveira et al. 2020), and slightly higher compared with *M. yunnanensis* (120 genes) (Yang et al. 2019).

Phylogenetic relations according to cpDNA genomes between species were well resolved (bootstrap support >85) (Figure 3). Four *Myristica* species formed a monophyletic clade, showing a similar topology as previously reported (de Oliveira et al. 2020).

Discussion and conclusions

The cp genome structures of *M. argentea* and *M. fatua* have been analyzed and found to exhibit typical structural characteristics of angiosperm cp genomes, including one LSC and SSC region, as well as two IR regions. We identified a different length of *atpF*, *clpP*, and *ndhD*, genes in both species, with a difference from 1 to 3 bp. The cpDNA-based phylogenetic tree supports *M. argentea* as a sister species to *M. fragrans*, and the close taxonomic relationship between genera *Knema* and *Myristica*. This study will underpin the relationships between *Myristica* species and provide important details for evolutionary research and intra-specific diversity analysis.

Acknowledgements

We thank Gregory M. Plunkett, Sean Thackurdeen, and Michael J. Balick from the New York Botanical Garden (NYBG) for having provided samples of *M. fatua*.

Author contributions

JK and JD conducted the field work; JK, NS, and JD conceptualized and designed research; JK and MC performed lab experiment; JK and NS analyzed data; JK wrote original draft of the manuscript; all authors read and approved the final manuscript.

Ethical approval

This research was done with the authorization of Indonesian National Research and Innovation Agency (BRIN) with research permit to JD (144/SIP/FRP/E5/Ditk.KI/V/2018) and collecting permit to JK (B-3382/IPH.1/KS.02.04/IX/2019 and B-75/IV/KS.01.04/3/2022).

Disclosure statement

The authors report there are no competing interests to declare.

Funding

This work was supported by the Indonesian Endowment Fund for Education (LPDP) [0003719/AFR/D/BUDI-2018]; Agropolis Fondation [ID 1502-503 and 1502-504] through the «Investissements d'avenir» Program [Labex Agro: ANR-10-LABX-0001-01], under the frame of I-SITE MUSE [ANR-16-IDEX-0006]. Field work for obtaining *M. argentea* specimen was authorized by Indonesian National Research and Innovation Agency (BRIN) with research permit to JD [144/SIP/FRP/E5/Ditk.KI/V/2018] and collecting permit to JK [B-3382/IPH.1/KS.02.04/IX/2019 and B-75/IV/KS.01.04/3/2022]. The authors acknowledge the ISO 9001 certified IRD i-Trop HPC (member of

the South Green Platform) at IRD Montpellier for providing HPC resources that have contributed to the research results reported within this paper. URL: <https://bioinfo.ird.fr/>; <http://www.southgreen.fr>.

ORCID

Jakty Kusuma  <http://orcid.org/0000-0002-3842-9756>
 Marie Couderc  <http://orcid.org/0000-0001-7297-155X>
 Nora Scarcelli  <http://orcid.org/0000-0003-4382-9583>
 Jérôme Duminil  <http://orcid.org/0000-0002-2500-824X>

Data availability statement

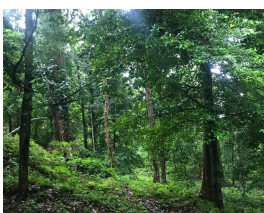
The genome sequence data on this study are openly available in GenBank of NCBI (<https://www.ncbi.nlm.nih.gov/>) indicated in Bioproject (PRJNA904151), BioSample (*M. argentea*: SAMN31831556; *M. fatua*: SAMN31831555), and SRA (*M. argentea*: SRR22428069; *M. fatua*: SRR22428070). The species are registered under the accession numbers OP866724 (*M. argentea*) and OP866725 (*M. fatua*).

