universität wien

DISSERTATION

Titel der Dissertation

„Evolution und Biodiversität von neukaledonischen *Diospyros*“

verfasst von

Mag. rer. nat. Barbara Turner

angestrebter akademischer Grad

Doctor of Philosophy (PhD)

Wien, 2014

Studienkennzahl lt. Studienblatt:	A 094 437
Dissertationsgebiet lt. Studienblatt:	Biologie
Betreuerin:	ao. Univ.-Prof. i.R. Dipl.-Biol. Mag. Dr. Mary Rosabelle Samuel

ACKNOWLEDGEMENTS

Although a thesis is the finishing-work of one person, it is the work of many persons. At this point I want to thank everybody who supported me throughout my academic career and especially during my work on this thesis.

Warm thanks go to Prof. Rosabelle Samuel, the supervisor of my work. She takes her duties as supervisor very serious and supports her students wherever she can. Beside the supervision of my work she gave me the opportunity to participate at conferences and get first experience in teaching.

Thanks to all co-authors for their excellent professional support. Special thanks go to Dr. Jérôme Munzinger who imparted me parts of his comprehensive knowledge about *Diospyros* and New Caledonia. Thanks to Dr. Ovidiu Paun who taught me a lot about AFLP, RAD and analysis of population-genetics data, as well as to Prof. Mark. W. Chase for his help with all questions about plant systematics and data evaluation and of course for his help with formulating English texts.

Thanks to Verena Klenja who did most of the lab work for me. Thanks also to Ing. Elfriede Grasserbauer and Mag. Gudrun Kohl for their help and advices with lab work. Thanks to the team of IRD AMAP Noumea for their help in New Caledonia. Apart from already mentioned persons I want to cordially thank Vanessa Hequet for her support.

Thanks to Dr. Fatemeh Maghuly as well as the rest of the team of IAM of the University of Natural Resources Vienna for their help with Southern Blotting trials.

Very cordially thanks go to Dr. Michael H.J. Barfuss for his professional, mental and amicable support with my work. If one counts the entire thesis he help with, he deserves at least ten PhD titles. Colleagues and students of the Department of Botany and Biodiversity Research of the University Vienna supported me both professionally and morally. For this I want to thank among others Prof. Hanna Schneeweiss, Prof. Gerald Schneeweiss, Prof. Jürg Schönenberger, Dr. Hermann Voglmayr, Dr. Fancisco Balao Robles, Dr. Khatere Emadzade, Dr. Anne-Caroline Cosendai, Dr. Stefan Safer and Dr. Anton Russell.

Beside all this professional support I want to thank my family and friends who were always there for me when I needed them and had appreciation for my short time resources. Special thanks go to my mum and my grand-parents who always motivated me and taught me how important it is to fight for your dreams and wishes.

Thanks to my husband Wolfgang who is always there for me, who stands to me in every situation and has brought up much appreciation for my time at university.

This project was funded by the Austrian Science Fund (FWF; project P22159, grant given to ao. Univ.-Prof. Dr. Mary Rosabelle Samuel).

DANKSAGUNG

Obwohl eine Dissertation die Abschluss-Arbeit einer Person ist, ist die ganze Arbeit doch jene vieler Personen. An dieser Stelle möchte ich mich bei all jenen bedanken, die mich während meiner akademischen Laufbahn und speziell bei der Arbeit an dieser Dissertation unterstützt haben.

Ein großer Dank geht an Frau Prof. Rosabelle Samuel, die Betreuerin meiner Arbeit. Sie nimmt ihre Aufgabe sehr ernst und unterstützt ihre Studenten wo sie kann. Neben der Betreuung meiner Arbeit, hat sie mir auch die Möglichkeit gegeben an Konferenzen teilzunehmen und erste Erfahrungen im Unterrichten zu sammeln.

Danke an alle Co-Autoren, für die exzellente fachliche Unterstützung. Speziell bedanken möchte ich mich bei Dr. Jérôme Munzinger, der mir einen Bruchteils seines umfassenden Wissens über *Diospyros* und Neu Kaledonien vermitteln konnte. Danke an Dr. Ovidiu Paun der mir vieles über AFLP, RAD und die Analyse von Populations-genetischen Daten beigebracht hat, sowie an Prof. Mark W. Chase für seine Hilfe bei sämtlichen Fragen zu Pflanzensystematik und Datenauswertung und seiner Hilfe beim Formulieren der englischen Texte.

Danke an Verena Klenja, die den Großteil der Laborarbeit für mich gemacht hat. Danke auch an Ing. Elfriede Grasserbauer und Mag. Gudrun Kohl für die ihre Hilfe und Ratschläge bei der Laborarbeit. Danke an das Team des IRD AMAP Noumea für ihre Hilfe bei der Arbeit in Neu Kaledonien. Außer bei bereits erwähnten Personen möchte ich mich ganz herzlich bei Vanessa Hequet für ihre Unterstützung bedanken.

Bedanken möchte ich mich bei Dr. Fatemeh Maghuly sowie dem Rest des Teams des IAM der Universität für Bodenkultur Wien für ihre Hilfe bei den Southern-Blotting Versuchen.

Ganz herzlich bedanken möchte ich mich bei Dr. Michael H.J. Barfuss für die fachliche, mentale und freundschaftliche Unterstützung meiner Arbeit. Wenn man alle Dissertation, bei denen er mitgeholfen hat zusammenzählt, verdient er mindestens zehn Doktortitel. Viele Mitarbeiter und Studenten des Departments für Botanik und Biodiversitätsforschung der Universität Wien haben mich sowohl wissenschaftlich als auch moralisch unterstützt und dafür möchte ich mich unter anderen bei Prof. Hanna Schneeweiss, Prof. Gerald Schneeweiss, Prof. Jürg Schönenberger, Dr. Hermann Voglmayr, Dr. Fancisco Balao Robles, Dr. Khatere Emadzade, Dr. Anne-Caroline Cosendai, Dr. Stefan Safer und Dr. Anton Russell bedanken.

Neben all der fachlichen Unterstützung möchte ich mich auch bei meiner Familie und meinen Freunden bedanken, die immer dann für mich hier waren, wenn ich sie gebraucht habe und viel Verständnis für meine knappen Zeitressourcen gezeigt haben. Speziell bedanken möchte ich mich bei meiner Mama und meinen Großeltern, die mich immer wieder aufs Neue motiviert haben und mir beigebracht haben, wie wichtig es ist, für seine Träume und Wünsche zu Kämpfen.

Danke an meinen Mann Wolfgang, der immer für mich da ist, in jeder Situation zu mir steht und sehr viel Geduld und Verständnis für meine Zeit auf der Uni aufgebracht hat.

Dieses Projekt wurde vom Österreichischen Fonds zur Förderung der wissenschaftlichen Forschung (FWF; Projekt P22159, Finanzierung vergeben an ao. Univ.-Prof. Dr. Mary Rosabelle Samuel) finanziert.

CONTENTS

ABSTRACT

Diospyros forms a large genus of woody flowering plants of the family Ebenaceae. New Caledonia is an archipelago in the south-western Pacific and harbours a great range of diverse habitats and a characteristic flora. In total there are 31 *Diospyros* species found in New Caledonia, of which 30 are endemic. Previous phylogenetic studies of the genus *Diospyros* based on sequences of plastid markers, showed the New Caledonian species to form three groups. Two of these groups contain only few species (two and, respectively, five), and the majority of species (24) forms the third group. The New Caledonian *Diospyros* species are morphologically diverse and occupy ecologically different habitats. This thesis focuses on the species rich third group, which is addressed as group 3 in this abstract.

In our first study, including sequences of two nuclear and four plastid markers, the species of group 3 are shown to be closely related. *Diospyros vieillardii* is clearly shown to be sister to the rest of this group. Apart from this, individuals of only four species formed unique groups. The morphological species concept, on which the species have been described, is supported by our AFLP results. However, with AFLP data we were not able to elucidate the phylogenetic relationships between these closely related species. Analysis of molecular variance (AMOVA) of the AFLP data set showed that most of the genetic variation occurs within the species rather than among them. Species delimitations inferred from next generation sequencing technique RAD (Restriction site associated DNA) are comparable to those obtained from AFLP data. Phylogenetic trees based on thousands of RAD-derived SNPs are much better resolved than those based on Sanger sequencing of nuclear and plastid markers. Most of the 21 included species formed monophyletic groups in AFLP and RAD analyses. The observed phylogenetic relationships do not follow an ecological structure, pointing to a role of environmental heterogeneity of New Caledonia in shaping speciation events in this group. Functional annotations of genomic regions consistently exhibiting high differentiation between pairs of sister species occurring on different substrates (e.g. *D. flavocarpa* – *D. umbrosa*, *D. labillardierei* – *D. trisulca*) pointed to genes involved in binding and transporting compounds to/through the cell membrane. Species from group 3 revealed nearly 3-fold larger genome sizes compared to *Diospyros* species from other groups. Chromosome counts showed no indication of polyploidy in this group. The increase in genome size in these species led us to investigate the repeated elements of these genomes. Whole genome sequencing using next generation sequencing techniques showed that the larger genomes generally contain more copies of repeated elements such as LTR/*gypsy* elements, without a significant enrichment for a particular element type. Up to now no species specific repeat elements have been identified. Beside the repeated elements we were able to obtain as a by-product whole plastid sequences from the low-coverage whole genome sequencing. The obtained plastomes were compared to

the plastid sequence of *Camellia sinensis*. The plastid genomes of *Diospyros* and *Camellia* are highly similar in size, structural organization and gene content. Dating analyses based on DNA sequence and RAD data showed that the crown group 3 is around seven million years old and the group with low statistical support in the RAD based analysis to be around four million years. *Diospyros* are woody plants with a generation time of several years, thus we can estimate, that not more than 500,000 generations passed since the most recent common ancestor of the latter *Diospyros* group. The low number of generations after the original long distance dispersal event, together with the rapid radiation across different habitats can explain the presence of the low genetic divergence in this group.

KURZFASSUNG

Diospyros ist eine große Gattung holziger Blütenpflanzen und zählt zur Familie Ebenaceae. Neu Kaledonien ist eine Inselgruppe in süd-westlichen Pazifik und beherbergt eine große Anzahl an unterschiedlichen Habitaten und eine charakteristische Flora. In Summe findet man 31 *Diospyros*-Arten in Neu Kaledonien, von denen 30 endemisch sind. Vorhergehende Studien der Gattung *Diospyros*, basierend auf Sequenzen von Plastidenmarkern, zeigten, dass die Neu Kaledonischen Arten drei Gruppen bilden. Zwei dieser Gruppen beinhalten nur wenige Arten (zwei bzw. fünf) und die Mehrheit der Arten (24) bildet die dritte Gruppe. Die Neu Kaledonischen *Diospyros*-Arten sind morphologisch unterschiedlich und bewohnen ökologisch verschiedene Habitate. Diese Arbeit konzentriert sich auf die artenreiche dritte Gruppe, die in dieser Zusammenfassung als „Gruppe 3“ bezeichnet wird.

Im unseren ersten Untersuchungen von Sequenzen von zwei Kern- und vier Plastidenmarkern zeigten sich die Arten der Gruppe 3 als sehr nahe miteinander verwandt. *Diospyros vieillardii* ist klar als Schwester zum Rest der Gruppe herausgekommen. Davon abgesehen formten nur bei vier Arten die jeweiligen Individuen einheitliche Gruppen. Das morphologische Artkonzept, nach welchem diese Arten beschrieben wurden, wird von AFLP-Daten unterstützt. Nichts desto trotz haben uns die AFLP-Daten nicht geholfen die Verwandtschaftsverhältnisse zwischen den nahe verwandten Arten aufzuklären. Analysen der molekularen Varianz (AMOVA) der AFLP-Daten, zeigten, dass der Großteil der genetischen Variation innerhalb der Arten vorkommt und nicht zwischen diesen. Aus RAD-Daten (Restriktionsstellen Assoziierte DNA; Sequenziermethode nächster Generation) abgeleitete Artabgrenzungen sind ähnlich derer von AFLP-Daten. Stammbäume basierend auf SNPs aus RAD-Daten sind wesentlich besser aufgelöst und besser unterstützt, als jene basierend auf den Sanger-Sequenzdaten der Kern- und Plastidenmarkern. Die meisten der 21 inkludierten Arten formen sowohl in den AFLP- als auch in den RAD-Analysen, monophyletische Gruppen. Die beobachteten, phylogenetischen Verwandtschaftsverhältnisse folgen keiner ökologischen Struktur, was darauf hindeutet, dass die ökologische Vielfalt Neu Kaledoniens eine Rolle bei Artbildungsprozessen hatte. Funktionelle Annotierungen von genomische Regionen mit konstant hohen Unterschieden zwischen Schwesterarten, welche auf unterschiedlichen Substraten vorkommen (z.B. *D. flavocarpa* – *D. umbrosa*, *D. labillardierei* – *D. trisulca*), deuten auf Gene, die in Bindung und Transport von Substanzen an/durch die Zellmembran, hin. Genomgrößenmessungen zeigten um fast 3-fach höhere Genomgrößen bei Arten der Gruppe 3 im Vergleich zu anderen *Diospyros*-Arten. Chromosomenzählungen haben keine Anzeichen von Ploidyploidie in dieser Gruppe geliefert. Dieser Anstieg der Genomgröße veranlasste uns die wiederholten Elemente dieser Genome zu untersuchen. Gesamtgenomsequenzierungen mittels einer Sequenziermethode nächster Generation zeigten, dass die größeren Genome

generell mehr wiederholte Elemente wie LTR/*gypsy* beinhalten, aber ohne signifikanter Anreicherung eines bestimmten Element-Typs. Bis jetzt wurden noch keine artspezifischen wiederholten Elemente identifiziert. Abgesehen von den wiederholten Elementen, konnten wir als Nebenprodukt auch ganze Plastidensequenzen aus den Daten der Gesamtgenomsequenzierung mit niedriger Abdeckung, sammeln. Diese Sequenzen konnten zu einen ganzen Plastidengenom zusammengesetzt werden. Die Plastome von *Diospyros* wurden mit der Plastidensequenz von *Camellia sinensis* verglichen. Die Plastidengenome von *Diospyros* und *Camellia* sind in ihrer Größe, strukturellen Organisation und Gen-Gehalt, sehr ähnlich. Datierungsanalysen basierend auf DNA-Sequenz- und RAD-Daten, zeigten, dass die Krongruppe 3 ca. sieben Millionen Jahre alt ist und die Gruppe mit geringer statistischer Unterstützung in den RAD-Analysen ca. vier Millionen Jahre alt ist. *Diospyros* sind holzige Gewächse mit einer Generationszeit von mehreren Jahren. Daher schätzen wir, dass es seit dem letzten gemeinsamen Vorfahren der zuletzt genannten *Diospyros*-Gruppe ca. 500.000 Generationen gegeben hat. Diese geringe Anzahl an Generationen, nach dem originalen Verbreitungsereignis über lange Distanzen, gemeinsam mit einer schnellen Radiation über die unterschiedlichen Habitate, kann das Vorhandensein der niedrigen genetischen Diversität, erklären.

INTRODUCTION

Genus *Diospyros* and family Ebenaceae

Diospyros L. is the most species rich genus of the pantropical family Ebenaceae, which are included in Ericales (APG III 2009). Within Ericales Ebenaceae are placed in a clade with Sapotaceae, Maesaceae, Theophrastaceae, Primulaceae and Myrsinaceae (Schönenberger *et al*. 2005). The family Ebenaceace can be divided into two subfamilies (Duangjai *et al*. 2006), Lissocarpoideae and Ebenoideae. Subfamily Lissocarpoideae is monogeneric, the genus *Lissocarpa* comprises 8 species which are found in the tropics of north western South America (Wallnöfer 2004). Subfamily Ebenoideae include *Diospyros* (>500 species, pantropical), *Euclea* (18 species in Africa) and *Royena* (17 species in Africa), with the latter two genera forming a clade which is sister to *Diospyros*. This classification of Ebenaceae into two subfamilies and four genera has been also supported by palynological data (Geeraets *et al*. 2009).

Diospyros is a large genus comprising more than 500 species, with the majority (~300 species) being distributed in Asia and the Pacific region. Individuals of this genus are shrubs or trees which occur in most tropical and subtropical habitats where they are often important and characteristic elements of the vegetation. The leaves are entire, arranged alternate and having flat glands on the lower surface. The flowers are actinomorphic, 3-8 merous, the calyx is adenated to the contorted corolla and they are often grouped together in axil born cymes. The superior ovary develops into a berry to which the calyx stays attached. There are monoecious (both hermaphroditic and single-sex) and dioecious species. *Diospyros* species are diploid with 2n = 2x = 30 chromosomes (e.g. Tamura *et al*. 1998; White 1992). Several species of *Diospyros* are of economic value as they have edible fruits (persimmons; e.g. *D. kaki*, *D. lotus* and *D. virginiana*) or precious timbers (ebony wood; e.g. *D. ebenum*). Species with edible fruits are often polyploid; e.g. commercial strands of *D. kaki* are hexaploid (2n = 90).

Current status of knowledge about *Diospyros* in New Caledonia

Recent molecular studies on the whole genus *Diospyros* found 11 mostly well-resolved clades within this genus (Duangjai *et al*. 2009; Fig. 2). In New Caledonia, there are 31 *Diospyros* species described, of which all but one are endemic (White 1992, 1993) and they are found in three clades (Fig. 2). The first group of New Caledonian *Diospyros* species (5 species; *D. balansae*, *D. brassica*, *D. macrocarpa*, *D. margaretae*, *D. oubatchensis*) forms a clade with Australian species (Fig. 2, clade II). The second group, consisting of species from Asia, America, Pacific Islands and New Caledonia, includes two widespread New Caledonian species (*D. olen* being endemic and *D. fasciculosa* is found throughout the southern Pacific, Fig. 2, Clade IX). These two species are not sister species, accounting for two colonisation events of


A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y

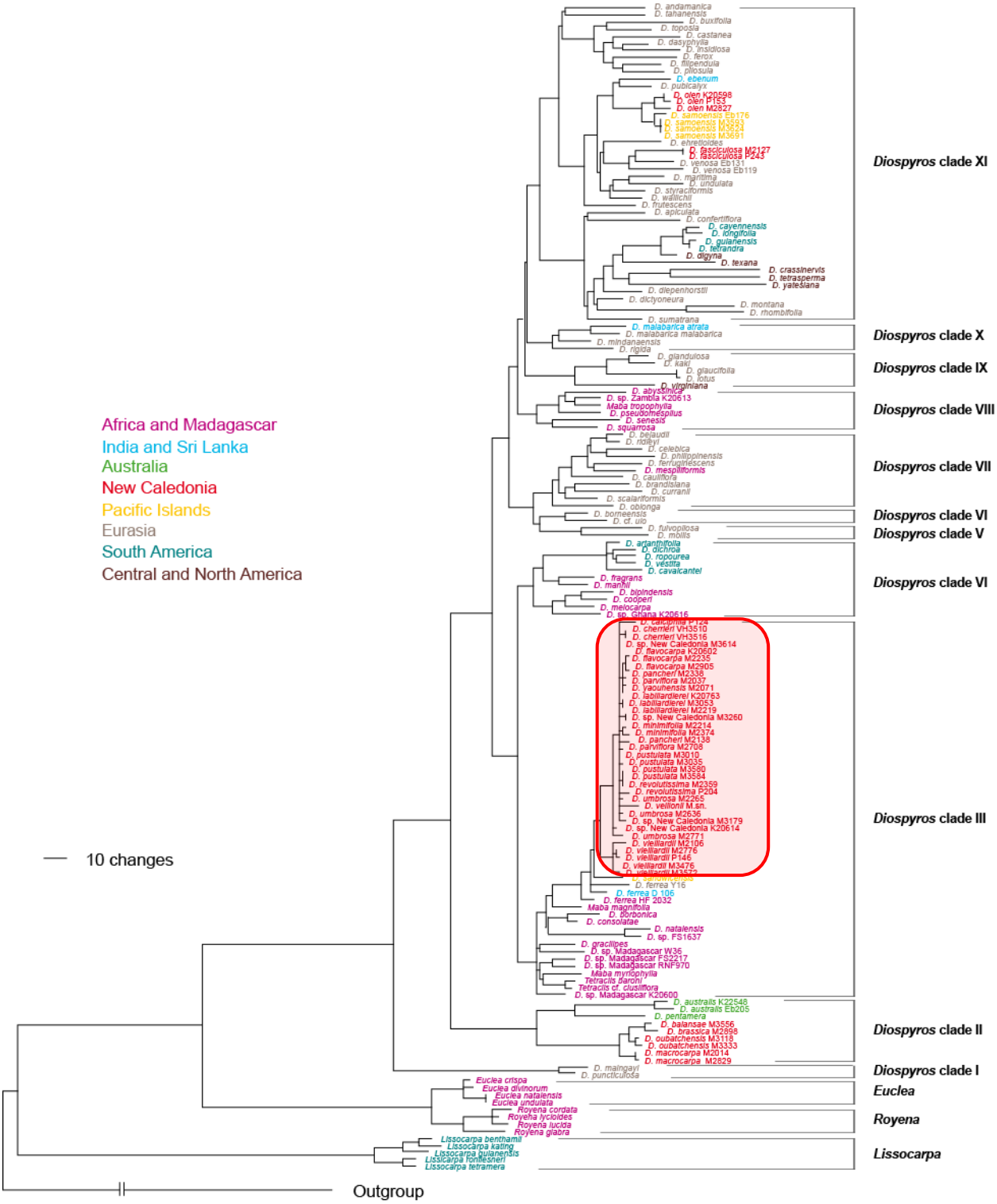


Figure 2: Phylogenetic tree of genus *Diospyros* and related genera based on plastid markers. Red square indicates study group of this work. (from Duangjai *et al*., 2009).