References

- Cai C-N, Ma H, Ci X-Q, Conran JG, Li J. 2021. Comparative phylogenetic analyses of Chinese *Horsfieldia* (Myristicaceae) using complete chloroplast genome sequences. *J Syst Evol.* 59(3):504–514. doi: [10.1111/jse.12556](https://doi.org/10.1111/jse.12556).
- de Oliveira S, Duijm E, ter Steege H. 2020. Chloroplast genome of the nutmeg tree: *Myristica fragrans* Houtt. (Myristicaceae). bioRxiv.
- de Wilde WJJO. 2014. *Flora Malesiana*. 1st ed. Leiden, The Netherlands: Pensoft Publishers; p. 360–362.
- Govaerts R, Nic Lughadha E, Black N, Turner R, Paton A. 2021. The world checklist of vascular plants, a continuously updated resource for exploring global plant diversity. *Sci Data.* 8(1):215. doi: [10.1038/s41597-021-00997-6](https://doi.org/10.1038/s41597-021-00997-6).
- Jin J-J, Yu W-B, Yang J-B, Song Y, dePamphilis CW, Yi T-S, Li D-Z. 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biol.* 21(1):241. doi: [10.1186/s13059-020-02154-5](https://doi.org/10.1186/s13059-020-02154-5).
- Katoh K, Rozewicki J, Yamada KD. 2019. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform.* 20(4):1160–1166. doi: [10.1093/bib/bbx108](https://doi.org/10.1093/bib/bbx108).
- Lim T. 2012. *Myristica fatua*. In: *Edible medicinal and non-medicinal plants*. 1st ed. Netherlands: Springer; p. 572–574. doi: [10.1007/978-94-007-2534-8](https://doi.org/10.1007/978-94-007-2534-8).
- Mao C-L, Zhang F-L, Yang T, Li X-Q, Wu Y. 2019. The complete chloroplast genome sequence of *Myristica yunnanensis* (Myristicaceae). *Mitochondrial DNA Part B.* 4(1):1871–1872. doi: [10.1080/23802359.2019.1612711](https://doi.org/10.1080/23802359.2019.1612711).
- Mariac C, Luong V, Kapran I, Mamadou A, Sagnard F, Deu M, Chantereau J, Gerard B, Ndjeunga J, Bezançon G, et al. 2006. Diversity of wild and cultivated pearl millet accessions (*Pennisetum glaucum* [L.] R. Br.) in Niger assessed by microsatellite markers. *Theor Appl Genet.* 114(1): 49–58. doi: [10.1007/s00122-006-0409-9](https://doi.org/10.1007/s00122-006-0409-9).
- Mariac C, Scarcelli N, Pouzadou J, Barnaud A, Billot C, Faye A, Kougbéadjou A, Maillol V, Martin G, Sabot F, et al. 2014. Cost-effective enrichment hybridization capture of chloroplast genomes at deep multiplexing levels for population genetics and phylogeography studies. *Mol Ecol Resour.* 14(6):1103–1113. doi: [10.1111/1755-0998.12258](https://doi.org/10.1111/1755-0998.12258).
- Migliore J, Kaymak E, Mariac C, Couvreur TLP, Lissambou B-J, Piñeiro R, Hardy OJ. 2018. Pre-Pleistocene origin of phylogeographical breaks in African rain forest trees: new insights from *Greenwayodendron* (Annonaceae) phylogenomics. *J Biogeogr.* 46:212–223. doi: [10.1111/jbi.13476](https://doi.org/10.1111/jbi.13476).
- Scarcelli N, Mariac C, Couvreur TLP, Faye A, Richard D, Sabot F, Berthouly-Salazar C, Vigouroux Y. 2016. Intra-individual polymorphism in chloroplasts from NGS data: where does it come from and how to handle it? *Mol Ecol Resour.* 16(2):434–445. [Database] doi: [10.1111/1755-0998.12462](https://doi.org/10.1111/1755-0998.12462).
- Shi L, Chen H, Jiang M, Wang L, Wu X, Huang L, Liu C. 2019. CPGAVAS2, an integrated plastome sequence annotator and analyzer. *Nucleic Acids Res.* 47(W1):W65–W73. doi: [10.1093/nar/gkz345](https://doi.org/10.1093/nar/gkz345).

- Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*. 30(9):1312–1313. doi: [10.1093/bioinformatics/btu033](https://doi.org/10.1093/bioinformatics/btu033).
- Wang J, Li Y, Wang Q, Fan W. 2019. Characterization of the complete chloroplast genome of *Michelia maudiae* (Magnoliaceae). *Mitochondrial DNA B Resour*. 4(2):2146–2147. doi: [10.1080/23802359.2019.1623110](https://doi.org/10.1080/23802359.2019.1623110).
- Yang T, Mao C-L, Li X-Q, Zhang F-L, Zhao Q, Wu Y. 2019. The complete chloroplast genome sequence of *Knema furfuracea* (Myristicaceae). *Mitochondrial DNA B Resour*. 5(1):325–326. doi: [10.1080/23802359.2019.1703570](https://doi.org/10.1080/23802359.2019.1703570).
- Zhang F-L, Mao C-L, Li X-Q, Yang T, Wu Y. 2019. The complete chloroplast genome sequence of *Horsfieldia amygdalina* (Myristicaceae). *Mitochondrial DNA B Resour*. 4(2):3923–3924. doi: [10.1080/23802359.2019.1688122](https://doi.org/10.1080/23802359.2019.1688122).
- Zheng S, Poczai P, Hyvönen J, Tang J, Amiryousefi A. 2020. CHLOROPLOT: an online program for the versatile plotting of organelle genomes. *Front. Genet*. 11:576124. doi: [10.3389/fgene.2020.576124](https://doi.org/10.3389/fgene.2020.576124).