Left page: Plate 1: *Diospyros* species from New Caledonia.
A: *D. balansae*, B: *D. brassica*, C: *D. calciphila*, D: *D. cherrieri*, E: *D. erudita*, F: *D. fasciculosa*, G: *D. flavocarpa*, H: *D. glans*, I: *D. impolita*, J: *D. labillardierei*, K: *D. macrocarpa*, L: *D. margaretae*, M: *D. minimifolia*, N: *D. oubatchensis*, O: *D. pancheri*, P: *D. parviflora*, Q: *D. perplexa*, R: *D. pustulata*, S: *D. olen*, T: *D. revolutissima*, U: *D. tireliae*, V: *D. umbrosa*, W: *D. veillonii*, X: *D. veillardii*, Y: *D. yahouensis*.
Picture credits are given at the end of References.

New Caledonia. The third group includes species distributed from Madagascar over Indian Ocean Islands, South-East Asia, South Pacific Islands (including New Caledonia, 24 species) to Hawai'i. The New Caledonian taxa belonging to this clade (clade III) seem to be closely related. One of the closest relatives to this group is *D. vera*, a widespread species found in Africa and in the whole Indian Ocean and western South-Pacific region. Phylogenetic analyses analysed using 8000 base pairs from the plastid genome showed no resolution among the New Caledonian species of this third group (Duangjai *et al*. 2009). New Caledonia has been colonised by *Diospyros* at least four times. Similar, multiple colonisation events are also found among other organisms in New Caledonia (e.g. Murienne *et al*. 2005).

Diospyros is observed in all kinds of vegetation in New Caledonia except mangrove; the species range from sea level up to ca. 1250 m (New Caledonia's highest point, Mount Panié is 1628 m). There are several micro-endemics restricted to just a small area (White, 1992). Most of the New Caledonian *Diospyros* species are morphologically clearly defined and appear related to edaphic factors, occurring on just one kind of substrate. However, several *Diospyros* species occur in sympatry in many localities.

New Caledonia

New Caledonia is an island group located in the south-western Pacific about 1,300 km east of Australia (Fig. 3), ranging from around 19° to 23° south with an land area of ca. 19,000 km^2. It consists of the main island Grande Terre (ca. 16,000 km^2), Îles Belep (in the north), Île des Pins (in the south), Loyalty Islands (in the east) and several other smaller islands. The New Caledonian climate is tropical to subtropical. The main island is split by a mountain range into a humid eastern portion (2000-4000 mm precipitation per year) and a dry western part (1000 mm precipitation per year) with winds and rain coming from the south east. The continental part of New Caledonia (mainly Grande Terre) separated from Gondwana during late Cretaceous (ca. 80 million years ago, mya; McLoughlin 2001). During the Palaeocene to late Eocene, this continental sliver was submerged for at least 20 million years (myr), and a thick layer of oceanic mantle accumulated (Pelletier 2006). After Grande Terre re-emerged in the late Eocene (37 mya), this heavy-metal rich oceanic material covered most of the land area. Today, around 1/3 of the main island is still covered with ultramafic substrates. Because Grande Terre was entirely submerged, it is highly unlikely that lineages that were already present in this region before the split from Gondwana could have survived locally. Current hypotheses suggest that biota present today are derived from elements/ancestors that reached New Caledonia via long distance dispersal (e.g. Morat *et al*. 2012; Pillon 2012; Grandcolas *et al*. 2008) mainly from Australia, New Guinea and Malaysia. Hypotheses of other islands between Australia and New Caledonia having served as stepping stones or refuges for Gondwanan taxa now endemic (e.g. *Amborella*) have been proposed by a few authors (Ladiges and Cantrill 2007), but there is no consensus of

how many, when they existed or how large they might have been. New Caledonia is one of the 34 biodiversity hotspots (Mittermeier *et al*. 2004; Myers *et al*. 2000), and nearly 75% of the native flora is endemic (Morat *et al*. 2012), which is the fourth highest found on islands (Lowry 1998). Among these endemic taxa there are 98 genera and three families, Amborellaceae, Oncothecaceae and Phellinaceae (Morat *et al*. 2012). One of the reasons hypothesised for the high level of endemism found in New Caledonia is the ultramafic substrate (Pillon *et al*. 2010), which acted as filter for species which were already pre-adapted to the ultramafic soils.

Speciation and evolution on oceanic islands

Oceanic islands are regularly relatively small land masses, geographically isolated, with a known geological age, and harbour special biota with high levels of endemism. These characteristics make oceanic island a natural laboratory for the study of evolution, which has fascinated generations of biologists (e.g. Darwin 1842, 1859; Wallace 1881; MacArthur and Wilson 1967; Carlquist 1974; Grant 1996; Whittaker 1998; Bramwell and Caujapé-Castells 2011). In most cases only few individuals (represented by their diaspores) from the original population reach the new habitat (island). Those few individuals, forming the founder population in the new habitat, represent only a fraction of the genetic diversity of the original population. This fact of diversity reduction is termed bottleneck-effect. If those dispersal events happened more recent or if the respective biota have long generation times we have to consider low genetic diversity within such groups. The environmental conditions in the new habitat are often different from those in the original habitat. For successful colonisation of a new habitat, lineages have to adapt to the altered conditions and these adaptations have to happen in short time. Adaptive radiation (speciation through rapid adaptation to different ecological conditions), is an often observed phenomenon in biota on oceanic islands. During adaptive radiation the initial founder population divides into several lineages adapted to different ecologic realities. These diverging lineages are morphologically and/or physiologically distinct, accumulating some genetic differences, but the more conspicuous pattern is partitioning of the gene pool into restricted genetic lineages (Schluter 1996). Several cases of rapid radiation have been documented on oceanic islands (e.g. New Caledonia: *Araucaria*, Gaudeul *et al*. 2012; New Zealand: *Ourisa*, Meudt *et al*. 2009; Hawai'i: silverswords, Baldwin and Sanderson 1998; lobeliads, Givnish *et al*. 2009; *Bidens*, Knope *et al*. 2012; Canary Islands: *Aeonium*, Mort *et al*. 2002; *Echium*, Böhle *et al*. 1996). The whole process of adaptive radiation is also termed cladogenic speciation. Apart from the above mentioned cladogenic speciation, there is another speciation process common on islands, termed anagenetic speciation. During anagenetic speciation, an initial founder lineage simply transforms genetically and morphologically through time without further specific differentiation forming just one species (Stuessy *et al*. 1990; Stuessy *et al*. 2006; Stuessy 2007; Whittaker *et al*. 2008).

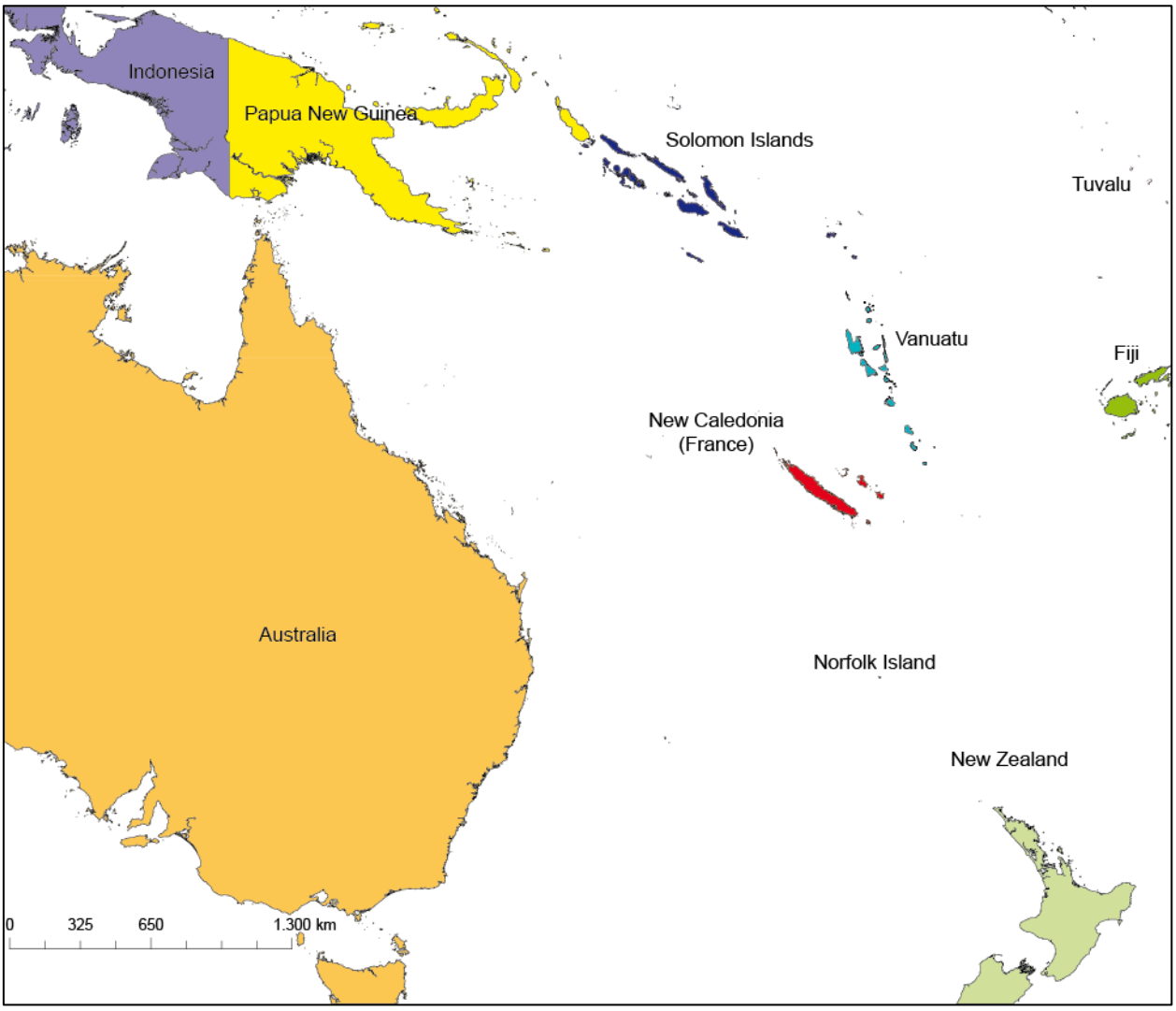


Figure 3: Map of the south Pacific region. New Caledonia is marked in red.

AIMS

Here in this project we aimed to answer following questions concerning the New Caledonian *Diospyros* species:

1. Relationships between the species.
2. How old are they and when did their ancestors arrive in New Caledonia?
3. Is there hybrid speciation?
4. Is polyploidy a cause for speciation?
5. Are the morphologically defined species genetically separated from each other?
6. What causes the differences in genome size?
7. Are there any repeated elements unique for the different species?

Answers to those questions are presented in four chapters in the present thesis. In chapter one, questions about phylogenetic relationships (1), age (2), hybrid speciation (3), and ploidy level (4) are dealt with. To answer questions 1-3 we used fast evolving plastid and/or low copy nuclear markers. To elucidate ploidy level/chromosome numbers and genome size we used Feulgen staining of chromosome preparations from root tips and flow cytometry. Questions about species boundaries (5) were assessed using AFLP and RADseq. Both approaches are dealt with in separate chapters (chapters two and three). To find repeated elements causing the genome size differences (6 and 7) we conducted whole genome sequencing with low coverage using the Ilumina technology. Details of this this work is given in chapter four.

References

APG III. 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. Botanical Journal of the Linnean Society 161: 105-121.

Baldwin BG, Sanderson MJ. 1998. Age and rate of diversification of the Hawaiian silversword alliance (Compositae). Proceedings of the National Academy of Sciences USA 95: 9402-9406.

Böhle U-R, Hilger HH, Martin WF. 1996. Island colonization and evolution of the insular woody habit in *Echium* L. (Boraginaceae). Proceedings of the National Academy of Sciences, USA 93: 11740-11745.

Bramwell D, Caujapé-Castells J. 2011. The biology of island floras. Cambridge University Press, Cambridge, UK.

Carlquist S. 1974. Island biology. Columbia University Press, New York, US.

Darwin C. 1842. The structure and distribution of coral reefs. Smith, Elder and Co., London, UK.

Darwin C. 1859. On the origin of species by means of natural selection. John Murrey, London, UK.

Duangjai S, Samuel R, Munzinger J, Forest F, Wallnöfer B, Barfuss MHJ, Fischer G, Chase MW. 2009. A multi-locus plastid phylogenetic analysis of the pantropical genus *Diospyros* (Ebenaceae), with an emphasis on the radiation and biogeographic origins of the New Caledonian endemic species. Molecular Phylogenetics and Evolution 52: 602-620.

Duangjai S, Wallnöfer B, Samuel R, Munzinger J, Chase MW. 2006. Generic delimitation and relationships in Ebenaceae *sensu lato*: evidence from six plastid DNA regions. American Journal of Botany 93: 1808-1827.

Gaudeul M, Rouhan G, Gardner MF, Hollingsworth PM. 2012. AFLP markers provide insights into the evolutionary relationships and diversification of New Caledonian *Araucaria* species (Araucariaceae). American Journal of Botany 99: 68-81.

Geeraerts A, Raeymaekers JAM, Vinckier S, Pletsers A, Smets E, Huysmans S. 2009. Systematic palynology in Ebenaceae with focus on Ebenoideae: Morphological diversity and character evolution. Review of Palaeobotany and Palynology 153: 336-353.

Givnish TJ, Millam KC, Mast AR, Paterson TB, Theim TJ, Hipp AL, Henss JM, Smith JF, Wood KR, Sytsma KJ. 2009. Origin, adaptive radiation and diversification of the Hawaiian lobeliads (Asterales: Campanulaceae). Proceedings of the Royal Society B – Biological Sciences 276: 407-146.

Grandcolas P, Murienne J, Robillard T, Desutter-Grandcolas L, Jourdan H, Guilbert E, Deharveng L. 2008. New Caledonia: a very old Darwinian island? Philosophical transactions of the Royal Society of London. Series B – Biological sciences 363: 3309-3317.

Grant PR. 1996. Evolution on islands. Oxford University Press, Oxford, UK.

Knope ML, Morden CW, Funk VA, Fukami T. 2012. Area and the rapid radiation of Hawaiian *Bidens* (Asteraceae). Journal of Biogeography 39: 1206-1216.

Ladiges PY, Cantrill D. 2007. New Caledonia–Australian connections: biogeographic patterns and geology. Australian Systematic Botany 20: 383-389.

Lowry II PP. 1998. Diversity, endemism and extinction in the flora of New Caledonia: a review. In: Peng CF, Lowry II PP. (Eds.), Rare, threatened, and endangered floras of Asia and the Pacific rim. Institute of Botany, Taipei, Taiwan, pp. 181-206.

MacArthur RH, Wilson EO. 1967. The theory of island biogeography. Princeton University Press, Princeton, US.

McLoughlin S. 2001. The breakup history of Gondwana and its impact on pre-Cenozoic floristic provincialism. Australian Journal of Botany 49: 271-300.

Meudt HM, Lockhart PJ, Bryant D. 2009. Species delimitation and phylogeny of a New Zealand pant species radiation. BMC Evolutionary Biology 9: 11.

Mittermeier RA, Gil PR, Hoffmann M, Pilgrim J, Brooks T, Mittermeier CG, Lamoreux J, da Fonseca GAB. 2004. Hotspots revisited: Earth's biologically richest and most endangered terrestrial ecoregions. CEMEX, Mexico City.

Morat P, Jaffré T, Tronchet F, Munzinger J, Pillon Y, Veillon J-M, Chalopin M. 2012. Le référentiel taxonomique Florical et les caractéristiques de la flore vasculaire indigène de la Nouvelle-Calédonie. Adansonia 34: 177-219.

Mort ME, Soltis DE, Soltis PS, Francisco-Ortega J, Santos-Guerra A. 2002. Phylogenetics and evolution of the Macaronesian clade of Crassulaceae inferred from nuclear and chloroplast sequence data. Systematic Botany 27: 271-288.

Murienne J, Grandcolas P, Piulachs M, Bellés X, D'Haese C, Legendre F, Pellens R, Guilbert E. 2005. Evolution on a shaky piece of Gondwana: is local endemism recent in New Caledonia? Cladistics 21: 2-7.

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J. 2000. Biodiversity hotspots for conservation priorities. Nature 403: 853-858.

Pelletier B. 2006. Geology of the New Caledonia region and its implications for the study of the New Caledonian biodiversity. In: Payri C, Richer de Forges B. (Eds.), Compendium of marine species from New Caledonia, Documents Scientifiques et Techniques II4, Institut de Recherche pour le Développement Nouméa, pp. 17-30.

Pillon Y, Munzinger J, Amir H, Lebrun M. 2010. Ultramafic soils and species sorting in the flora of New Caledonia. Journal of Ecology 98: 1108-1116.

Pillon Y. 2012. Time and tempo of diversification in the flora of New Caledonia. Botanical Journal of the Linnean Society 170: 288-298.

Schluter D. 1996. Adaptive radiation along genetic lineages of least resistance. Evolution 50: 1766-1774.

Schönenberger J, Anderberg AA, Sytsma KJ, 2005. Molecular phylogenetics and patterns of floral evolution in the Ericales. International Journal of Plant Sciences 166: 265-288.

Stuessy T, Crawford DJ, Marticorena C. 1990. Patterns of phylogeny in the endemic vascular flora of the Juan Fernandez Islands, Chile. Systematic Botany 15: 338-346.

Stuessy TF, Jakubowsky G, Salguero-Gómez R, Pfosser M, Schluter PM, Fer T, Sun B-Y, Kato H. 2006. Anagenetic evolution in island plants. Journal of Biogeography 33: 1259-1265.

Stuessy TF. 2007. Evolution of specific and genetic diversity during ontogeny of island floras: the importance of understanding process for interpreting island biogeographic patterns. In: Ebach MC, Tangney RS (Eds.) Biogeography in a changing world. CRC Press, Boca Raton, pp. 117-133.

Tamura M, Tao R, Yonemori K, Utsunomiya N, Sugiura A. 1998. Ploidy level and genome size of several *Diospyros* species. Journal of the Japanese Horticultural Society 67, 306-312.

Wallace AR. 1881. Island life. Macmillan and Co., London, UK.

Wallnöfer B. 2004. A revision of *Lissocarpa* Benth. (Ebenaceae subfam. Lissocarpoideae (Gilg in Engler) B.Walln.). Annalen des Naturhistorischen Museums Wien 105B: 515-564.

White F. 1992. Twenty-two new and little known species of *Diospyros* (Ebenaceae) from New Caledonia with comments on section *Maba*. Bulletin du Muséum national d'Histoire naturelle 4ème série – section B, Adansonia 2: 179-222.

White F. 1993. Flore de la Nouvelle-Calédonie et Dépendances. 19. Ébénacées. Muséum National d'Histoire Naturelle, Paris, France.

Whittaker RJ, Triantis KA, Ladle RJ. 2008. A general dynamic theory of oceanic island biogeography. Journal of Biogeography 35: 977-994.

Whittaker RJ. 1998. Island biogeography. Oxford University Press, Oxford, UK.

Picture credits of figure 1:

A, G, M, N, O, P, T, U: Jérôme Munzinger

B, E, F, K, L, R, V, W, X: Daniel & Iréne Létocart

C: Benoît Henry

D: Céline Chambrey

H: Jean-Louis Ruiz

I, Q, S, Y: Vanessa Hequet

J: Barbara Turner

CHAPTER 1

Molecular phylogenetics of New Caledonian *Diospyros* (Ebenaceae) using plastid and nuclear markers

Barbara Turner, Jérôme Munzinger, Sutee Duangjai, Eva M. Temsch, Reinhold Stockenhuber, Michael H. J. Barfuss, Mark W. Chase, Rosabelle Samuel

Status: published, Molecular Phylogenetics and Evolution 69 (2013) pp. 740-763

Contribution: Collection of material, collection of data, analysis of data, phylogenetic analysis, manuscript writing/editing

Molecular Phylogenetics and Evolution 69 (2013) 740–763

Contents lists available at SciVerse ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Molecular phylogenetics of New Caledonian *Diospyros* (Ebenaceae) using plastid and nuclear markers ☆

Barbara Turner [a,*], Jérôme Munzinger [b], Sutee Duangjai [c], Eva M. Temsch [a], Reinhold Stockenhuber [d], Michael H.J. Barfuss [a], Mark W. Chase [e,f], Rosabelle Samuel [a]

[a] Department of Systematic and Evolutionary Botany, Faculty of Life Sciences, University Vienna, Rennweg 14, 1030 Wien, Austria
[b] IRD, UMR AMAP, TA A51/PS2, 34398 Montpellier Cedex 5, France
[c] Department of Forest Biology, Faculty of Forestry, Kasetsart University, Bangkok, Thailand
[d] Department of Evolutionary Biology and Environmental Sciences, University Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland
[e] Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3DS, UK
[f] School of Plant Biology, The University of Western Australia, Crawley, WA 6009, Australia

ARTICLE INFO

Article history:
Received 25 January 2013
Revised 1 July 2013
Accepted 3 July 2013
Available online 12 July 2013

Keywords:
Endemism
Genome size
Island flora
Low-copy nuclear markers
Molecular dating

ABSTRACT

To clarify phylogenetic relationships among New Caledonian species of *Diospyros*, sequences of four plastid markers (*atpB*, *rbcL*, *trnK–matK* and *trnS–trnG*) and two low-copy nuclear markers (*ncpGS* and *PHYA*) were analysed. New Caledonian *Diospyros* species fall into three clades, two of which have only a few members (1 or 5 species); the third has 21 closely related species for which relationships among species have been mostly unresolved in a previous study. Although species of the third group (NC clade III) are morphologically distinct and largely occupy different habitats, they exhibit little molecular variability. *Diospyros vieillardii* is sister to the rest of the NC clade III, followed by *D. umbrosa* and *D. flavocarpa*, which are sister to the rest of this clade. Species from coastal habitats of western Grande Terre (*D. cherrieri* and *D. veillonii*) and some found on coralline substrates (*D. calciphila* and *D. inexplorata*) form two well-supported subgroups. The species of NC clade III have significantly larger genomes than found in diploid species of *Diospyros* from other parts of the world, but they all appear to be diploids. By applying a molecular clock, we infer that the ancestor of the NC clade III arrived in New Caledonia around 9 million years ago. The oldest species are around 7 million years old and the youngest ones probably much less than 1 million years.

1. Introduction

New Caledonia is an island group located in the southwestern Pacific about 1300 km east of Australia, ranging from around 19° to 23° south with an land area of ca. 19,000 km^2. It consists of the main island Grande Terre (ca. 16,000 km^2), Iles Belep (in the north), Ile des Pins (in the south), Loyalty Islands (in the east) and several other smaller islands. The continental part of New Caledonia (mainly Grande Terre) separated from Gondwanan during late Cretaceous (ca. 80 million years ago, mya; McLoughlin, 2001). During the Palaeocene to late Eocene, this continental sliver was submerged for at least 20 million years (myr), and a thick layer of oceanic mantle accumulated (Pelletier, 2006). After Grande Terre re-emerged in the late Eocene (37 mya), this heavy-metal rich oceanic material covered most of the land. Today, around 1/3 of the main island is still covered with ultramafic substrates. Because Grande Terre was totally submerged, it is highly unlikely that lineages that were already present in this area before the split from Gondwanan could have survived locally. Current hypotheses suggest that biota present today are derived from elements/ancestors that reached New Caledonia via long distance dispersal (e.g. Morat et al., 2012; Pillon, 2012; Grandcolas et al., 2008) mainly from Australia, New Guinea and Malaysia. Hypotheses of other islands between Australia and New Caledonia having served as stepping stones or refuges for Gondwanan taxa now endemic (e.g. *Amborella*) have been proposed by a few authors (Ladiges and Cantrill, 2007), but there is no consensus of when they existed or how large and numerous they might have been. The New Caledonian climate is tropical to subtropical. The main island is split by a mountain range into a humid eastern portion (2000–4000 mm precipitation per year) and a dry western part (1000 mm precipitation per year) with prevailing winds and rain coming from the south east. New Caledonia is one of the 34 biodiversity hotspots (Mittermeier et al., 2004; Myers et al., 2000), and nearly 75% of the native flora is endemic (Morat et al., 2012), which is the fourth highest for an island (Lowry, 1998). Among these endemic taxa are 98 genera and three families, Amborellaceae, Oncothecaceae and Phellinaceae (Morat et al.,

* Corresponding author.
E-mail address: barbara.turner@univie.ac.at (B. Turner).

1055-7903/$ - see front matter © 2013 The Authors. Published by Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ympev.2013.07.002

2012). One of the reasons hypothesised for the high level of endemism found in New Caledonia is the ultramafic substrates, which have acted as a filter for colonising species that were already pre-adapted to this special soil (Pillon et al., 2010).

Ebenaceae are pantropical and belong to the order Ericales (APG, 2009); the majority of species occur in Africa (incl. Madagascar) and the Indo-Pacific region. Duangjai et al. (2006) divided Ebenaceae into two sub families, Lissocarpoideae and Ebenoideae. Lissocarpoideae are monogeneric (*Lissocarpa*, 8 species in northwestern South America), and Ebenoideae include *Diospyros*, *Euclea* (18 species in Africa) and *Royena* (17 species in Africa). This classification of Ebenaceae in two subfamilies and four genera has been also supported by palynological data (Geeraerts et al., 2009).

In this paper, we use the circumscription of *Diospyros* as proposed by Duangjai et al. (2006). *Diospyros* is the largest genus of Ebenaceae with more than 500 species, making it also one of the largest angiosperm genera. The greatest species of diversity is in Asia and the Pacific region (~300 species). Fruits of some species (persimmons; e.g. *D. kaki*, *D. lotus* and *D. virginiana*) are edible, and ebony wood (e.g. *D. ebenum*) is one of the most expensive timbers. Species of *Diospyros* are shrubs or trees that occur in most tropical and subtropical habitats, where they are often important and characteristic elements. Duangjai et al. (2009) found 11 mostly well-resolved clades within *Diospyros*. In New Caledonia, there are 31 described *Diospyros* species, of which all but one are endemic, and they belong to three clades (Duangjai et al., 2009; Fig. 4, clades II, III and XI). The first clade (clade II) contains five species from New Caledonia that are related to Australian species of *Diospyros*. The second clade (clade III) includes species from Hawai'i, Indian Ocean islands and 24 taxa from New Caledonia, within which the species from New Caledonia form a sublcade, here termed NC clade III. Although Duangjai et al. (2009) analysed more than 8000 base pairs of plastid DNA, low variability and little resolution was found among these endemic New Caledonian species. The third clade (clade XI), consisting of taxa from Asia, America, Pacific Islands and New Caledonia, includes two *Diospyros* species from New Caledonia, one endemic and the other found throughout the southern Pacific. These two species are not sister species, accounting for two more colonisations of New Caledonia (i.e. four in total). Similar, multiple colonisation events are also found among other organisms in New Caledonia (e.g. Murienne et al., 2005). *Diospyros* is observed in all types of New Caledonian vegetation except mangrove; the species range from sea level up to ca. 1250 m (New Caledonia's highest point is 1628 m). There are several micro-endemics restricted to just a small area (White, 1992). Most New Caledonian *Diospyros* species from clade III are morphologically clearly defined and restricted by edaphic factors and occur on just one substrate type. For example, *D. labillardierei* (Fig. 1D) is distinctive with its long narrow leaves and *Salix*-like habit; it is a rheophyte on non-ultramafic substrates. *Diospyros veillonii* (Fig. 1F) is a remarkable species with coralloid inflorescence axes (unique among New Caledonian *Diospyros*) and large leaves, but is known from only a single locality in dry forest on black clay soil. Other species have broader distributions and ecologies, such as *D. parviflora* (Fig. 1J), which grows on both ultramafic and non-ultramafic substrates and is widespread throughout Grande Terre and Balabio Island in dense humid forests as well as in more open and dry vegetation. Some species can have similar ecological requirements, but are morphologically well differentiated; for example *D. vieillardii* (Fig. 1A) has a calyx narrower than its prune-like fruit, whereas *D. glans* (Fig. 1N) has a thick calyx much wider than its fruit, but both grow in maquis vegetation and co-occur at some sites.

For establishing phylogenetic relationships, sequences of low-copy nuclear genes are not as often used as regions from the plastid genome, often due to methodological difficulties. Low-copy genes are present in one or few copies in the genome, and primers are often highly specific for individual groups, requiring them therefore to be newly designed for each study. On the other hand, low-copy nuclear markers are normally highly informative and as they are biparentally inherited they may also help detect recent hybridization (e.g. Moody and Rieseberg, 2012). However, in a study of Hawaiian endemics in two unrelated genera, Pillon et al. (2013) found that although two low-copy nuclear loci displayed a high level of variability, they also exhibited heterozygosity, intraspecific variation, and retention of ancient alleles; allele coalescence was older than the species under study. Nonetheless, we hoped that inclusion of low-copy nuclear genes might provide additional insight into species relationships and thus included two such loci. Phytochrome A (*PHYA*) belongs to the gene family of the phytochromes, which has eight members across the seed plants (*PHYA–PHYE* in angiosperms and *PHYN–PHYP* in gymnosperms); *PHYN*/*PHYA*, *PHYO*/*PHYC* and *PHYP*/*PHYBDE* are orthologs, the rest being paralogs of the others (Mathews et al., 2010). Genes of this family encode photoreceptor proteins that mediate developmental responses to red and far red light. The three main paralogs (*PHYA*, *PHYB* and *PHYC*) are different enough to be amplified with specific primers (Zimmer and Wen, 2012). Sequences of *phy* genes have been used successfully across the flowering plants (e.g. Mathews et al., 2010; Nie et al., 2008; Bennett and Mathews, 2006) for phylogenetic reconstruction. The gene *PHYA* used in this study consists of four exons and three introns. Glutamine synthetase (*GS*), codes for a protein involved in nitrogen assimilation. There are two main types of *GS* genes, cytosolic- and chloroplast-expressed. Chloroplast-expressed glutamine synthetase (*ncpGS*) consists of 12 exons and 11 introns and has been shown to be a single-copy gene in plants (Emshwiller and Doyle, 1999). This combination of coding and non-coding regions has been shown to be highly informative for inferring phylogenetic trees of various groups (e.g. Oxalidaceae, Emshwiller and Doyle, 1999; *Passiflora*, Yockteng and Nadot, 2004; *Spiraeanthemum*, Pillon et al., 2009a; *Codia*, Pillon et al., 2009b; *Achillea millefolium*, Guo et al., 2012).

Beside phylogenetic relationships, the age of clades is of interest. In many cases, there are no fossils available for direct dating of a group of interest in a particular region, which is often the case for islands and is certainly true for New Caledonia (the few fossils recorded to date are older than the last emergence of the island and are not certain to be angiosperms; Salard and Avias, 1968). Rates of DNA divergence are generally consistent with a molecular clock (Zuckerkandl and Pauling, 1965), and therefore DNA data contain information about the relative ages of taxa. When substitution rates (e.g. Silvestro et al., 2011; Alba et al., 2000) or fossils belonging to defined clades (e.g. Pirie and Doyle, 2012; Magallón, 2010) are taken into consideration, the relative ages obtained can be transformed into absolute ages. Placement of fossils in the correct position in the phylogenetic tree is crucial for accurate interpretation (Forest, 2009). Some previous studies have has been published on the subject of the age of asterids (e.g. Millán-Martínez, 2010; Bell et al., 2010; Bremer et al., 2004) to which Ericales belong, and fossil *Diospyros* are known from some localities (mainly in India and North America), but none has been found in New Caledonia. *Austrodiospyros cryptostoma* (Basinger and Christophel, 1985), a fossil from Australia has many morphological similarities to *D. australis* of clade II (Duangjai et al., 2009). It is thus far the only fossil belonging to a clade that includes *Diospyros* species from New Caledonia. We treat *A. cryptostoma* as member of clade II in this study.

Genome sizes vary nearly 2400-fold across angiosperms (Pellicer et al., 2010). Most variation in DNA amount is caused by different amounts of non-coding, repetitive DNA, such as pseudogenes, retrotransposons, transposons and satellite repeats (Leitch, 2007; Bennett and Leitch, 2005; Parisod et al., 2009; Petrov, 2001). Genome sizes and chromosome numbers of *Diospyros* are within the range of those of other members of Ericales (Bennett and

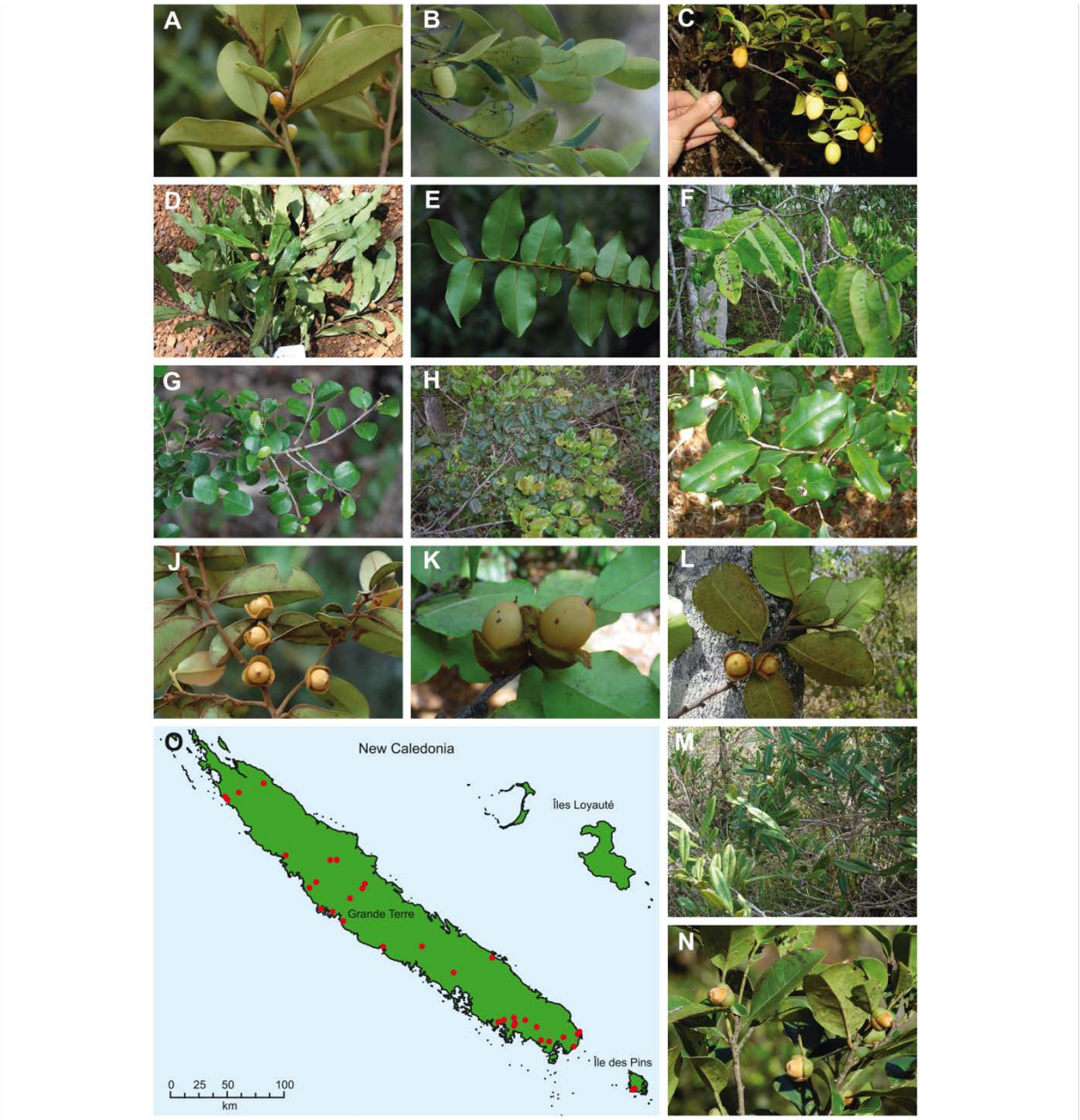


Fig. 1. Examples of *Diospyros* species from New Caledonia (A–N) and Map of New Caledonia with collection points (O). A: *D. vieillardii*; B: *D. umbrosa*; C: *D. flavocarpa*, D: *D. labillardierei*; E: *D. pancheri*, F: *D. veillonii*; G: *D. minimifolia*; H: *D. pustulata*; I: *D. cherrieri*; J: *D. parviflora*; K: *D.perplexa*; L: *D. yaouhensis*; M: *D. revolutissima*; N: *D. glans*; O: Map of New Caledonia with sampling localities. Photographs taken by: C. Chambrey (I), V. Hequet (F, K, L), J. Munzinger (A, B, C, E, G, H, J, M, N) and B. Turner (D).

Leitch, 2010). Nuclear DNA amounts in *Diospyros* range from 0.78 pg (1C-value) in diploid *D. rhodocalyx* up to 4.06 pg in nonaploid *D. kaki* cultivars (Tamura et al., 1998). The basic chromosome number in *Diospyros* is $2n = 2x = 30$, and most species seem to be diploid (e.g. Tamura et al., 1998; White, 1992). There are some reports of polyploid *Diospyros*, mostly from cultivated species (e.g. *D. rhombifolia* 4*x*, *D. ebenum* 6*x*, *D. kaki* 6*x* and 9*x*, *D. virginiana* 6*x* and 9*x*; Tamura et al., 1998). White (1992) provided chromosome counts for nine New Caledonian species of *Diospyros* (*D. calciphila*, *D. fasciculosa*, *D. flavocarpa*, *D. minimifolia*, *D. olen*, *D. parviflora*, *D. umbrosa*, *D. vieillardii* and *D. yaouhensis*), all of which are diploid.

Duangjai et al. (2009) found little sequence variation in the markers investigated among many species from NC clade III, which could indicate recent diversification. White (1992), who described most the New Caledonian *Diospyros* species, suspected some hybridization was taking place. The main aim of this study was to clarify relationships among New Caledonian *Diospyros* species, especially of those belonging to clade III (Duangjai et al., 2009).

Furthermore, if we were able to find more variable than those previously studied, we wanted to elucidate potential factors underlying speciation (e.g. ecological speciation, hybrid speciation and introgression) and understand better differences in speciation rates of the clades that reached New Caledonia independently. We used low-copy nuclear markers, *PHYA* and *ncpGS* because they offered the prospect of resolving relationships within this clade and detecting possible hybrid species. We also included samples from nine additional species that were not available for the study of Duangjai et al. (2009). Moreover, we conducted dating analyses to obtain estimates of the ages for the lineages to which New Caledonian *Diospyros* species belong. We also present chromosome numbers and genome sizes of some additional New Caledonian species of *Diospyros*; we wished to examine further the hypothesis that polyploidy (perhaps involving hybridization) might have played a role in producing diversity in this comparatively species-rich clade.