INTRODUCTION



- ❖ *Myristica fragrans* Houtt and its sister species (*M. fatua* Houtt, *M. argentea* Warb) commonly known as Indonesian nutmegs are important food tree species (FTS)
- ❖ Distributed in all Indonesian archipelago, in the wild or agroforestry systems
- ❖ Provide micronutrients
- ❖ Food, industrial, and medicinal purposes
- ❖ Worth more than \$50 million (FAO, 2017)

- THREATENED!**
- ❖ Loss of FTS habitats and biodiversity caused by high level of deforestation
 - ❖ Their genetic diversity is poorly documented which impedes development of efficient conservation and sustainable management strategies



Conceptual framework

- ❖ Conservation efforts are urgently needed to ensure the safeguarding and use of genetic resources of native populations of Indonesian nutmegs
- ❖ Determine the origin and genetic diversity distribution of cultivated Indonesian nutmeg population is important to develop conservation strategies

Fig 3. Seed variability of (a) *M. fatua*; (b) *M. fragrans*; and (c) *M. argentea* and their distribution in Indonesia and beyond (Fig 6, Fig 7, and Fig 8, respectively)

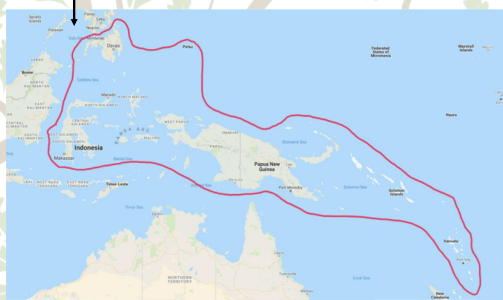


Fig 4. Natural distribution of *M. fatua*

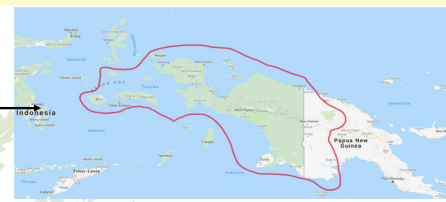


Fig 5. Natural distribution of *M. argentea*

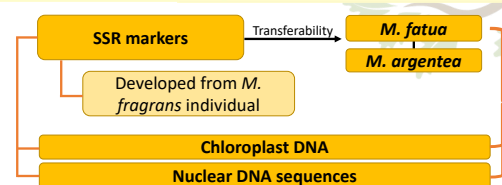


Fig 6. Natural distribution of *M. fragrans*. Inset: (a) natural origin of *M. fragrans*, and (b) Ambon Island

Objectives

- ❖ Characterize the distribution of the genetic diversity of Indonesian nutmeg species in Indonesia
- ❖ Orientate conservation, sustainable management and use of genetic resources by using Indonesian nutmeg as a model species

METHODS



Distribution of genetic diversity within and among Indonesian nutmeg species

INITIAL FINDINGS

Sampling, DNA extraction, SSR markers development



Fig 7. Sampling area of genetic materials in (a) Ambon Island; and (b) Banda Archipelago

- ❖ Total 334 samples, consisted of 252 *M. fragrans*, 55 *M. fatua*, and 12 *M. argentea*
- ❖ 48 SSR primer pairs tested (Myr_1 to Myr_48)

INITIAL FINDINGS

SSRs polymorphism test

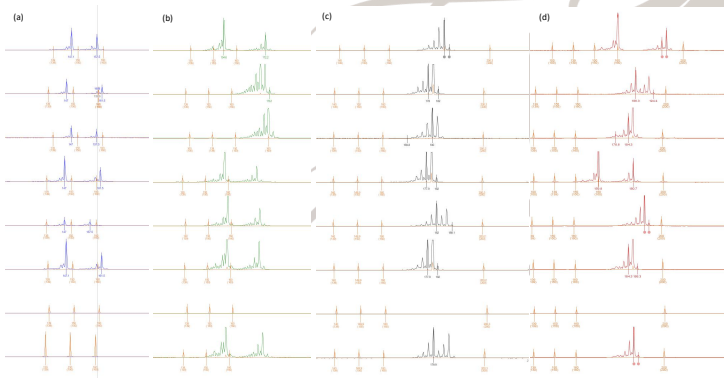


Fig 8. Peak results and polymorphism as shown by marker (a) Myr_13; (b) Myr_23; (c) Myr_26; and (d) Myr_28 on 7 different individuals.