2. Materials and methods

2.1. Material

Material from New Caledonian *Diospyros* species was collected by B. Turner (BT), J. Munzinger (JM), Yohan Pillon (YP) or Vanessa Hequet (VH). When fertile, a voucher was made with several duplicates sent to various herbaria. When sterile, one voucher per population was taken; this was compared to already existing collections in Noumea Herbarium (NOU) from the same location and referred to that species if similar. One putatively new species was detected while doing fieldwork for this project, here called *D.* sp. Pic N'ga. Other Ebenaceae samples are from previous studies (Duangjai et al., 2009). Outgroup taxa and a few *Diospyros* samples were taken from the Royal Botanic Gardens, Kew, DNA Bank (http://apps.kew.org/dnabank/homepage.html). Compared to the sampling of Duangjai et al. (2009), we added material of the following New Caledonian species: *D. erudita*, *D. glans*, *D. impolita*, *D. inexplorata*, *D. margaretae*, *D. tireliae*, *D. tridentata*, *D. trisulca* and *D. veillonii* (for details see Table 1). The three un-sampled species from New Caledonia (*D. fastidiosa*, *D. nebulosa* and *D. neglecta*) are rare and have not been seen after their description.

2.2. DNA extraction

For DNA extraction the sorbitol/high-salt CTAB method (Tel-Zur et al., 1999), modified for 2 ml micro-centrifuge tubes, was used. Tubes containing silica gel-dried material were frozen with liquid nitrogen (to keep material frozen during grinding to avoid enzymatic action) and then ground with glass-beads to a fine powder. Prior to extraction, ground material was washed three times with sorbitol buffer.

2.3. PCR and cycle sequencing

We sequenced four plastid regions: *atpB*, *rbcL*, *trnK–matK* (partial *trnK* intron and complete *matK* gene) and *trnS–trnG*, which collectively represent approximately 6.5 kb. Primers and PCR conditions are those of Duangjai et al. (2009). We added 136 accessions to the matrix of Duangjai et al. (2009).

Chloroplast-expressed glutamine synthetase (*ncpGS*) was amplified with primers designed for this study (GScpDio1F and GScpDioR; Table 4). Initial *Diospyros* sequences for primer design were obtained with the primers and PCR protocol of Yockteng and Nadot (2004). Primers were situated at the end of exon 7 (forward) and beginning of exon 11 (reverse), amplifying a fragment between 700 and 715 bp (Fig. 2). Primers used for PCR were also used for cycle sequencing (Tables 2 and 3).

Initial PCR products and sequences of *PHYA* were obtained with the locus-specific primers of Mathews and Donoghue (1999; *PHYA* upstream [2nd] and *PHYA* downstream [1st]). As these primers were not specific enough, we cloned the PCR products (see Section 2.4) to be able to design *Diospyros*-specific *PHYA* PCR and sequencing primers (PhyADioF, PhyADioR, PhyADioFi and PhyADioRi; Table 4; Fig. 3). However, as the new PCR primers designed for *Diospyros* did not amplify consistently, we used a two-step amplification protocol. In the first PCR, the universal *PHYA* primers were used, and then a second nested PCR was performed with the newly designed primers and the product from the first PCR as template. All primers are located in exon 1 of *PHYA* flanking a region of 1187 bp in length. PCR conditions and composition are provided in Tables 2 and 3. For cycle-sequencing, we used the two internal primers and the external reverse primer.

PCR products were cleaned with a mixture of exonuclease I and alkaline phosphatase (10 units exo I and one unit FastAP, both from Thermo Scientific) and incubated at 37 °C for 45 min followed by 15 min at 80 °C to inactivate enzymes. Cycle sequencing reactions were performed with 0.8 µl BigDye Terminator v3.1 (AB, Life Technologies), 1.0 µl primer (3.2 µM), 1.6 µl 5× sequencing buffer and 6.6 µl cleaned-up PCR product using 35 cycles of 96 °C for 10 s, 50 °C for 5 s and 60 °C for 3 min. Sequences were produced on a capillary sequencer (3730 DNA Analyzer, AB, Life Technologies) following the manufacturer's protocols.

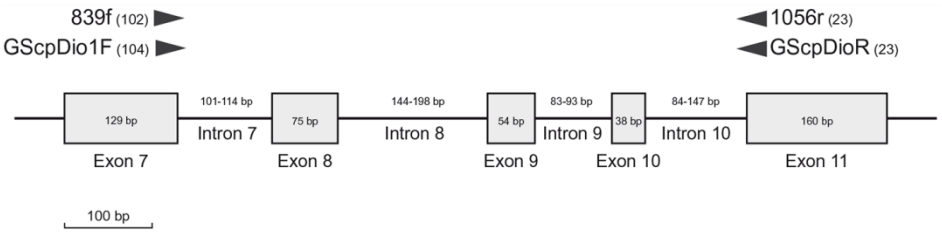


Fig. 2. Schematic diagram of exon 7–exon 11 of *ncpGS* with primer positions and length of exons and introns. Numbers in parentheses give 5′ end of primers.

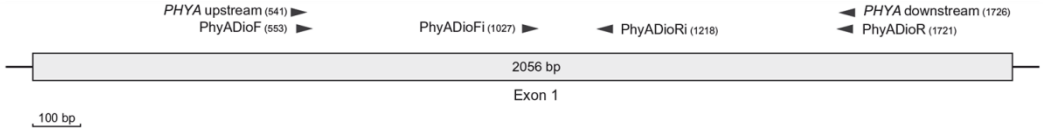


Fig. 3. Schematic diagram of exon 1 of *phyA* with primer positions and length of exon. Numbers in parentheses give 5′ end of primers.

2.4. Cloning

Cloning was needed to produce *PHYA* from some accessions; these where than used for development of more specific primers. In addition, cloning of samples was necessary when we failed to obtain good sequences with the *Diospyros*-specific primers. PCR products were obtained using the universal *PHYA* primers, and after gel purification (Inivsorb Spin DNA Extraction Kit, Invitek), cleaned products were cloned using the pGEM-T Easy cloning system (Promega), following the manufacturer's protocol. Cloned fragments were amplified using M13-f47 and M13-r48 primers and the following PCR conditions: initial denaturation 94 °C for 3 min, 35 cycles of denaturation 94 °C for 30 s, annealing 62 °C for 30 s and extension 72 °C for 2 min followed by a final extension at 72 °C for 7 min.

2.5. Sequence assembly, and editing, and phylogenetic analyses

Assembly and editing of sequences was done with the SeqMan Pro of the Lasergene v8.1 software package (DNASTAR); alignment was conducted with MUSCLE v3.8 (Edgar, 2004) and inspected visually using BioEdit v7.0.4 (Hall, 1999). Discrimination between the two copies of *PHYA* that were recovered from some species was done based on the alignment, and the 'wrong' (highly divergent) copy was excluded from further analyses. To test congruence between the data sets, ILD (incongruence length difference) test (Farris et al., 1994) implemented in PAUP* v4b10 (Swofford, 2003; termed the "partition homogeny test") was carried out with 100 replicates. To speed up this analysis, the neo-endemic clade (where resolution is low due to lack of variability and therefore congruence is unlikely to be detected) was reduced to two accessions (*D.* sp. Pic N'ga BT318 and *D. vieillardii* BT025). Results of the ILD test indicated congruence of the four plastid data sets, and therefore the plastid data sets were combined; jModeltest indicated the same model could be used in all analyses without partitioning. Phylogenetic analyses were performed using PAUP* v4b10 (Swofford, 2003) for maximum parsimony (MP) and RaxML (Stamatakis, 2006) for maximum likelihood (ML) analyses. For both methods, bootstrap with 1000 replicates was performed to estimate clade support. For Bayesian inference, the program BEAST v1.7.4 (Drummond et al., 2012) was used. Parsimony and Bayesian analyses were run on the Bioportal computer cluster of the University Oslo (www.bioportal.uio.no), and likelihood analyses were run on CIPRS Science Gateway (http://www.phylo.org/portal2/; Miller et al., 2010). Estimation of evolutionary models and values was conducted with jModeltest v2.0.1 (Darriba et al., 2012; Guindon and Gascuel, 2003). For the Bayesian analyses the general time reversible nucleotide substitution model (GTR; Tavaré, 1986) with among site rate variation modelled with a gamma distribution (GTR + Γ) was used for *ncpGS*, whereas for plastid data the same model was used but with a proportion of invariable sites (GTR + Γ + I). For *PHYA* the Hasegawa–Kishino–Yano nucleotide substitution model (HKY; Hasegawa et al., 1985) was used with among site rate variation modelled with a gamma distribution and a proportion of invariable sites (HKY + Γ + I). Base frequencies (uniform), substitution rates between bases (gamma shape 10), alpha (gamma shape 10), kappa (gamma shape 10) and *p*-inv (uniform) were inferred by Modeltest from each data set. We used a relaxed uncorrelated log-normal clock model (Drummond et al., 2006). As speciation model, we used a Yule model (Gernhard, 2008; Yule, 1925). For further details see Supplementary material S1. Two independent Metropolis-coupled Markov chain Monte Carlo (MCMC) analyses each with 20 million generations were run sampling each 1000th generation. The initial 10% of trees obtained from each MCMC run were removed as burn in; the remaining trees of both runs were used to calculate a maximum clade credibility tree.

2.6. Dating the tree

To obtain an overarching dated tree, we used parts (*atpB* and *rbcL* sequences of Cornales and Ericales) of the data set of Bell et al. (2010) and combined it with our matrix. This matrix consisted of two plastid markers (*atpB* and *rbcL*), which were analysed as two partitions. Dating analyses were run in BEAST with an uncorrelated log-normal relaxed clock under the GTR + Γ + I model. The tree was calibrated with two fossils, *Paleoenkianthus sayrevillensis* (90 myr; Nixon and Crepet, 1993) as minimum age for Ericales and *A. cryptostoma* (34 myr; Basinger and Christophel, 1985) as minimum age for *Diospyros* clade II. Both groups (Ericales and *Diospyros* calde II) were defined as monophyletic, including the stem. Following tmrca (time of most recent common ancestor) settings used were: log normal prior distribution with a mean of 1.5, log standard deviation of 0.5 and an offset of 89 (Ericales) and 33 (*Diospyros* clade II). Priors for the molecular clock were: ucld.stdev: log normal, mean 0.9, log stdev 1, initial value 0.5, mean in real space; ucld.mean: CTMC rate reference (Ferreira and Suchard, 2008, initial value 1. Details of settings for BEAST analysis are provided in Section 2.5 (above) and Supplementary material S2. In addition to the plastid marker dating, we also conducted an analysis with our combined data set. We used the same settings as for the Bayesian analysis, but we added two calibration points: *A. cryptostoma* at 34 myr (Basinger and Christophel, 1985) as minimum age for *Diospyros* clade II and the split of *Diospyros* and its sister clade, *Euclea* plus *Royena*, 42 myr, which is the minimum age of that node based on dating exercises with the plastid markers. All settings for the molecular clock were the same as those for the plastid data set. The input file used for dating the combined analysis is provided in Supplementary material S3.

2.7. Chromosome counts of Diospyros

Chromosome preparations were made using Feulgen staining following the protocol from Weiss-Schneeweiss et al. (2009). Root tips were collected from plant material growing in the Botanical Garden of the University of Vienna (HBV) and a private garden in New Caledonia. To arrest mitotic spindles, root tips were treated with 0.002 M 8-hydroxquinoline for 2 h at room temperature and 2 h at 4 °C (always in darkness because 8-hydroxquinoline is light sensitive). Pre-treated material was fixed for 12 h at room temperature in 3:1 ethanol:acetic acid and then stored at −20 °C until examined. Fixed root tips were washed in distilled water to remove fixative, hydrolysed in 5 N HCl for 30 min, washed again with distilled water and stained with Schiff's reagent for approximately 2 h in the dark. Squash preparations were made under a coverslip in a drop of 45% acetic acid. Counts could only be made for few species because obtaining young, actively growing root-tips from New Caledonian *Diospyros* is difficult. Collecting root-tips from forest trees and shrubs is not possible because there are too many roots in the soil to determine which is from the plant of interest. An alternative method is to grow seedlings in the lab/greenhouse. Obtaining seeds from tropical plants is not easy because these species do not produce fruit at a specific time of the year and flowering is diffuse (only few flowers produced at a time), so one would have to visit the plants regularly for at least 1 year to collect seed material. The logistics of this in process in New Caledonia were difficult. In addition, we found germination of seeds and maintenance of *Diospyros* seedlings highly problematic. Fortunately, the material we were able to obtain is well distributed among the genome sizes obtained, so we can conclude more than would otherwise be possible.

2.8. Genome size estimations of Diospyros

Genome size was determined using flow cytometry performed on leaf material. Fresh tissue was used from plants growing in the

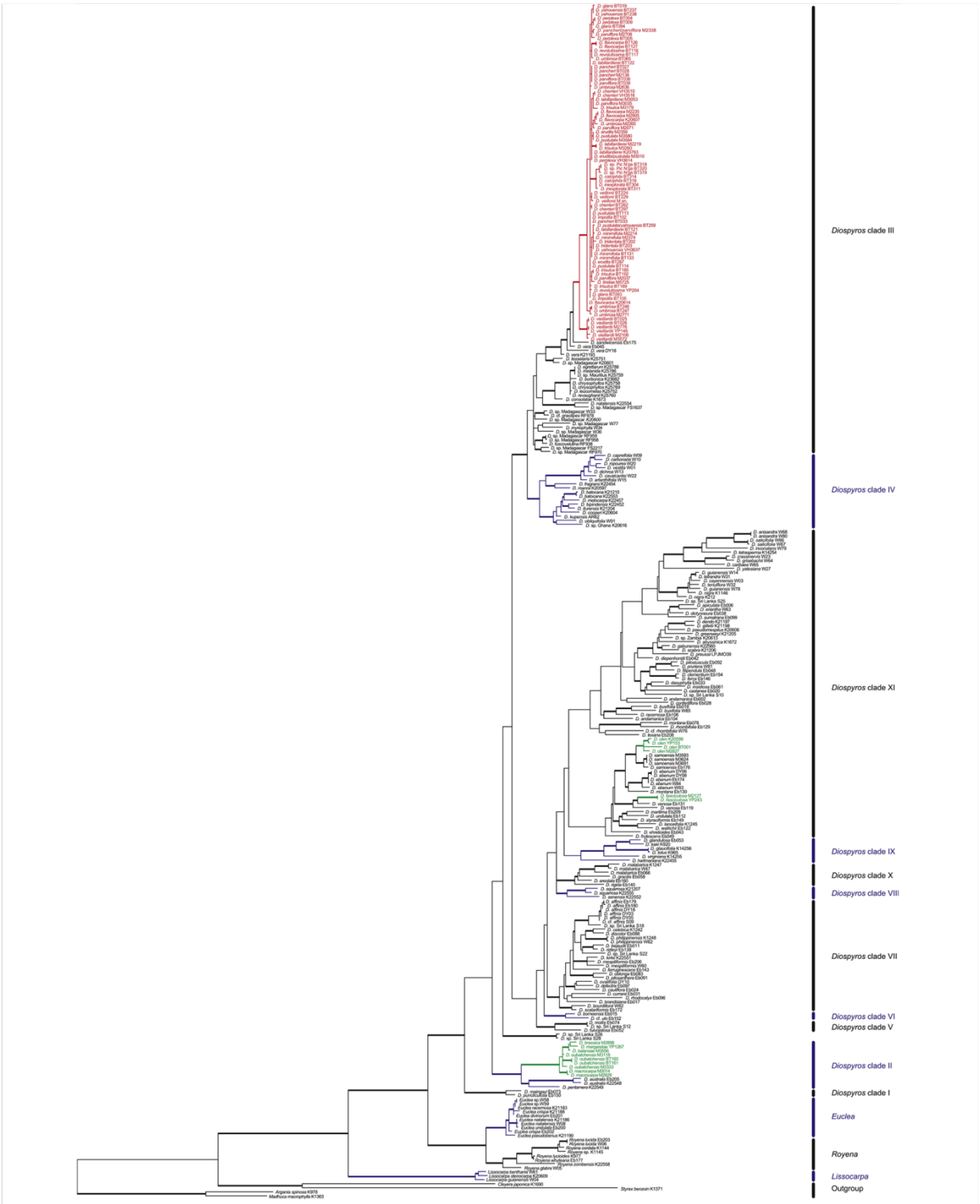


Fig. 4. One of 210 equally parsimonius trees of the plastid data set. Clades are named according to Duangjai et al. (2009). Bold branches have more than 70% support in all three analysis. New Caledonian taxa are coloured, red represents clade III NC.

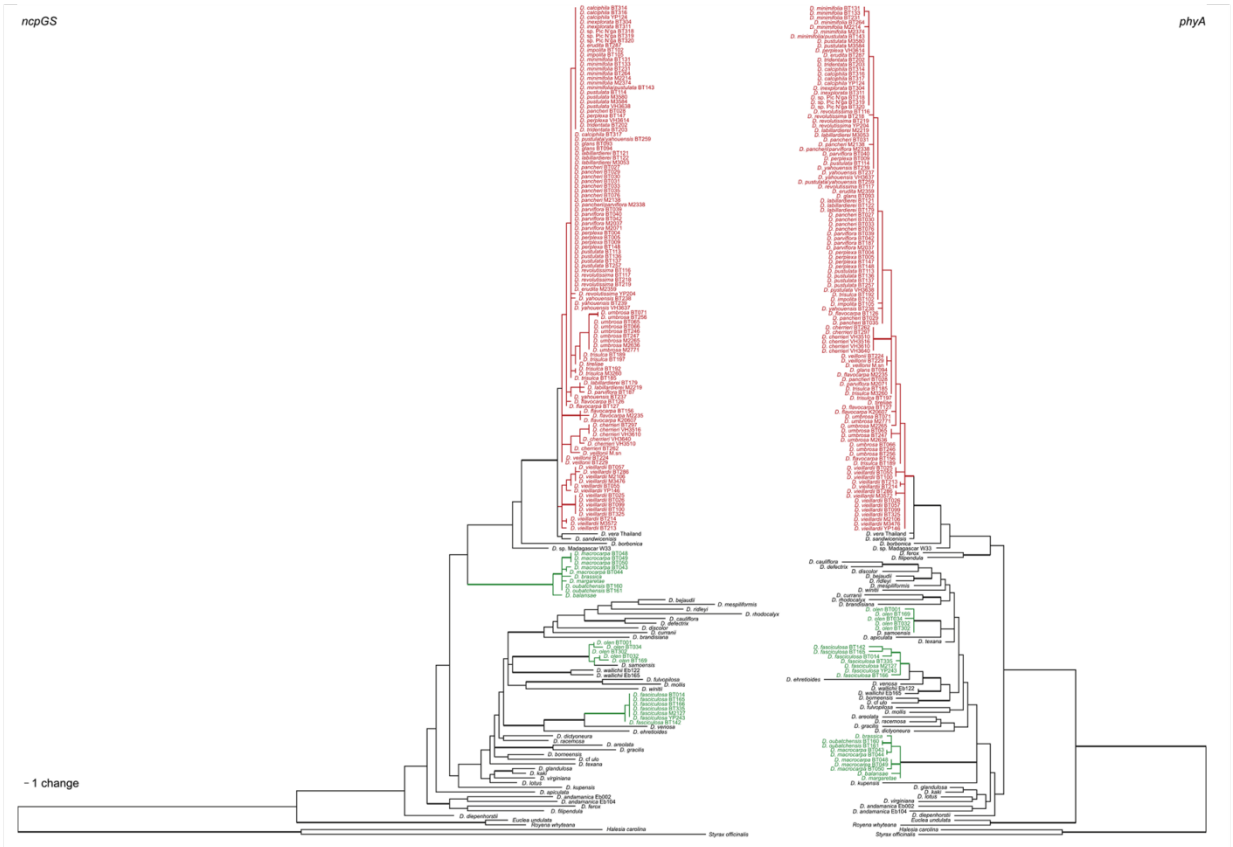


Fig. 5. Maximum parsimony trees inferred from the nuclear data sets, branch length scaled to same value on both trees. Bold branches have more than 70% support in all three analysis. New Caledonian taxa are coloured, red represents clade III NC. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

HBV. In addition, recently collected silica-gel dried material from New Caledonia was used for several measurements because it was not possible to transport fresh leaf material from New Caledonia to the laboratory. Samples were chopped in Otto I buffer (Otto et al., 1981) together with leaves of the internal standard species, *Solanum pseudocapsicum*, 1C = 1.30 pg (Temsch et al., 2010) or *Pisum sativum* Kleine Rheinländeriń, 1C = 4.42 pg (Greilhuber and Ebert, 1994), according to the method of Galbraith et al. (1983). The isolate was filtered through a 30 µm nylon mesh, and RNA was digested with 15 mg/l RNase A for 30 min at 37 °C. Subsequently, DNA was stained in propidium iodide (50 mg/l) supplemented with Otto II buffer (Otto et al., 1981). Mean fluorescence intensity of a total of 15,000 particles was measured with a CyFlow cytometer (Partec, Münster, Germany) equipped with a green laser (Cobolt Samba, Cobolt AB, Stockholm, Sweden); the 1C-value was calculated according to the formula: $(MFI_{object}/MFI_{Standard}) \times 1C\text{-}value_{Standard}$, where MFI is the mean fluorescence intensity of the G1 nuclei population. Statistical significance of asymmetry between the results obtained from *Diospyros* species belonging to clade III and those from clades VII–XI was tested using SPSS 15.0 (SPSS, Chicago; IL, USA) and the non-parametric Mann–Whitney U-test because of non-homogeneity of variances between the two groups of variables (Levene's test for equality of variances, $p < 0.05$).

3. Results

The data characteristics and statistics from the maximum parsimony analyses of all three individual and the combined data sets are provided in Table 5. Since the focus of this paper is the New Caledonian *Diospyros* species from clade III, only results pertaining to this group will be discussed in detail. The other species have been included to (i) investigate the utility of these markers for resolving phylogenetic relationships within *Diospyros* and (ii) further evaluate the hypothesis (proposed by Duangjai et al., 2009) that not all New Caledonian *Diospyros* resulted from a single colonisation event.

3.1. Plastid markers

Parsimony analysis of the plastid data set produced 210 equally parsimonious trees, one of which (randomly selected) is shown to demonstrate comparative levels of divergence (Fig. 4). Clade names correspond to those of Duangjai et al. (2009). Resolution among the New Caledonian taxa of clade III is low, but monophyly of these taxa is strongly supported: bootstrap percentage MP (BMP) 88; bootstrap percentage ML (BML) 97; Bayesian posterior probability (BPP) 0.95. Furthermore, *D. vieillardii* (BMP 99, BML 98, BPP 1.00) and its position as sister (BMP 97, BML 96, BPP 1.00) to the rest of the clade are well supported. Within the NC clade III, only one group of three taxa (*D. calciphila*, *D. inexplorata* and *D.* sp. Pic N'ga) is supported in all three analyses (BMP 91, BML 92, BPP 1.00); this includes all accessions of each species forming unique clusters. There are a few more, weakly supported small groups in which individuals of one population fall together, but they are not consistent among the three analyses and fail to include all accessions of these species.

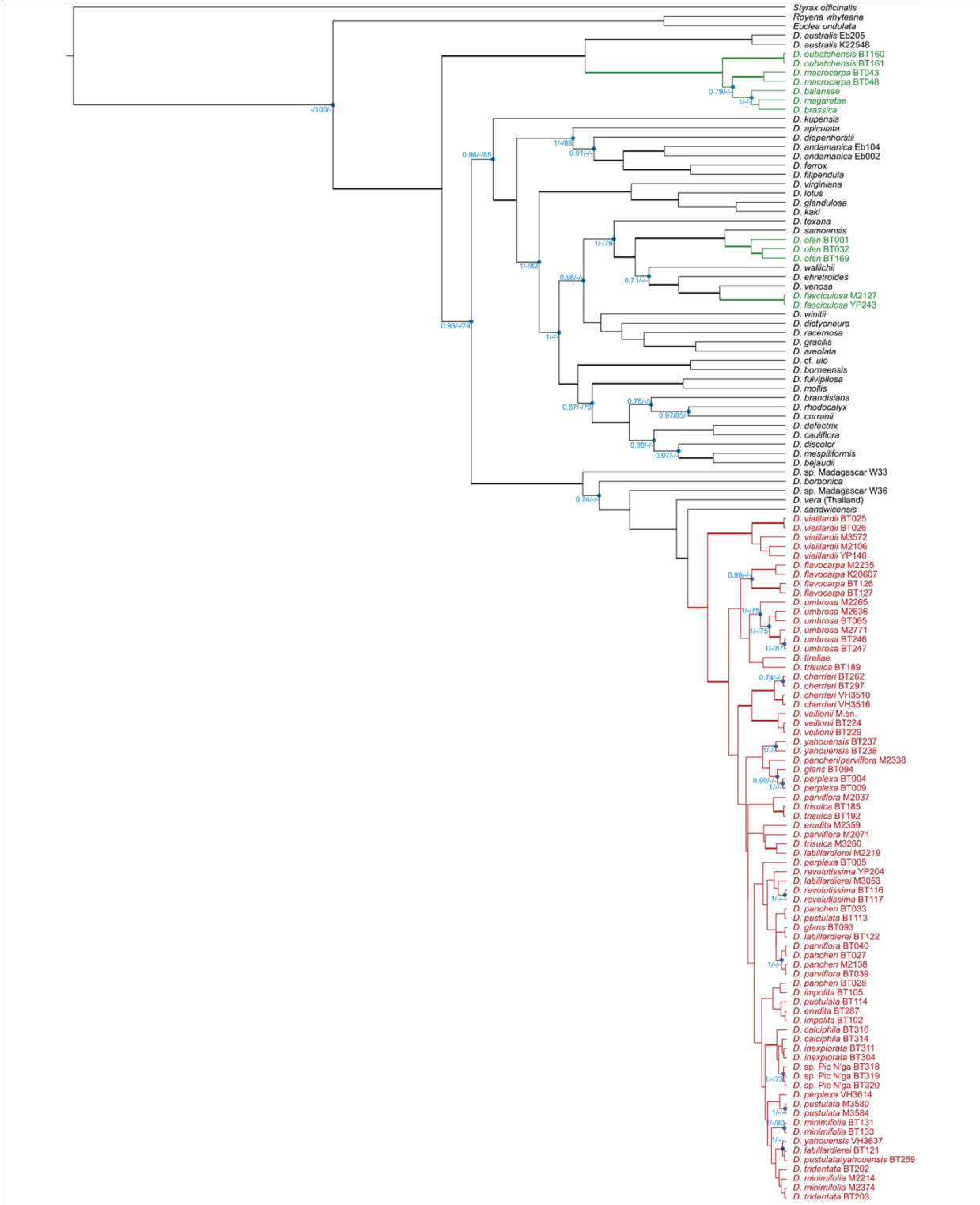


Fig. 6. Baysian maximum clade credibility tree inferred from the combined data set. Bold branches have more than 70% support in all three analysis, nodes with at least one support value ⩾70% are indicated with blue dots (BPP/BMP/BML). New Caledonian taxa are coloured, red represents clade III NC. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

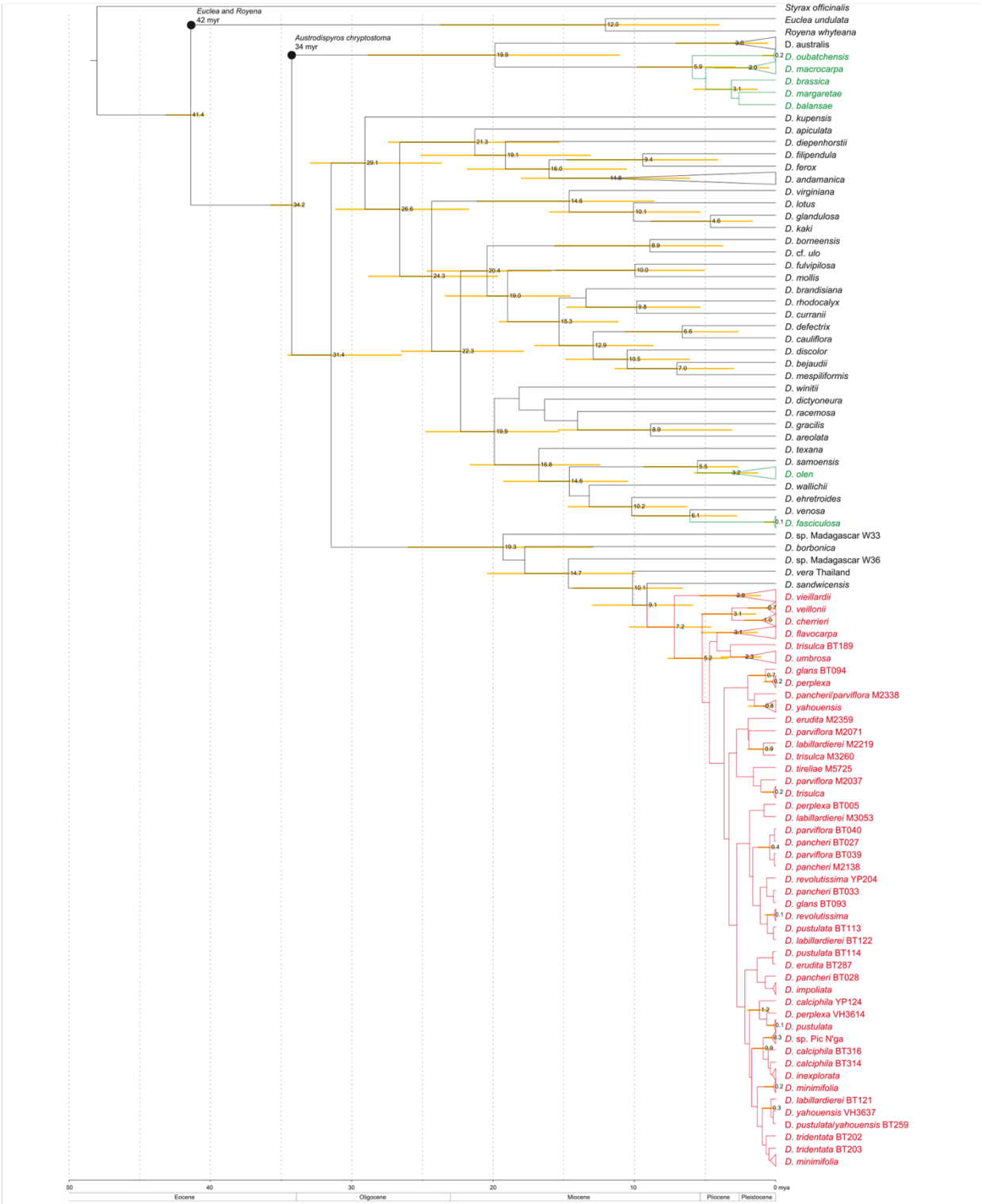


Fig. 7. Chronogram based on the combined data set. Ages are given (in million years) for nodes with more than 0.85 BPP. Nodes which were calibrated are marked with a black dot. Yellow bars represent the 95% highest posterior density interval. New Caledonian taxa are coloured, red represents clade III NC.

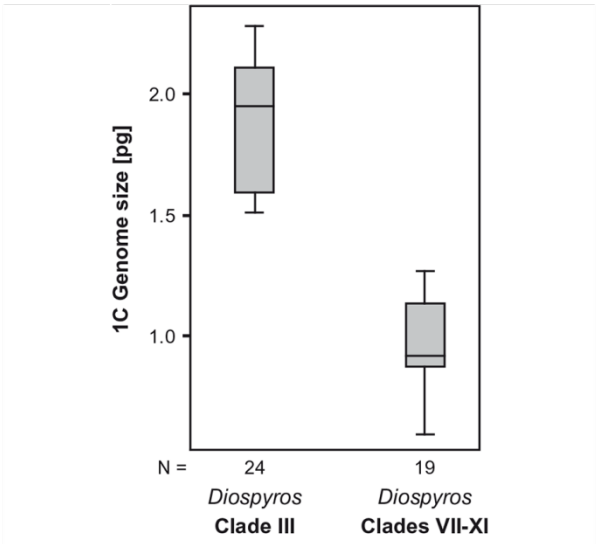


Fig. 8. Boxplot of genome size differences between taxa from clade III and those from clades VII–XI.

3.2. Low-copy nuclear markers

Nuclear markers contained proportionally more parsimony informative characters (*ncpGS* 2.7%, *PHYA* 1.2%) than the plastid markers (0.7%), but variation was still low. Some species form groups (Fig. 5), but they lack bootstrap and Bayesian posterior probability support. Among the three methods of analysis used for the *ncpGS* data set, Bayesian inference provides the best resolution (tree not shown), placing *D. vieillardii* (BBP 0.99) sister (BPP 1.00) to the rest of the NC clade. The relationship between *D. veillonii* and *D. cherrieri* (BPP 0.84) is weakly supported, but their position as subsequent sister of the rest of this clade is well supported (BPP 0.99). All individuals of *D. umbrosa* form a group with two individuals of *D. trisulca* (BBP 0.91). This set of accessions is subsequently sister (BBP 0.95) to the rest of the clade, within which there is no resolution. In the *PHYA* tree, there are only a few clades with strong support regardless of method of analysis. Clade III (BMP 100, BML 100, BPP 1.00) as monophyletic unit is confirmed, as well as the monophyly of NC clade III within it (BMP 77, BML 78, BPP 1.00). All included individuals of *D. cherrieri* fall together (BMP 84, BML 81, BPP 1.00) in the *PHYA* analyses. Only a single copy of *ncpGS* was recovered from all accessions investigated, as well as from most of the accessions of *PHYA*. Species from which two copies of *PHYA* were obtained when cloned are found in clades IX, X and XI (Fig. 4). The paralogous (divergent) copies of *PHYA* were easily detected and excluded from the phylogenetic analyses.

3.3. Combined data set

The ILD test found the trees of the plastid and low-copy nuclear markers to be congruent with *p*-values of 0.01, which indicates that combined analysis was appropriate. In trees inferred from the combined data set (Fig. 6), species of clade III were highly supported (BMP 100, BML 100, BPP 1.00); they include the species of NC clade III, Indian Ocean islands, Thailand and Hawai'i. *Diospyros vera* is sister to *D. sandwicensis* (BMP 100, BML 100, BPP 1.00) and then the NC clade III. NC clade III is moderately to well supported (BMP 83, BML 96, BPP 0.96). The position of *D. vieillardii* (BMP 100, BML 99, BPP 1.00) as sister to the rest of the clade is strongly supported (BMP 92, BML 98, BPP 1.00). All accessions of each of the two species, *D. umbrosa* (BMP < 70, BML 75, BPP 1.00) and *D. flavocarpa* (BMP < 70, BML < 70, BPP 0.99), form unique groups, which together are sister (BMP 100, BML 100, BPP 1.00) to the rest of the group. A sister relationship between *D. cherrieri* (BMP 96, BML 99, BPP 1.00) and *D. veillonii* (BMP 78, BML 86, BPP 1.00) is supported (BMP 75, BML 88, BPP 1.00). A clade comprising *D. calciphila*, *D. inexplorata* (both on coralline substrates) and *D.* sp. Pic N'ga (ultramafic substrate) is well supported (BMP 97, BML 99, BPP 1.00).

3.4. Dating analysis

We performed two dating analyses. The first one was based on a joint matrix of our plastid sequences together with the data set of Bell et al. (2010), which included many families across the whole Ericales with Cornales as outgroup. This dating analysis was used to get age estimates for the crown node of Ebenaceae, the two subfamilies Ebenoideae and Lissocarpoideae, the split of the three genera of Ebenoideae (*Diospyros* versus *Euclea*/*Royena*) and the main clades of *Diospyros*. The second dating analysis was based on our combined data set, which was used to infer ages of clades and species within *Diospyros*. The dating analysis of the over-arching matrix of plastid markers (Fig. S4, Supplementary material) indicates that the two subfamilies of Ebenaceae, Lissocarpoideae and Ebenoideae, diverged around 54 mya (42–65; 95% highest posterior density interval). The split of *Diospyros* from its sister genera, *Euclea* plus *Royena* occurred around 42 mya (35–50). The following conclusions are based on the dating analysis of the combined data set (Fig. 7). The Australian clade of *Diospyros* (clade II, Fig. 4), including five species from New Caledonia, separated from the rest of the genus around 34 mya (33–36), the New Caledonian and Australian members of this clade diverged around 20 mya (11–29). Divergence among the New Caledonian members began only about 6 mya (3–10). The two large main groups (clades V–XI and clade III, Fig. 4) diverged about 32 mya (25–35). The last common ancestor of *D. fasciculosa* and *D. olen* existed around 15 mya (11–19). *Diospyros olen* is around 5 myr (3–9) old and *D. fasciculosa* about 6 myr (3–10). Lineages of clade III started to diversify about 19 mya (13–21). Lineages forming NC clade III arrived in New Caledonia around 9 mya (6–13). *Diospyros vieillardii* is around 7 myr (5–10) old. The clade comprising *D. cherrieri* and *D. veillonii* is around 5 myr (3–8) old, and the two species separated around 3 mya (1–5). The clade including *D. flavocarpa*, *D. umbrosa* and one accession of *D. trisulca* is 5 myr (3–7) old. *Diospyros flavocarpa* is around 4 myr (3–6) old. The relationship between *D. umbrosa* and *D. trisulca* is not highly supported, but suggests an age of around 3 myr (2–5) for *D. umbrosa*. The group comprising *D. calciphila*, *D. inexplorata* and *D.* sp. Pic N'ga appears to be around 2 myr (1–3) old and started to diversify around 0.9 mya (0.5–2). Resolution between other species is too limited to say anything about their ages.

3.5. Chromosome counts and genome size

Chromosome counts made for *Diospyros fasciculosa*, *D. inconstans*, *D. macrocarpa*, *D. minimifolia*, *D. pentamera*, *D. pustulata*, *D. texana*, *D. veillonii* and *D. yatesiana* indicate that they are all diploid, $2n = 30$. The counts from the underlined species are here reported for the first time in literature. The other counts confirm results of White (1992).

Measurements of genome size showed differences among the New Caledonian species of *Diospyros*. *Diospyros olen* has with 1C = 0.86 pg, the smallest genome of the New Caledonian *Diospyros* species examined, followed by *D. fasciculosa* with 1C = 1.13 pg (both clade XI). The investigated species from the NC clade III have larger genomes (mean value 1C = 1.90 pg) than the two mentioned above (Table 6). We were not able to examine New Caledonian

Table 1
Table of accessions; showing all individuals used in this study. Sequences provided by S. Duangjai are indicated.