Transferability test

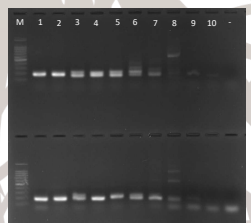


Fig 9. Bands result of transferability test by using Myr_23 (upper part) and Myr_25 (lower part) to other *Myristica* species. Line 1 to 10 are *M. argentea*, *M. fragrans* (positive control), *M. fatua*, *M. argentea*, *M. fatua*, *M. succedanea*, *M. lancifolia*, *M. elliptica*, *M. crassipes*, and *M. pachyphilla*, respectively.

48 primers tested and 16 are polymorphic in *M. fragrans* and present moderate to good transferability to other *Myristica* species.

CONCLUSION

- ❖ Screening genetic diversity of *M. fragrans*, *M. fatua*, and *M. argentea* is ongoing
- ❖ This will provide a comprehensive information to contribute to conservation and sustainable management strategies
- ❖ Genetic data will also allow us determining the geographic origin of *M. fragrans*
- ❖ We will extend the study to the whole *Myristica* family (phylogeny) to understand the evolutionary history of Malesiana flora

Reference

FAO. 2017. FAOSTAT database collections. FAO, Rome. Retrieved online (<http://faostat.fao.org>)

Acknowledgements

I would like to express my heartfelt gratitude to my parents, particularly my wife, who have instilled in me the value of gratitude. They taught me that after putting in effort, coupled with prayers, everything will be alright. Firstly, I am immensely thankful to the Higher Power above, who has always watched over and taken care of me. Delving deep into the study of evolution and bioinformatics has taught me that life is beautifully orchestrated by a script written by The One. In my opinion, He has created a perfect script with countless iterations and probabilities. Thus, when mortals act independently, the code adapts and executes its iterative condition accordingly. Eureka! That's how it works! Well, at least that's how it works in my imagination.

Furthermore, this thesis would not have been completed without the immense support and effort from my supervisor, Jérôme Duminil. Back in March 2017, I reached out to him via email, seeking his collaboration for the NUSANTARA program (which unfortunately did not materialize). To my surprise, he replied swiftly and willingly offered his assistance, going beyond the scope of the NUSANTARA program. In short, we embarked on a fruitful trip to Ambon, collecting nutmeg samples. It was during this time that I truly learned what it means to be a "true" researcher. When I arrived in Montpellier with limited knowledge of population genetics from an ecological perspective, he welcomed me at the airport and reassured me that I was not alone. In that moment, I believed that I could accomplish this task. Throughout my journey, he has taught me to think critically and nurture a perpetual curiosity. Our discussions have expanded my mind, and he has provided me with tremendous academic and personal support. I am truly grateful to him.

When Jérôme informed me that a woman would be my second supervisor, Nora Scarcelli, I was grateful for the gender balance. Nora's generosity was evident when she gave me two free tickets to visit the Musée Fabré during my first week in Montpellier—a remarkable experience. She introduced me to the fascinating world of bioinformatics. I must highlight her incredible dedication when she taught me bioinformatics through an email course for a full three weeks during the Covid-19

lockdown. Additionally, she possesses an exceptional ability to perceive atomic details in research and manuscripts that I may overlook. I am deeply appreciative of her contributions.

I would also like to express my gratitude to my thesis committee members—Gaël Kergoat, Thomas Couvreur, Pierre Gérard, and Deden D. Matra—for their valuable inputs during my *comité de suivi*. Their criticism, insights, and suggestions have been instrumental in shaping this thesis. I would like to express my gratitude to all those who have played a significant role in shaping this thesis. I extend my heartfelt thanks to my thesis committee members, Amandine Cornille, Myriam Heuertz, Nicolas Hubert, and Jacques David. Your invaluable input and insightful discussions have greatly contributed to expanding my understanding of evolutionary and ecological topics.