Taxon	Acc.-nr.	Origin	Voucher	Herbarium	*atpB*	*rbcL*	*matK* & *trnK* intron	*trnS–trnG*	*ncpGS*	*PHYA*
D. abyssinica (Hiern) F. White	K1672	Africa	Gilbert & Sebseke 8803	K	DQ923883	EU980646	DQ923990	EU981061		
D. affinis Thwaites	DY03	Sri Lanka	Yakandawala 03	PDA	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291310		
D. affinis	DY05	Sri Lanka	Yakandawala 05	PDA	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291311		
D. affinis	DY18	Sri Lanka	Yakandawala 18	PDA	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291312		
D. affinis	Eb179	Sri Lanka	Samuel s.n.	PDA	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291313		
D. affinis	Eb180	Sri Lanka	Samuel s.n.	PDA	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291314		
D. cf. *affinis*	S09	Sri Lanka	Samuel 09	PDA	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291315		
D. andamanica (Kurz) Bakh.	Eb002	Thailand	Duangjai 068	KUFF, W	DQ923884	EU980645	DQ923991	EU981060	KF291447	KF291624
D. andamanica	Eb104	Thailand	Duangjai 162	KUFF, W	DQ923950	EU980755	DQ924057	EU981170	KF291448	KF291625
D. anisandra S.F. Blake	W68	Guatemala	Wallnöfer 6012	W	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291316		
D. anisandra	W80	Guatemala	Frisch 2006-1	W	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291317		
D. apiculata Hiern	Eb006	Thailand	Duangjai 072	KUFF	EU980813	EU980647	EU980936	EU981062	KF291449	KF291626
D. areolata King & Gamble	Eb160	Brunei	Duangjai et al. 33	BRUN, W, WU	Duangjai unpubl	Duangjai unpubl	Duangjai unpubl	KF291318	KF291450	KF291627
D. artanthifolia Mart. ex Miq.	W15	Peru	Pirie 62	W	DQ923885	EU980648	DQ923992	EU981063		
D. australis (R.Br) Hiern	Eb205	Australia	Wallnöfer & Duangjai 13944	WU	DQ923887	EU980650	DQ923994	EU981065		
D. australis	K22548	Australia	Forster 7848	K	DQ923886	EU980649	DQ923993	EU981064		
D. balansae Guillaumin	M3556	New Caledonia	Munzinger 3556	NOU015466	EU980814	EU980651	EU980937	EU981066	KF291451	KF291628
D. batocana Hiern	K21210	Namibia	Steyl 88	K	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291319		
D. batocana	K22553	Zambia	Pope et al. 2196	K	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291320		
D. bejaudii Lecomte	Eb011	Thailand	Duangjai 075	KUFF, W	DQ923888	EU980652	DQ923995	EU981067	KF291452	KF291629
D. bipindensis Gürke	K22452	Gabon	Stone & Niangadouma 3554	MO	DQ923889	EU980653	DQ923996	EU981068		
D. borbonica I. Richardson	K23682	Reunion	Chase REU10042	REU, WU	EU980815	EU980654	EU980938	EU981069	KF291453	KF291630
D. borneensis Hiern	Eb015	Thailand	Duangjai 079	KUFF, W	DQ923890	EU980655	DQ923997	EU981070	KF291454	KF291631
D. bourdillonii Brandis	W82	India	DeFranceschi 18.12.2006	W	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291321		
D. brandisiana Kurz	Eb017	Thailand	Duangjai & Sinbumrung 007	KUFF, W	DQ923891	EU980656	DQ923998	EU981071	KF291455	KF291632
D. brassica F. White	M2898	New Caledonia	Munzinger 2898	NOU007949	DQ923892	EU980657	DQ923999	EU981072	KF291456	KF291633
D. buxifolia (Blume) Hiern	Eb018	Thailand	Duangjai 081	KUFF, W	EU980816	EU980658	EU980939	EU981073		
D. buxifolia	W85	India	DeFranceschi 18.12.2006	W	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291322		
D. calciphila F. White	BT314	New Caledonia	Munziner et al. 6650	MPU, NOU, P	KF291801	KF291860	KF291919	KF291323	KF291457	KF291634
D. calciphila	BT316	New Caledonia	Munziner et al. 6650	MPU, NOU, P	KF291802	KF291861	KF291920	KF291324	KF291458	KF291635
D. calciphila	BT317	New Caledonia	Munziner et al. 6653	MPU, NOU, P					KF291459	KF291636
D. calciphila	YP124	New Caledonia	Pillon 124	NOU006325					KF291460	KF291637
D. capreifolia Mart. ex Hiern	W09	French Guiana	Prévost & Sabatier 3476	W	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291325		
D. carbonaria Benoist	W10	French Guiana	Prévost & Sabatier 3470	W	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291326		
D. caribaea (A.DC.) Standl.	W65	Cuba	Abbott 19004	W	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291327		
D. castanea (Craib) Fletcher	Eb020	Thailand	Duangjai 083	KUFF, W	DQ923893	EU980660	DQ924000	EU981075		
D. cauliflora Blume	Eb024	Thailand	Duangjai 087	KUFF, W	DQ923894	EU980661	DQ924001	EU981076	KF291461	KF291638
D. cavalcantei Sothers	W22	French Guiana	Prévost et al. 4671	W	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291328		
D. cayennensis A.DC.	W03	French Guiana	Prévost 3430	W	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291329		
D. celebica Bakh.	K1242	Indonesia	Chase 1242	K	DQ923897	EU980664	DQ924004	EU981079		

(*continued on next page*)

Table 1 (*continued*)

Taxon	Acc.-nr.	Origin	Voucher	Herbarium	*atpB*	*rbcL*	*matK* & *trnK* intron	*trnS–trnG*	*ncpGS*	*PHYA*
D. cherrieri F. White	BT262	New Caledonia	Chambrey & Turner 16	NOU079551, WU062860	KF291803	KF291862	KF291921	KF291330	KF291463	KF291640
D. cherrieri	BT297	New Caledonia	Chambrey & Turner 17	NOU079547	KF291804	KF291863	KF291922	KF291331	KF291464	KF291641
D. cherrieri	VH3510	New Caledonia	Hequet 3510	NOU015245	EU980818	EU980665	EU980941	EU981080	KF291465	KF291642
D. cherrieri	VH3516	New Caledonia	Hequet 3516	NOU015251	EU980819	EU980666	EU980942	EU981081	KF291466	KF291643
D. cherrieri	VH3610	New Caledonia	Hequet 3610	NOU016962					KF291467	KF291644
D. cherrieri	VH3640	New Caledonia	Hequet 3640	NOU017014					KF291468	KF291645
D. chrysophyllos Poir.	K25758	Mauritius	Page 45	MAU	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291332		
D. chrysophyllos	K25769	Mauritius	Page 71	MAU	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291333		
D. clementium Bakh.	Eb154	Brunei	Duangjai et al. 24	BRUN, W, WU	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291334		
D. confertiflora (Hiern) Bakh.	Eb028	Thailand	Duangjai 091	KUFF, W	DQ923898	EU980667	DQ924005	EU981082		
D. consolatae Chiov.	K1673	Africa	Beentje 2168	K	DQ923899	EU980668	DQ924006	EU981083		
D. cooperi (Hutchinson & Dalziel) F. White	K20604	Ghana	Merello et al. 1350	MO	DQ923900	EU980669	DQ924007	EU981084		
D. crassinervis (Krug & Urb.) Standl.	W23	Cuba	Rainer s.n.	W	DQ923901	EU980670	DQ924008	EU981085		
D. curranii Merr.	Eb031	Thailand	Duangjai 094	KUFF, W, WU	DQ923902	EU980671	DQ924009	EU981086	KF291469	KF291646
D. dasyphylla Kurz	Eb033	Thailand	Duangjai 096	KUFF, W	DQ923903	EU980672	DQ924010	EU981087		
D. defectrix Fletcher	Eb097	Thailand	Duangjai 155	KUFF, WU	KF291805	KF291864	KF291923	KF291335	KF291470	KF291647
D. dendo Welw. ex Hiern	K21197	Central African Republic	Harris & Fay 1594	K	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291336		
D. dichroa Sandwith	W13	French Guiana	Sabatier et al. 4457	W	DQ923904	EU980673	DQ924011	EU981088		
D. dictyoneura Hiern	Eb038	Thailand	Duangjai 100	KUFF, W	EU980674	EU980820	EU980943	EU981089	KF291471	KF291648
D. diepenhorstii Miq.	Eb042	Thailand	Duangjai 103	KUFF, W	DQ923905	EU980675	DQ924012	EU981090	KF291472	KF291649
D. discolor Willd.	Eb088	Thailand	Duangjai 146	KUFF, WU	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291337	KF291473	KF291650
D. ebenum J. Koenig ex Retz	DY06	Sri Lanka	Yakandawala 06	PDA	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291338		
D. ebenum	DY08	Sri Lanka	Yakandawala 08	PDA	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291339		
D. ebenum	Eb174	Sri Lanka	Samuel s.n.	WU	EU980677	EU980821	EU980944	EU981092		
D. ebenum	W83	India	Ramesh Diosass-2	W	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291340		
D. ebenum	W84	India	DeFranceschi 21.12.2006	W	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291341		
D. egrettarum I. Richardson	K25788	Mauritius	Page 122	MAU	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291342		
D. ehretioides Wall. ex G. Don	Eb043	Thailand	Duangjai 104	KUFF, W	DQ923907	EU980678	DQ924014	EU981093	KF291474	KF291651
D. eriantha Charmp. ex Benth	W63	Taiwan	Chung & Anderberg 1401	HAST	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291343		
D. erudita F. White	BT287	New Caledonia	Chambrey & Turner 20	NOU	KF291806	KF291865	KF291924	KF291344	KF291475	KF291652
D. erudita	M2359	New Caledonia	Munzinger et al. 2359	NOU003840	EU980845	EU980739	EU980968	EU981154	KF291476	KF291653
D. erudita/*pustulata*	M3010	New Caledonia	Munziner et al. 3010	NOU008358	EU980841	EU980735	EU980964	EU981150		
D. fasciculosa (F. Muell.) F. Muell.	BT014	New Caledonia	Munzinger et al. 6617	NOU					KF291477	KF291654
D. fasciculosa	BT142	New Caledonia	MacKee 27341	NOU022840					KF291478	KF291655
D. fasciculosa	BT165	New Caledonia							KF291479	KF291656
D. fasciculosa	BT166	New Caledonia							KF291480	KF291657
D. fasciculosa	BT335	New Caledonia							KF291481	KF291658
D. fasciculosa	M2127	New Caledonia	Munzinger 2127	NOU003604	DQ923908	EU980679	DQ924015	EU981094	KF291482	KF291659
D. fasciculosa	YP243	New Caledonia	Pillon et al. 243	NOU010096	EU980822	EU980680	EU980945	EU981095	KF291483	KF291660
D. ferox Bakh.	Eb146	Brunei	Duangjai et al. 012	BRUN, W, WU	DQ923909	EU980681	DQ924016	EU981096	KF291484	KF291661

(*continued on next page*)

Table 1 (*continued*)

D. ferruginescens Bakh.	Eb143	Brunei	Duangjai et al. 007	BRUN, W, WU	DQ923911	EU980685	DQ924018	EU981100		
D. filipendula Pierre ex Lecomte	Eb048	Thailand	Duangjai 109	KUFF	DQ923912	EU980686	DQ924019	EU981101	KF291485	KF291662
D. flavocarpa (Vieill. ex P. Parm.) F. White	BT126	New Caledonia	Munzinger et al. 6625	NOU	KF291807	KF291866	KF291925	KF291345	KF291486	KF291663
D. flavocarpa	BT127	New Caledonia	Munzinger et al. 6625	NOU	KF291808	KF291867	KF291926	KF291346	KF291487	KF291664
D. flavocarpa	BT156	New Caledonia	Munzinger et al. 6632	NOU					KF291488	KF291665
D. flavocarpa	K20607	New Caledonia	McPherson & Lowry 18563	NOU022877	DQ923913	EU980687	DQ924020	EU981102	KF291489	KF291666
D. flavocarpa	K20614	New Caledonia	Lowry et al. 5783	NOU023319	EU980870	EU980782	EU980993	EU981197		
D. flavocarpa	M2235	New Caledonia	Munzinger 2235	NOU006659	EU980825	EU980688	EU980948	EU981103	KF291490	KF291667
D. flavocarpa	M2905	New Caledonia	Munzinger et al. 2905	NOU007977	EU980826	EU980689	EU980949	EU981104		
D. fragrans Gürke	K22454	Gabon	SIMAB 010610	MO	DQ923914	EU980690	DQ924021	EU981105		
D. frutescens Blume	Eb049	Thailand	Duangjai 110	KUFF, W	EU980827	EU980691	EU980950	EU981106		
D. fulvopilosa Fletcher	Eb052	Thailand	Duangjai 113	KUFF, W	DQ923915	EU980692	DQ924022	EU981107	KF291491	KF291668
D. fuscovelutina Baker	RF938	Madagascar	RF 938	W	DQ923979	EU980803	DQ924088	EU981218		
D. gabunensis Gürke	K22560	Tanzania	Bidgood et al. 2890	K	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291347		
D. gilletii De Wild	K21198	Cameroon	Harris & Fay 884	K	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291348		
D. glandulosa Lace	Eb053	Thailand	Duangjai 114	KUFF, W	DQ923916	EU980693	DQ924023	EU981108	KF291492	KF291669
D. glans F. White	BT019	New Caledonia			KF291809	KF291868	KF291927	KF291349		
D. glans	BT093	New Caledonia	Turner et al. 093	MPU	KF291810	KF291869	KF291928	KF291350	KF291493	KF291670
D. glans	BT094	New Caledonia	Turner et al. 094	MPU	KF291811	KF291870	KF291929	KF291351	KF291494	KF291671
D. glaucifolia Metcalf	K14256	China	Chase 14256	K	DQ923917	EU980694	DQ924024	EU981109		
D. cf. *gracilipes* Hiern	RF978	Madagascar	RNF 978	W	DQ923918	EU980695	DQ924025	EU981110		
D. gracilis Fletcher	Eb058	Thailand	Duangjai 019	BK, BKF, KUFF, WU	KF291812	KF291871	KF291930	KF291352	KF291495	KF291672
D. greenweyi F. White	K21205	Somalia	Friis et al. 4991	K	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291353		
D. grisebachii (Heirn) Standl.	W64	Cuba	Abbott 18937	W	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291354		
D. guianensis (Aubl.) Gürke	W14	French Guiana	Prévost & Sabatier 4029	W	DQ923919	EU980696	DQ924026	EU981111		
D. guianensis	W78	French Guiana	Mori 25921	NY, W	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291355		
D. hartmaniana S. Knapp	K22455	Panama	McPherson & Richardson 15959	MO	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291356		
D. impolita F. White	BT102	New Caledonia	Schmid 5010	NOU019538	KF291813	KF291872	KF291931	KF291357	KF291496	KF291673
D. impolita	BT105	New Caledonia	Schmid 5010	NOU019538	KF291814	KF291873	KF291932	KF291358	KF291497	KF291674
D. inconstans Jacq.	W79	Ecuador	Rainer 1682	W	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291359		
D. inexplorata F. White	BT304	New Caledonia	MacKee 22791	NOU005818	KF291815	KF291874	KF291933	KF291360	KF291498	KF291675
D. inexplorata	BT311	New Caledonia	MacKee 22791	NOU005818	KF291816	KF291875	KF291934	KF291361	KF291499	KF291676
D. insidiosa Bakh.	Eb061	Thailand	Duangjai 120	KUFF, W	DQ923920	EU980697	DQ924027	EU981112		
D. iturensis (Gürke) Letouzey & F. White	K21204	Cameroon	Harris & Fay 1513	K	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291362		
D. kaki L.f.	K920	Japan	Chase 920	K	DQ923921	EU980698	DQ924028	EU981113	KF291500	KF291677
D. kirkii Hiern	K22551	Zimbabwe	Poilecot 7650	K	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291363		
D. kupensis Gosline	AR62	Cameroon	Russell 62	K	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.		KF291501	KF291678
D. labillardierei F. White	BT121	New Caledonia	Munzinger et al. 6624	NOU	KF291817	KF291876	KF291935	KF291364	KF291502	KF291679
D. labillardierei	BT122	New Caledonia	Munzinger et al. 6624	NOU	KF291818	KF291877	KF291936	KF291365	KF291503	KF291680
D. labillardierei	BT179	New Caledonia							KF291504	KF291681
D. labillardierei	K20763	New Caledonia	McPherson & Munzinger 18038	MO	DQ923922	EU980699	DQ924029	EU981114		
D. labillardierei	M2219	New Caledonia	Munzinger 2219	NOU006657	EU980828	EU980700	EU980951	EU981115	KF291505	KF291682
D. labillardierei	M3053	New Caledonia	Munzinger 3053	NOU008407	EU980829	EU980701	EU980952	EU981116	KF291506	KF291683

(*continued on next page*)

Table 1 (*continued*)

Taxon	Acc.-nr.	Origin	Voucher	Herbarium	*atpB*	*rbcL*	*matK* & *trnK* intron	*trnS–trnG*	*ncpGS*	*PHYA*
D. lanceifolia Roxb.	K1245	Indonesia	Chase 1245	K	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291366		
D. leucomelas Poir.	K25752	Mauritius	Page 16	MAU	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291367		
D. lotus L.	D16	Living coll. HBV	Turner D16	Living coll. HBV					KF291507	KF291684
D. lotus	K965	Living coll. Kew 1882-3501	Chase 965	K	DQ923924	EU980703	DQ924031	EU981118		
D. macrocarpa (Vieill.) Hiern	BT043	New Caledonia							KF291508	KF291685
D. macrocarpa	BT044	New Caledonia							KF291509	KF291686
D. macrocarpa	BT048	New Caledonia							KF291510	KF291687
D. macrocarpa	BT049	New Caledonia							KF291511	KF291688
D. macrocarpa	BT050	New Caledonia							KF291512	KF291689
D. macrocarpa	M2014	New Caledonia	Munzinger 2014	NOU003637	EU980830	EU980704	EU980953	EU981119		
D. macrocarpa	M2829	New Caledonia	Munzinger 2829	NOU008233	DQ923925	EU980705	DQ924032	EU981120		
D. maingayi (Hiern) Bakh.	Eb073	Thailand	Duangjai 131	KUFF, W	DQ923926	EU980706	DQ924033	EU981121		
D. malabarica (Desr.) Kostel.	Eb066	Thailand	Duangjai 006	KUFF, W	EU980708	DQ923928	DQ924035	EU981123		
D. malabarica	K1247	Indonesia	Chase 1247	K	DQ923927	EU980707	DQ924034	EU981122		
D. malabarica	W47	South East Asia, cult. USA	Abbott 14325	W	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291368		
D. mannii Hiern	K20597	Ghana	Merello et al. 1348	MO	DQ923929	EU980709	DQ924036	EU981124		
D. margaretae F. White	YP1267	New Caledonia	Pillon 1267	NOU049432, WU062863	KF291819	KF291878	KF291937	KF291369	KF291513	KF291690
D. maritima Blume	Eb209	Malaysia	Wallnöfer 13948	W	DQ923930	EU980710	DQ924037	EU981125		
D. melanida Poir.	K25786	Mauritius	Page 112	MAU	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291370		
D. melocarpa F. White	K22457	Gabon	SIMAB 012319	MO	DQ923931	EU980711	DQ924038	EU981126		
D. mespiliformis Hochst. Ex A.DC.	Eb206	Tropical Africa	Wallnöfer & Duangjai 13945	W	DQ923932	EU980712	DQ924039	EU981127	KF291514	KF291691
D. mespiliformis	W60	Senegal	Prinz 2005-5	W	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291371		
D. minimifolia F. White	BT131	New Caledonia	Dagostini 203	NOU019556	KF291820	KF291879	KF291938	KF291372	KF291515	KF291692
D. minimifolia	BT133	New Caledonia	Dagostini 203	NOU019556	KF291821	KF291880	KF291939	KF291373	KF291516	KF291693
D. minimifolia	BT231	New Caledonia	Veillon 7206	NOU019554					KF291517	KF291694
D. minimifolia	BT264	New Caledonia	Chambrey & Turner 24	NOU079549, WU062872					KF291518	KF291695
D. minimifolia	M2214	New Caledonia	Munzinger 2214	NOU006263	EU980831	EU980714	EU980954	EU981129	KF291519	KF291696
D. minimifolia	M2374	New Caledonia	Munzinger 2374	NOU006677	EU980832	EU980715	EU980955	EU981130	KF291520	KF291697
D. minimifolia/pustulata	BT143	New Caledonia							KF291521	KF291698
D. mollis Griff.	Eb074	Thailand	Duangjai 132	KUFF, W	DQ923934	EU980716	DQ924041	EU981131	KF291522	KF291699
D. montana Roxb.	Eb078	Thailand	Duangjai 136	KUFF, W	DQ923935	EU980717	DQ924042	EU981132		
D. montana	Eb130	Thailand	Duangjai & Sinbumrung 017	KUFF, W	DQ923943	EU980733	DQ924050	EU981148		
D. myriophylla (H. Perrier) G.E. Schatz & Lowry	W34	Madagascar	Sieder 209	W	DQ923974	EU980797	DQ924083	EU981212		
D. natalensis (Harv.) Brenan	K22554	Zambia	Bingham 10635	K	DQ923936	EU980718	DQ924043	EU981133		
D. nigra (J.F. Gmel.) Perrier	K212	Cult. Mexico	Chase 212	NCU	DQ923906	EU980676	DQ924013	EU981091		
D. nigra	K1146	Cult. Mexico	Chase 1146	K	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291374		

(*continued on next page*)

Table 1 (*continued*)