In the laboratory, I had a remarkable experience working alongside Marie Couderc, who possessed a Midas touch and magical pipetting skills. Despite our initial language limitations during my first months in France, we managed to communicate effectively using basic human communication, such as sign language, resulting in fruitful and amusing interactions. I am also thankful for the assistance provided by Leïla Zekroui and Cedric Mariac during my lab work, as they guided me in finding efficient methods.

I extend my gratitude to all of Jérôme's students: Aurore Rimlinger, who introduced me to the concept of "Aurore's pony horse" and provided immense help when I struggled with R; Yao Tze Leong, my partner in political gossip, with whom I had enjoyable conversations about our respective countries; Boniface Yogom, my strong Cameroonian friend—had you stayed longer, I would have become fluent in French; Franca Meguem, thank you for the bracelet—it gave me courage; Victoria Malaval, our time together in the lab was filled with fun, and you taught me many slang French words, which was entertaining; Gaël Bouka, your encouragement to pursue a postdoc as soon as possible has motivated me; and Pablo Dupiol, I hope you continue mastering R and bioinformatics—trust me, it will be a rewarding journey.

I am grateful to fellow students Mariano Kpatenon, who helped me solve bash scripting problems, and Sebastian Escobar, who assisted in developing effective

scripts. Aby Cisse, thank you for our engaging conversations on various topics. I also extend my thanks to fellow postdocs Andrew Helmstetter and Francis J. Nge for their expertise in bioinformatics and his assistance in developing and creating scripts for genomic data inference.

I would like to acknowledge the contributions of Phillippe Cubry from Team DYNADIV, who assisted me with demographic inference analysis. I am grateful to Miguel de Navascuès from INRAe and the University of Uppsala, Sweden, for his valuable suggestions and scripts during the analysis of demographic and population dynamics data. I also thank Pierre Mournet and Ronan Rivallan in CIRAD for their assistance during DNA extraction and genotyping. I thank Sylvain Santoni, Audrey Weber, and Morgane Ardisson for their help during genomic libraries activities. I would also like to thank my collaborators at the INOBU Foundation, John Watts and Amalia Paramitha, for their assistance in completing the Papuan nutmeg samples. Special thanks to the people at MNHN, including Myriam Gaudeuil and Thierry Deroin, who supported me in collecting herbarium samples. I had a great time with Thierry during my stay at MNHN. I would also like to express my appreciation to Roxali Bijmoer and Marnell Scherrenberg from Naturalis in Leiden, The Netherlands, and Gregory M. Plunkett, Sean Thackurdeen, and Michael J. Balick from the New York Botanical Garden. Lastly, a thank you to Jérôme Munzinger from the Herbar de l'Université de Montpellier (MPU).

I am indebted to Mr. Attaché Prof. Warsito and his wife, Madame Attaché Luh Anik Mayani, Mr. Consulaire Asianto Simamora and his wife, and Mr. Consulaire Arif Basalamah and his wife, as well as all the diplomats, including Alex Rudini, Théofilus Harefa, Gemma Ramadhan, and Agung Darmawan, for their assistance in resolving my passport and citizenship issues.

To my fellow Parisians—Eduardo Erlangga and Airis Sabriena, Aditia Gunawan and his wife, Tohir Mustofa, Acintya Ratna, Katarina Mellyna, and Gloria Truly—I extend my thanks for being part of this journey. To all *journée des doctorants indonesiennes* members, thank you very much for sharing. A special thanks to the Mahar Schützenberger committee, Hélène Schützengerber, and Dominique Perrin, for trusting me as one of your fellows.

Finally, to all my Indonesian friends—José Rizal and family, Joely Samadi, Agus Subianto and family, as well as Agus Pego—who shared laughter and joy during our free time. Emmanuel Paradis and Sintia Paradis, I cannot thank you enough for your unwavering support throughout my journey, and thanks for the book as well. Your belief in me and constant encouragement have been a source of strength and inspiration. To Youvita and Alexander Aviantara, I am deeply thankful for your unparalleled kindness. Your selflessness and generosity have touched my soul in ways words cannot express. Santi Silas and Charly Silas, I am forever grateful for the delectable Indonesian food you have graciously provided. Each meal was a delightful taste of home, and I cherish the moments we shared around the table. Tante Grace, our shared moments of craziness will forever be etched in my mind. Your vibrant spirit and zest for life brought immeasurable joy to our time together. Jane Couillerot and Jacques Couillerot, your acts of kindness will forever be etched in my memory. The warmth and compassion you have shown me are truly remarkable. Bunda Itjeu, your kindness and affection will forever be cherished. Your loving care and support have made a profound impact on my life, and I will always hold you dear in my heart. To Wika and Andon, thank you very much for being our ally in Nîmes, providing us a shelter whenever we came to your city, especially after watching Slipknot’s concert. To Mbak Arum, Mbak Nina, Mbak Rosmida, Eddy, and Teh Inong, thank you from the depths of my heart for taking care of Dina. Your love and support have provided her with comfort and strength, and I am forever grateful for your presence in her life. To Gabriel, Angelique, Githa, Nicole, Firza, Ruth, Angela, Nurul, Indra, Hadi, and Royan, the laughter we shared will forever echo. Each moment spent together was filled with joy, and I am grateful for the friendship and happiness you have brought into my life.