D. obliquifolia (Hiern ex Gürke) F. White	W91	Cameroon	Rainer 6.3.2007	W	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291375		
D. oblonga Wall. Ex G. Don.	Eb083	Thailand	Duangjai 141	KUFF, W	DQ923937	EU980719	DQ924044	EU981134		
D. olen Hiern	BT001	New Caledonia	Munzinger et al. 6609	NOU	KF291822	KF291881	KF291940	KF291376	KF291523	KF291700
D. olen	BT032	New Caledonia							KF291524	KF291701
D. olen	BT034	New Caledonia							KF291525	KF291702
D. olen	BT169	New Caledonia	Munzinger et al. 6634	NOU					KF291526	KF291703
D. olen	BT302	New Caledonia							KF291527	KF291704
D. olen	K20598	New Caledonia	Lowry et al. 5628	MO, NOU004840	DQ923938	EU980720	DQ924045	EU981135		
D. olen	M2827	New Caledonia	Munzinger 2827	NOU008235	EU980833	EU980721	EU980956	EU981136		
D. olen	YP153	New Caledonia	Pillon 153	NOU006438	EU980834	EU980722	EU980957	EU981137		
D. oubatchensis Kosterm.	BT160	New Caledonia	LeCore et al. 768	NOU079472	KF291823	KF291882	KF291941	KF291377	KF291528	KF291705
D. oubatchensis	BT161	New Caledonia	LeCore et al. 768	NOU079472	KF291824	KF291883	KF291942	KF291378	KF291529	KF291706
D. oubatchensis	M3118	New Caledonia	Munzinger 3118	NOU009675	EU980835	EU980723	EU980958	EU981138		
D. oubatchensis	M3333	New Caledonia	Munzinger 3333	NOU011201	EU980836	EU980724	EU980959	EU981139		
D. ovalifolia Wight	DY10	Sri Lanka	Yakandawala 10	PDA	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291379		
D. pancheri Kosterm.	BT027	New Caledonia	Munzinger et al. 6619	NOU	KF291825	KF291884	KF291943	KF291380	KF291530	KF291707
D. pancheri	BT028	New Caledonia	Munzinger et al. 6619	NOU	KF291826	KF291885	KF291944	KF291381	KF291531	KF291708
D. pancheri	BT029	New Caledonia	Munzinger et al. 6619	NOU					KF291532	KF291709
D. pancheri	BT030	New Caledonia	Munzinger et al. 6620	NOU					KF291533	KF291710
D. pancheri	BT031	New Caledonia	Munzinger et al. 6620	NOU					KF291534	KF291711
D. pancheri	BT033	New Caledonia	Munzinger et al. 6620	NOU	KF291827	KF291886	KF291945	KF291382	KF291535	KF291712
D. pancheri	BT035	New Caledonia	Munzinger et al. 6620	NOU					KF291536	KF291713
D. pancheri	BT076	New Caledonia							KF291537	KF291714
D. pancheri	M2138	New Caledonia	Munzinger 2138	NOU003868	EU980837	EU980725	EU980960	EU981140	KF291538	KF291715
D. pancheri/*parviflora*	M2338	New Caledonia	Munzinger 2338	NOU006586	EU980838	EU980726	EU980961	EU981141	KF291539	KF291716
D. parviflora (Schltr.) Bakh.	BT038	New Caledonia			KF291828	KF291887	KF291946	KF291383		
D. parviflora	BT039	New Caledonia			KF291829	KF291888	KF291947	KF291384	KF291540	KF291717
D. parviflora	BT040	New Caledonia							KF291541	KF291718
D. parviflora	BT042	New Caledonia							KF291542	KF291719
D. parviflora	BT187	New Caledonia	Munzinger et al. 6636	NOU					KF291543	KF291720
D. parviflora	M2037	New Caledonia	Munzinger 2037	NOU002519	EU980839	EU980727	EU980962	EU981142	KF291544	KF291721
D. parviflora	M2071	New Caledonia	Munzinger 2071	NOU002608	EU980869	EU980776	EU980992	EU981191	KF291545	KF291722
D. parviflora	M2708	New Caledonia	Munzinger 2708	NOU006658	EU980728	EU980840	EU980963	EU981143		
D. parviflora	M3035	New Caledonia	Munzinger 3035	NOU008397	EU980842	EU980736	EU980965	EU981151		
D. pentamera (Woolls & F. Muell.) F. Muell.	K22549	Australia	Forster & Booth 25525	K	DQ923939	EU980729	DQ924046	EU981144		
D. perplexa F. White	BT004	New Caledonia	Munzinger et al. 6611	NOU	KF291830	KF291889	KF291948	KF291385	KF291546	KF291723
D. perplexa	BT005	New Caledonia	Munzinger et al. 6611	NOU	KF291831	KF291890	KF291949	KF291386	KF291547	KF291724
D. perplexa	BT009	New Caledonia	Munzinger et al. 6611	NOU	KF291832	KF291891	KF291950	KF291387	KF291548	KF291725

(*continued on next page*)

Table 1 (*continued*)

Taxon	Acc.-nr.	Origin	Voucher	Herbarium	*atpB*	*rbcL*	*matK* & *trnK* intron	*trnS–trnG*	*ncpGS*	*PHYA*
D. perplexa	BT147	New Caledonia	Munzinger et al. 6630	NOU					KF291549	KF291726
D. perplexa	BT148	New Caledonia	Munzinger et al. 6630	NOU					KF291550	KF291727
D. perplexa	VH3614	New Caledonia	Hequet et al. 3614	NOU016957	EU980873	EU980786	EU980996	EU981201	KF291551	KF291728
D. philippinensis A.DC.	K1248	Indonesia	Chase 1248	K	DQ923940	EU980730	DQ924047	EU981145		
D. philippinensis	W62	Taiwan	Chung & Anderberg 1400	HAST	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291388		
D. pilosanthera Blanco	Eb091	Thailand	Duangjai 149	KUFF, W	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291389		
D. pilosiuscula G. Don	Eb092	Thailand	Duangjai 150	KUFF, W	DQ923941	EU980731	DQ924048	EU981146		
D. preussii Gürke	LPJMO39	Cameroon	LPJMO39	YA	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291390		
D. pruriens Dalzell	W81	India	DeFranceschi 18.12.2006	W	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291391		
D. pseudomespilus Mildbr.	K20606	Gabon	Walters et al. 956	MO	DQ923942	EU980732	DQ924049	EU981147		
D. puncticulosa Bakh.	Eb150	Brunei	Duangjai et al. 018	BRUN, W, WU	DQ923944	EU980734	DQ924051	EU981149		
D. pustulata F. White	BT113	New Caledonia			KF291833	KF291892	KF291951	KF291392	KF291552	KF291729
D. pustulata	BT114	New Caledonia			KF291834	KF291893	KF291952	KF291393	KF291553	KF291730
D. pustulata	BT136	New Caledonia	Munzinger et al. 6629	NOU					KF291554	KF291731
D. pustulata	BT137	New Caledonia	Munzinger et al. 6629	NOU					KF291555	KF291732
D. pustulata	BT257	New Caledonia	Cambrey & Turner 21	NOU079548, WU062871					KF291556	KF291733
D. pustulata	M3580	New Caledonia	Munzinger 3580	NOU016720	EU980843	EU980737	EU980966	EU981152	KF291557	KF291734
D. pustulata	M3584	New Caledonia	Munzinger 3584	NOU016734	EU980844	EU980738	EU980967	EU981153	KF291558	KF291735
D. pustulata	VH3638	New Caledonia	Hequet et al. 3638	NOU017016					KF291559	KF291736
D. pustulata/ *yahouensis*	BT259	New Caledonia	Chambrey & Turner 26	WU062855	KF291835	KF291894	KF291953	KF291394	KF291560	KF291737
D. racemosa Roxb.	Eb106	Thailand	Duangjai 164	KUFF	EU980856	EU980759	EU980979	EU981174	KF291561	KF291738
D. revaughanii I. Richardson	K25760	Mauritius	Page 47	MAU	Duangjai unpubl	Duangjai unpubl	Duangjai unpubl	KF291395		
D. revolutissima F. White	BT116	New Caledonia	MacKee 22382	NOU023189	KF291836	KF291895	KF291954	KF291396	KF291562	KF291739
D. revolutissima	BT117	New Caledonia	MacKee 22382	NOU023189	KF291837	KF291896	KF291955	KF291397	KF291563	KF291740
D. revolutissima	BT218	New Caledonia	Munzinger et al. 6640	NOU					KF291564	KF291741
D. revolutissima	BT219	New Caledonia	Munzinger et al. 6640	NOU					KF291565	KF291742
D. revolutissima	YP204	New Caledonia	Pillon 204	NOU009155	EU980846	EU980740	EU980969	EU981155	KF291566	KF291743
D. rhodocalyx Kurz	Eb096	Thailand	Duangjai 154	KUFF, WU	KF291838	KF291897	KF291956	KF291398	KF291567	KF291744
D. rhombifolia Hemsl.	Eb129	Thailand	Duangjai & Sinbumrung 016	KUFF, W	DQ923945	EU980741	DQ924052	EU981156		
D. cf. *rhombifolia*	W76	Cult. USA, (South East Asia)	Abbott 20824	W	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291399		
D. ridleyi Bakh.	Eb138	Brunei	Duangjai et al. 002	BRUN, W, WU	DQ923946	EU980742	DQ924053	EU981157	KF291568	KF291745
D. rigida Hiern	Eb140	Brunei	Duangjai et al. 004	BRUN, W, WU	DQ923947	EU980743	DQ924054	EU981158		
D. ropourea B. Walln.	W20	French Guiana	Wallnöfer 13459	W	DQ923948	EU980744	DQ924055	EU981159		
D. salicifolia Humb. & Bonpl. ex Willd.	W66	Guatemala	Abbott 19765	W	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291400		
D. salicifolia	W67	Guatemala	Abbott 19777	W	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291401		
D. samoensis A. Gray	Eb176	Cult. Hawaii Bot Garden	Kiehn s.n.	WU	EU980745	EU980847	EU980970	EU981160		
D. samoensis	M3593	Vanuatu	Munzinger 3593	NOU080070	EU980848	EU980746	EU980971	EU981161		
D. samoensis	M3624	Vanuatu	Munzinger 3624	NOU080138, NOU080139	EU980849	EU980747	EU980972	EU981162	KF291569	KF291746
D. samoensis	M3691	Vanuatu	Munzinger 3691	NOU	EU980850	EU980748	EU980973	EU981163		

(*continued on next page*)

Table 1 (*continued*)

D. sandwicensis (A.DC.) Fosberg	Eb175	Cult. Hawaii Bot Garden	Kiehn s.n.	WU	EU980851	EU980749	EU980974	EU981164	KF291570	KF291747
D. scabra (Chiov.) Cufod.	K21206	Ethiopia	Wondefrash & Tefera 9622	K	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291402		
D. scalariformis Fletcher	Eb172	Thailand	Duangjai & Sinbumrung s.n.	KUFF, W	EU980750	EU980852	EU980975	EU981165		
D. senensis Klotzsch	K22552	Zambia	Bingham 11092	K	EU980853	EU980751	EU980976	EU981166		
D. squarrosa Klotzsch	K21207	Somalia	Friis et al. 4894	K	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291417		
D. squarrosa	K22555	Zambia	Bingham & Downie 11465	K	EU980854	EU980752	EU980977	EU981167		
D. styraciformis King & Gamble	Eb149	Brunei	Duangjai et al. 017	BRUN, W, WU	DQ923949	EU980753	DQ924056	EU981168		
D. sumatrana Miq.	Eb099	Thailand	Duangjai 157	KUFF, W	EU980855	EU980754	EU980978	EU981169		
D. tenuiflora A.C.Sm.	W32	Brazil	Maas et al. 9186	NY, W	DQ923923	EU980702	DQ924030	EU981117		
D. tesselaria Poir.	K25751	Mauritius	Page 15	MAU	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291418		
D. tetrandra Hiern	W31	French Guiana	Prévost & Sabatier 4713	W	DQ923951	EU980756	DQ924058	EU981171		
D. tetrasperma Sw.	K14254	Mexico	Chase 14254	K, W	DQ923952	EU980757	DQ924059	EU981172		
D. texana Scheele	Eb208	Middle America	Wallnöfer & Duangjai 13946	W	DQ923953	EU980758	DQ924060	EU981173	KF291575	KF291752
D. tireliae F. White	M5725	New Caledonia	Munzinger 5725	NOU051026	KF291843	KF291902	KF291961	KF291419	KF291576	KF291753
D. tridentata F. White	BT202	New Caledonia	Munzinger et al. 6639	NOU	KF291844	KF291903	KF291962	KF291420	KF291577	KF291754
D. tridentata	BT203	New Caledonia	Munzinger et al. 6639	NOU	KF291845	KF291904	KF291963	KF291421	KF291578	KF291755
D. trisulca F. White	BT185	New Caledonia	Hequet (leg. Butin) 3820	NOU031344	KF291846	KF291905	KF291964	KF291422	KF291579	KF291756
D. trisulca	BT189	New Caledonia	Hequet (leg. Butin) 3820	NOU031344	KF291847	KF291906	KF291965	KF291423	KF291580	KF291757
D. trisulca	BT192	New Caledonia	Hequet (leg. Butin) 3820	NOU031344	KF291848	KF291907	KF291966	KF291424	KF291581	KF291758
D. trisulca	BT197	New Caledonia	Munzinger et al. 6637	NOU					KF291582	KF291759
D. trisulca	M3179	New Caledonia	Munzinger 3179	NOU016896	EU980871	EU980784	EU980994	EU981199		
D. trisulca	M3260	New Caledonia	Munzinger 3260	NOU016891, WU062868	EU980872	EU980785	EU980995	EU981200	KF291583	KF291760
D. cf. *ulo* Merr.	Eb152	Brunei	Duangjai et al. 021	BRUN, W, WU	EU980857	EU980760	EU980980	EU981175	KF291462	KF291639
D. umbrosa F. White	BT065	New Caledonia			KF291849	KF291908	KF291967	KF291425	KF291584	KF291761
D. umbrosa	BT066	New Caledonia							KF291585	KF291762
D. umbrosa	BT071	New Caledonia							KF291586	KF291763
D. umbrosa	BT246	New Caledonia	McPherson 2144	NOU023234	KF291850	KF291909	KF291968	KF291426	KF291587	KF291764
D. umbrosa	BT247	New Caledonia	McPherson 2144	NOU023234	KF291851	KF291910	KF291969	KF291427	KF291588	KF291765
D. umbrosa	BT256	New Caledonia	McPherson 2144	NOU023234					KF291589	KF291766
D. umbrosa	M2265	New Caledonia	Munzinger 2265	NOU006679	EU980858	EU980761	EU980981	EU981176	KF291590	KF291767
D. umbrosa	M2636	New Caledonia	Munzinger 2636	NOU006678	EU980859	EU980762	EU980982	EU981177	KF291591	KF291768
D. umbrosa	M2771	New Caledonia	Munzinger 2771	NOU007912	EU980860	EU980763	EU980983	EU981178	KF291592	KF291769
D. undulata Wall. Ex G. Don	Eb112	Thailand	Duangjai 170	KUFF, W	DQ923954	EU980764	DQ924061	EU981179		
D. veillonii F. White	BT224	New Caledonia	Veillon 7919	NOU019582	KF291852	KF291911	KF291970	KF291428	KF291593	KF291770
D. veillonii	BT229	New Caledonia	Veillon 7919	NOU019582	KF291853	KF291912	KF291971	KF291429	KF291594	KF291771
D. veillonii	M.sn.	New Caledonia	Munzinger s.n.	Living coll. Hortus Veillonii	EU980861	EU980765	EU980984	EU981180	KF291595	KF291772
D. venosa Wall ex A.DC	Eb119	Thailand	Duangjai 177	KUFF, W	DQ923955	EU980767	DQ924062	EU981182	KF291596	KF291773
D. venosa	Eb131	Thailand	Duangjai 059	KUFF, W	EU980862	EU980766	EU980985	EU981181		
D. vera (Lour.) A. Chev.	DY16	Sri Lanka	Yakandawala 16	PDA	EU980823	EU980682	EU980946	EU981097		
D. vera	Eb045	Thailand	Duangjai 106	KUFF	DQ923910	EU980683	DQ924017	EU981098	KF291597	KF291774

(*continued on next page*)

Table 1 (*continued*)

Taxon	Acc.-nr.	Origin	Voucher	Herbarium	*atpB*	*rbcL*	*matK* & *trnK* intron	*trnS–trnG*	*ncpGS*	*PHYA*
D. vera	K21193	Central African Republic	Harris & Fay 2032	K	EU980824	EU980684	EU980947	EU981099		
D. vestita Benoist	W01	French Guiana	Molino 1849	W	DQ923956	EU980768	DQ924063	EU981183		
D. vieillardii (Hiern) Kosterm.	BT025	New Caledonia	Munzinger et al. 6618	NOU	KF291854	KF291913	KF291972	KF291430	KF291598	KF291775
D. vieillardii	BT026	New Caledonia	Munzinger et al. 6618	NOU	KF291855	KF291914	KF291973	KF291431	KF291599	KF291776
D. vieillardii	BT055	New Caledonia							KF291600	KF291777
D. vieillardii	BT057	New Caledonia							KF291601	KF291778
D. vieillardii	BT099	New Caledonia							KF291602	KF291779
D. vieillardii	BT100	New Caledonia							KF291603	KF291780
D. vieillardii	BT213	New Caledonia	MacKee 25141	NOU023242					KF291604	KF291781
D. vieillardii	BT214	New Caledonia	MacKee 25141	NOU023242					KF291605	KF291782
D. vieillardii	BT286	New Caledonia	Chambrey & Turner 13	NOU054004, WU062859					KF291606	KF291783
D. vieillardii	BT325	New Caledonia	Munzinger et al. 6657	NOU, P					KF291607	KF291784
D. vieillardii	M2106	New Caledonia	Munzinger 2106	NOU006676	EU980863	EU980769	EU980986	EU981184	KF291608	KF291785
D. vieillardii	M2776	New Caledonia	Munzinger 2776	NOU008207	EU980864	EU980770	EU980987	EU981185		
D. vieillardii	M3476	New Caledonia	Munzinger 3476	NOU012947					KF291609	KF291786
D. vieillardii	M3572	New Caledonia	Munzinger 3572	NOU016733	EU980866	EU980772	EU980989	EU981187	KF291610	KF291787
D. vieillardii	YP146	New Caledonia	Pillon 146	NOU006400	EU980867	EU980773	EU980990	EU981052	KF291611	KF291788
D. virginiana L.	K14255	USA	Chase 14255	K	DQ923957	EU980774	DQ924064	EU981189	KF291612	KF291789
D. wallichii King & Gamble ex King	Eb122	Thailand	Duangjai 180	KUFF, W	EU980868	EU980775	EU980991	EU981190	KF291613	KF291790
D. wallichii	Eb165	Brunei	Duangjai et al. 41	BRUN, W, WU					KF291614	KF291791
D. winitii Fletcher	Eb123	Thailand	Duangjai 181	KUFF, WU					KF291615	KF291792
D. yahouensis (Schltr.) Kosterm.	BT237	New Caledonia	Schlechter 15059	P00057340	KF291856	KF291915	KF291974	KF291432	KF291616	KF291793
D. yahouensis	BT238	New Caledonia	Schlechter 15059	P00057340	KF291857	KF291916	KF291975	KF291433	KF291617	KF291794
D. yahouensis	BT239	New Caledonia	Schlechter 15059	P00057340					KF291618	KF291795
D. yahouensis	VH3637	New Caledonia	Hequet et al. 3637	NOU017017	KF291858	KF291917	KF291976	KF291434	KF291619	KF291796
D. yatesiana Standl.	W27	Guatemala	Frisch s.n.	W	DQ923958	EU980777	DQ924065	EU981192		
D. sp. Pic N'ga	BT318	New Caledonia	Munzinger 6065	NOU	KF291839	KF291898	KF291957	KF291404	KF291572	KF291749
D. sp. Pic N'ga	BT319	New Caledonia	Munzinger 6065	NOU	KF291840	KF291899	KF291958	KF291405	KF291573	KF291750
D. sp. Pic N'ga	BT320	New Caledonia	Munzinger 6065	NOU	KF291841	KF291900	KF291959	KF291406	KF291574	KF291751
D. sp.	FS1637	Madagascar	Fischer & Sieder 1637	W	DQ923959	EU980778	DQ924066	EU981193		
D. sp.	FS2217	Madagascar	Fischer & Sieder 2217	W	DQ923960	EU980779	DQ924067	EU981194		
D. sp.	K20600	Madagascar	Rabenantoandro et al. 1246	MO	DQ923961	EU980780	DQ924068	EU981195		
D. sp.	K20601	Madagascar	Rabevohitra et al. 3660	MO	DQ923973	EU980796	DQ924082	EU981211		
D. sp.	K20613	Zambia	Zimba et al. 893	MO	DQ923962	EU980781	DQ924069	EU981196		
D. sp.	K20616	Ghana	Schmidt et al. 2207	MO	DQ923963	EU980783	DQ924070	EU981198		
D. sp.	K25759	Mauritius	Page 46	MAU	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291403		
D. sp.	RF958	Madagascar	RNF 958	W	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291407		
D. sp.	RF959	Madagascar	RNF 959	W	DQ923980	EU980804	DQ924089	EU981219		
D. sp.	RF970	Madagascar	RNF 970	W	DQ923964	EU980787	DQ924071	EU981202		
D. sp.	S10	Sri Lanka	Samuel 10	PDA	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291408		