Lastly, I want to express my deepest gratitude to my parents, my siblings, and my wife. Despite the thousands of miles that separate us, your unwavering support has been the foundation of my journey. To my wife, Dina Tri Wahyuni, words will never be enough to express my gratitude for your constant presence. Through the ups and downs, you have been by my side, come rain or shine. From the very beginning of our journey, we made a promise to start together and finish together.

Thank you for your unwavering love and support, which have been my guiding light. I present this thesis to you, and to our Lundi.

Although this marks the end of my thesis, I believe that my journey is far from over. There are still many paths to explore, discoveries to be made, and knowledge to acquire. The knowledge and experiences I have gained during this thesis will continue to shape my future endeavors.

To be continued...

For Dina, and Lundi...

Abstract. This thesis investigates the evolutionary history of Indonesian nutmegs (*Myristica fragrans* and *M. argentea*), which are spice tree species native respectively to the Moluccas Archipelago and Papua Island. We explored the influence of the environment and human utilization on the evolutionary history of the two species. Based on a large sampling covering both native and introduced populations of the two species and using a set of complementary molecular markers (newly developed microsatellite markers, chloroplast genome sequences and nuclear genes), we investigated (i) the geographical origin of *M. fragrans*, its divergence in its native range and its introduction history; (ii) the impact of climate change, effect of topography and human utilization on the evolutionary history of *M. argentea* in Papua Island. Our findings revealed that *M. fragrans* originated from the South Moluccas, and that human activities most probably contributed to a reduction of population size of the species. Most of the introduced populations were sharing genetic similarities with the South Moluccas, suggesting that this later region was a primary source of propagation material along the introduction history of the species. Further, our results suggest that the evolutionary history of *M. argentea* in Papua Island was shaped by past climate change, mountain ranges and riverine networks. Introduced populations of *M. argentea* were mainly sourced from one native population in South Fakfak. This study provides insights into the evolutionary history of *M. fragrans* and *M. argentea*, informing conservation strategies and addressing environmental challenges.

Keywords: nutmeg, phylogeography, genetic diversity, Moluccas, Papua, climate change

Résumé. Cette thèse étudie l'histoire évolutive des noix de muscade indonésiennes (*Myristica fragrans* et *M. argentea*), deux espèces d'arbres à épices originaires respectivement de l'archipel des Moluques et de Papouasie. Sur la base d'un large échantillonnage couvrant à la fois les populations indigènes et introduites des deux espèces, et en utilisant un ensemble de marqueurs moléculaires complémentaires (microsatellites développés *de novo*, séquences du génome chloroplastique et gènes nucléaires), nous avons étudié (i) l'origine géographique de *M. fragrans*, sa divergence dans son aire de répartition naturelle et l'histoire de son introduction ; (ii) l'impact du changement climatique, l'effet de la topographie et de l'utilisation humaine sur l'histoire évolutive de *M. argentea*. Nos résultats ont révélé que *M. fragrans* est originaire des Moluques du Sud et que les activités humaines ont très probablement contribué à une réduction de la taille de la population de l'espèce. La plupart des populations introduites partageaient des similitudes génétiques avec les Moluques du Sud, suggérant que cette région a été une source principale de matériel de propagation de l'espèce. En outre, nos résultats suggèrent que l'histoire évolutive de *M. argentea* dans l'île de Papouasie a été façonnée par les changements climatiques passés, les chaînes de montagnes et les réseaux fluviaux. Cette étude a permis de mieux comprendre l'histoire évolutive de *M. fragrans* et de *M. argentea* et d'éclairer les stratégies de conservation.

Mots clés : muscade, phylogéographie, diversité génétique, Moluques, Papouasie, changements climatiques.