(*continued on next page*)

Table 1 (*continued*)

D. sp.	S12	Sri Lanka	Samuel 12	PDA	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291409		
D. sp.	S18	Sri Lanka	Samuel 18	PDA	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291410		
D. sp.	S22	Sri Lanka	Samuel 22	PDA	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291411		
D. sp.	S25	Sri Lanka	Samuel 25	PDA	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291412		
D. sp.	S26	Sri Lanka	Samuel 26	PDA	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291413		
D. sp.	S28	Sri Lanka	Samuel 28	PDA	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291414		
D. sp.	W33	Madagascar	Sieder 440	W	KF291842	KF291901	KF291960	KF291415	KF291571	KF291748
D. sp.	W36	Madagascar	Sieder et al. 258	W	DQ923965	EU980788	DQ924072	EU981203		
D. sp.	W77	Madagascar	Sieder et al. 3079	W	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291416		
Euclea crispa (Thunb.) Gürke	Eb202	Living coll. HBV (EB 4/2)	Wallnöfer 13949	W	DQ923966	EU980789	DQ924073	EU981204		
Euclea crispa	K21188	Malawi	Chapman & Chapman 8085	K	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291435		
Euclea divinorum Hiern	Eb201	Cult. HBV (EB 2/1, Salisburg 69)	Wallnöfer & Duangjai 13947	W	DQ923967	EU980790	DQ924074	EU981205		
Euclea natalensis A.DC.	K21186	Zimbabwe	Timberlake & Cunliffe 4389	K	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291436		
Euclea natalensis	W08	South Africa	Kurzweil E514	W	DQ923968	EU980791	DQ924075	EU981206		
Euclea pseudobenus E. Mey. ex A.DC	K21190	Namibia	Ward 9205	K	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291437		
Euclea racemosa L.	K21183	Somalia	Thulin 10739	K	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291438		
Euclea sp.	W58	Tanzania	Kutalek 1-2001	W	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291439		
Euclea sp.	W59	Tanzania	Mbeyela 2-2001	W	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291440		
Euclea undulata Thunb.	Eb200	Cult. HBV (EB 5/2, 1973)	Wallnöfer 13897	W	EU980874	EU980792	DQ924076	EU981207	KF291620	KF291797
Royena cordata E. Mey ex A.DC	K1144	South Africa	Chase 1144	K	DQ923975	EU980799	DQ924084	EU981214		
Royena glabra L.	W05	South Africa	Kurzweil 2097	W	DQ923976	EU980800	DQ924085	EU981215		
Royena lucida L.	Eb203	South Africa	Wallnöfer & Duangjai 13943	W	DQ923977	EU980801	DQ924086	EU981216		
Royena lucida	W06	South Africa	Kurzweil E513	W	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291442		
Royena lycioides Desf. Ex A.DC	K977	South Africa	Chase 977	K	DQ923978	EU980802	DQ924087	EU981217		
Royena sp.	K1145	South Africa	Chase 1145	K	KF291859	KF291918	KF291977	KF291444		
Royena whyteana Hiern	Eb177	Africa	Kiehn s.n.	WU	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291443	KF291622	KF291799
Royena zombensis B.L. Burtt	K22558	Tanzania	Abdallah & Vollesen 95/106	K	Duangjai unpubl.	Duangjai unpubl.	Duangjai unpubl.	KF291445		
Lissocarpa benthamii Gürke	W61	Venezuela	Berry et al. 7217	PORT	DQ923969	EU980793	DQ924077	EU981208		
Lissocarpa guianensis Gleason	W04	Guyana	Arets s.n.	U	DQ923970	EU980794	DQ924078	EU981209		
Lissocarpa stenocarpa Steyerm.	K20609	Peru	Vásquez & Ortiz-Gentry 25233	MO	DQ923971	EU980795	DQ924079	EU981210		
Argania spinosa (L.) Skeels	K978	Morocco	Chase 978	K	DQ923981	EU980805	DQ924090	KF291308		
Cleyera japonica Thunb.	K1690	Japan	Chase 1690	K	DQ923985	EU980811	DQ924094	KF291309		
Halesia carolina L.	K910	USA	Chase 910	K					KF291621	KF291798
Madhuca macrophylla (Hassk.) H.J. Lam	K1363	Cult. Indonesia	Chase 1363	K	DQ923982	EU980806	DQ924091	KF291441		
Styrax benzoin Dryand.	K1371	Indonesia	Chase 1371	K	DQ923989	EU980809	DQ924098	KF291446		
Styrax officinalis L.	K872	Living coll. RGB Kew 1973-14474	Chase 872	K					KF291623	KF291800

species from clade II. Finally, across whole genus *Diospyros* there is a significant difference (Mann Whitney U test, $p < 0.001$) in genome size between clade III on the one hand and clades VI–XI on the other (Fig. 8). However, *D. pentamera* of clade II has a comparatively large genome (1C = 1.97 pg).

4. Discussion

Previous phylogenetic studies of *Diospyros* based on plastid markers demonstrated low levels of sequence divergence among New Caledonian species belonging to clade III (Duangjai et al., 2009), and inclusion of additional species in our investigation did not improve resolution in this group. Low-copy nuclear markers have been shown to be highly informative and useful for resolving phylogenetic relationships at lower taxonomic levels in some taxa (e.g. *Passiflora*: Yockteng and Nadot, 2004; *Paeonia*: Tank and Sang, 2001). The low-copy markers *ncpGS* and *PHYA* used here, however, did not improve resolution in this clade of 21 closely related species, thus preventing detection of hybrids and elucidation of geographical patterns (Fig. 5). There are also examples where

Table 2
PCR reactions.

ncpGS		1st *phyA*		2nd *phyA*	
18 μl	1.1xReddyMix (Thermo Scientific)	18 μl	1.1xReddyMix (Thermo Scientific)	18 μl	1.1xReddyMix (Thermo Scientific)
0.4 μl	Primer GScpDio1F (20 pM)	0.4 μl	Primer *PHYA* upstream (20 pM)	0.4 μl	Primer PhyADioF (20 pM)
0.4 μl	Primer GScpDioR (20 pM)	0.4 μl	Primer *PHYA* down-stream (20 pM)	0.4 μl	Primer PhyADioR (20 pM)
0.7 μl	Water	0.7 μl	Water	0.3 μl	Water
				0.4 μl	BSA (20 mg/ml)
0.5 μl	DNA	0.5 μl	DNA	0.5 μl	PCR product

BSA: bovine albumin serum (Thermo Scientific).

Table 3
PCR conditions.

ncpGS		1st *phyA*		2nd *phyA*	
95 °C for 2 min		95 °C for 2 min		95 °C for 2 min	
95 °C for 30 s		95 °C for 30 s		95 °C for 30 s	
58 °C for 30 s	35 cycles	52 °C for 30 s	35 cycles	60 °C for 30 s	35 cycles
72 °C for 2 min		70 °C for 2 min		72 °C for 1.5 min	
72 °C for 7 min		70 °C for 7 min		72 °C for 7 min	

Table 4
Primers used in this study.

Primer name	Fragment	Sequence (5'-3')	References
GScp839F	*ncpGS*	CACCAATGGGGAGGTTATGC	Yockteng and Nadot (2004)
GScp1056R	*ncpGS*	CATCTTCCCTCATGCTCTTTGT	Yockteng and Nadot (2004)
GscpDio1F	*ncpGS*	CCAATGGGGAGGTTATGCCTGGACAG	This study
GScpDioR	*ncpGS*	CATCTTCCCTCATGCTCTTTGTACTG	This study
PHYA upstream	*phyA*	GACTTTGARCCNGTBAAGCCTTAYG	Mathews and Donoghue (1999)
PHYA downstream	*phyA*	GDATDGCRTCCATYTCRTAGTC	Mathews and Donoghue (1999)
PhyADioF	*phyA*	GTBAAGCCTTAYGAAGTCCCGATGA	This study
PhyADioFi	*phyA*	GTCAAYGAGGGGGATGRAGAGGGAG	This study
PhyADioR	*phyA*	GCRTCCATYTCRTAGTCCTTCCAAG	This study
PhyADioRi	*phyA*	CTGATTYTCCAAYTCTAACTCCTTGTTGAC	This study

Table 5
Data characteristics and statistics from the maximum parsimony analyses of all three individual and the combined data sets.

	Combined plastid markers	*ncpGS*	*phyA*	Combined data set
Total no. of accessions	294	177	177	129
No. of outgroup accessions other than Ebenaceae	4	2	2	1
No. of outgroup accessions from Ebenaceae	21	2	2	2
No. of *Diospyros* accessions	269	173	173	126
No. of *Diospyros* species	149	64	64	64
No. of New Caledonian accessions	98	134	134	86
No. of New Caledonian species	28	28	28	28
No. of New Caledonian neoendemic accessions	83	112	112	74
No. of New Caledonian neoendemic species	21	21	21	21
Length of alignment	6556	1039	1187	8542
No. of variable characters	1880	532	374	1845
No. of parsimony informative characters	1126 (17.2%)	341 (32.8%)	223 (18.8%)	863 (10%)
No. of parsimony informative characters NCnc	44 (0.7%)	28 (2.7%)	14 (1.2%)	79 (0.9%)
Tree length of best parsimony tree (steps)	3808	1171	689	3259
Trees saved (parsimony analysis)	210	4810	1870	930
Consistency index	0.603	0.663	0.685	0.692
Retention index	0.857	0.857	0.893	0.848
Best fitting model	GTR + Γ + I	GTR + Γ	HKY + Γ + I	

low-copy nuclear markers were not able to fully resolve phylogenetic relationships between closely related species, especially on islands (e.g. Pillon et al., 2009a, 2013; Green et al., 2011). Nonetheless, the analysis based on combined plastid and nuclear data provides some resolution of relationships within the NC clade III. Of the 21 entities included in the analyses, seven species and one unidentified taxon formed well defined and inclusive clusters (Fig. 6). The remaining 14 species failed to form groups including all individuals of a particular species, but in many cases it was simply that some accessions were part of a polytomy and did not cluster consistently with any group.

In light of our results, members of the NC clade III appear little diverged but still form a strongly supported clade, which our dating analyses indicate are the result of recent rapid radiation. Only a few studies have examined the adaptive basis and processes involved in speciation in New Caledonia (e.g. Pillon et al., 2009b; Murienne et al., 2009). Rapid radiation has been observed in isolated areas such as islands (e.g. Givnish et al., 2009; Knope et al., 2012), high mountains (e.g. Hughes and Eastwood, 2006) and valleys (e.g. Givnish et al., 2007, 2011; Richardson et al., 2001). Island floras often show high levels of endemism and closely related species groups that result from a single colonisation event followed by rapid speciation, some of which have been hypothesised to represent adaptive radiations (e.g. Hawaiian silverswords, Baldwin and Sanderson, 1998; Hawaiian *Bidens*, Knope et al., 2012; *Araucaria* in New Caledonia, Gaudeul et al., 2012). The low levels of variation and resolution detected in the NC clade III prevent us from examining factors that may be promoting speciation on New Caledonia.

As all lineages of New Caledonian *Diospyros* seem to have arrived relatively recent on this island, the terms paleo-endemics

Table 6
Genome size of *Diospyros* and other genera from Ebenoideae. S.D.: standard deviation, N: number of measurements (replicates), *S.p.*: *Solanum pseudocapsicum*, *P.s.*: *Pisum sativum* 'Kleine Rheinländerin'.

Name	Acc. nr	1C-value	S.D.	N	Standard	Material
D. calciphila	BT313	1.99		1	S.p.	Dry
D. calciphila	BT316	1.97		1	P.s.	Silicagel
D. cherrieri	BT262	1.65	0.0092	5	S.p.	Dry
D. cherrieri	BT293	1.57	0.0117	2	S.p.	Silicagel
D. discolor	EBE100026	0.92	0.0020	3	S.p.	Fresh
D. erudita	BT260	2.17		1	S.p.	Dry
D. erudita	BT261	2.13	0.0367	3	S.p.	Dry
D. erudita	BT280	1.88	0.0253	3	S.p.	Silicagel
D. fasciculosa	BT012	1.19	0.0031	3	P.s.	Silicagel
D. fasciculosa	BT106	1.13	0.0064	3	P.s.	Silicagel
D. fasciculosa	BT144	1.22	0.0032	3	P.s.	Silicagel
D. fasciculosa	BT167	1.02	0.0318	4	P.s.	Silicagel
D. fasciculosa	BT212	1.09	0.0227	2	P.s.	Silicagel
D. fasciculosa	BT335	1.14	0.0300	3	P.s.	Silicagel
D. glans	BT019	2.03		1	S.p.	Silicagel
D. glans	BT093	2.02	0.0153	3	S.p.	Dry
D. impolita	BT101	1.79		1	P.s.	Silicagel
D. impolita	BT105	1.90	0.0132	3	P.s.	Silicagel
D. inconstans		1.13	0.0019	3	S.p.	Fresh
D. inexplorata	BT304	1.94	0.0693	3	P.s.	Silicagel
D. kaki	Sharon	2.29	0.0121	3	S.p.	Dry
D. lotus	EBE	0.87	0.0075	8	S.p.	Fresh
D. lotus	EBE03002	0.86	0.0012	3	S.p.	Fresh
D. mespiliformis	EBE000001	1.24	0.0029	3	P.s.	Fresh
D. mespiliformis	EBE100027	1.27	0.0035	3	P.s.	Fresh
D. minimifolia	BT230	1.57	0.0445	3	P.s.	Silicagel
D. olen	BT001	0.82	0.0062	3	S.p.	Silicagel
D. olen	BT036	0.87	0.0475	3	S.p.	Silicagel
D. olen	BT096	0.86	0.0042	3	S.p.	Dry
D. olen	BT186	0.90	0.0041	3	S.p.	Dry
D. pancheri	BT077	2.28	0.0129	3	S.p.	Dry
D. parviflora	BT085	2.16	0.0493	3	P.s.	Dry
D. pentamera	EBE030020	1.97	0.0020	3	S.p.	Fresh
D. perplexa	BT002	2.27		1	S.p.	Silicagel
D. pustulata	BT137	1.54	0.0490	2	P.s.	Silicagel
D. revolutissima	BT222	2.05	0.0148	4	P.s.	Dry
D. texana	EBE020015	0.89	0.8849	3	S.p.	Fresh
D. texana	EBE100029	0.89	0.0019	3	S.p.	Fresh
D. tridentata	BT205	2.21	0.0246	2	S.p.	Dry
D. tridentata	BT206	2.09		1	P.s.	Dry
D. umbrosa	BT171	1.61		1	P.s.	Silicagel
D. umbrosa	BT247	1.51	0.0894	3	P.s.	Silicagel
D. vieillardii	BT100	1.55		1	P.s.	Dry
D. vieillardii	BT216	1.57	0.0238	3	P.s.	Dry
D. yatesiana		0.60	0.0010	5	S.p.	Fresh
E. divinorum	EBE000002	1.98	0.0014	3	S.p.	Fresh
E. undulata	EBE100002	0.74	0.0220	3	S.p.	Fresh
R. whyteana	EBE030021	0.79	0.0031	3	S.p.	Fresh
R. whyteana	EBE030022	0.78	0.0007	3	S.p.	Fresh

Table 7
Main habitats of New Caledonian neoendemic *Diospyros* species.

Habitat		Species
Maquis on ultramafic substrates		D. *erudita*, *D. pancheri*, *D. parviflora*, *D. tireliae*, *D. vieillardii*
Dry forests on non-ultramafic substrates		*D. cherrieri*, *D. perplexa*, *D. yahouensis*
Humid forests at low elevations	Ultramafic substrates	*D. pancheri*, *D. parvilfora*, *D. umbrosa*
	Calcareous rocks	*D. tridentata*
Humid mountain forests on schist		*D. flavocarpa*, *D. labillardierei*, *D. trisulca*
Dry coastal forests	Black clays	*D. veillonii*, *D. minimifolia*, *D. pustulata* (the latter two can also occur on calcareous substrates)
	Schist	*D. impolita*
	Ultramafic substrates	*D. revolutissima*
	Various substrates	*D. pancheri*
Coastal forests on coralline substrates		*D. calciphila*, *D. impolita*
Humid forests on the east coast		*D. glans*

and neo-endemics used by Duangjai et al. (2009) were not used here. The common ancestor of clade III diverged about 19 mya (Fig. 7), and the earlier diverging species occur mainly in Africa and on islands of the western Indian Ocean (e.g. Madagascar). Our results in combination with the DIVA analysis from Duangjai et al. (2009) indicate that, from there, this group spread eastwards

via Southeast Asia, where it arrived around 15 mya, and then reached the Hawaiian Archipelago and New Caledonia around 9–10 mya. This time of colonisation is consistent with that found for other plant groups (reviewed in Pillon, 2012) and animals (e.g. Nattier et al., 2011). The close relationship of New Caledonian and Hawaiian endemic *Diospyros* shows that migration around the Pacific Ocean has taken place, but to make more definite conclusions about the direction of dispersal, data from species present on other islands between New Caledonia and Hawai'i are needed. In contrast to long-held hypotheses that many taxa are Gondwanan relicts (e.g. Lowry II, 1998; Swenson et al., 2001), our results suggest that all groups of New Caledonian *Diospyros* are much younger than 37 myr (when New Caledonia re-emerged) and arrived, like many others, via long-distance dispersal (e.g. Bartish et al., 2011; Espeland and Murienne, 2011; Murienne, 2009).

The closely related species of the NC clade III are distinguishable from one another by morphological characters (e.g. leaf, flower, fruit and calyx characters), and many of them are found in different habitats (e.g. humid/dry, different substrate types, different elevations, etc.). Leaf morphology shows adaptation to the environment in which a species occurs (e.g. species found in dry habitats have sclerophyllous leaves; for details of species descriptions see White, 1992, 1993). In most plant groups, closely related species rarely occur in sympatry, but not in New Caledonia where this seems to be a common pattern in several groups (J. Munzinger pers. obs.), including *Diospyros*. However, *Diospyros* has been reported to be one of the few genera outside New Caledonia (e.g. Madagascar) with several co-occurring species (pers. comm. P.S. Ashton). The habitats occupied by the New Caledonian *Diospyros* species belonging to clade III can be roughly divided into seven groups (Table 7). *D. vieillardii*, a common species found all over Grande Terre and the islands north of the main island in maquis vegetation, occurs on a variety of substrates, including ultramafic. *Diospyros umbrosa*/*D. flavocarpa* are sister to the remainder of the clade excluding *D. vieillardii* (Fig. 6). *D. umbrosa* occurs only on ultramafic substrates in comparatively humid forests mainly consisting of *Nothofagus* and *Araucaria*. *D. flavocarpa* is found on schist in middle elevation forests in northeastern Grande Terre. *Diospyros cherrieri* (a local endemic in dry forests on basalts at the western coast of Grande Terre, Fig. 1I) and *D. veillonii* (a local endemic in dry coastal forests on black clay on the western side of Grande Terre, Fig. 1F) are together sister to the rest of the clade (minus those mentioned above). The clade comprising *D. calciphila*, *D. inexplorata* (littoral forests on coralline substrates) and *D.* sp. from Pic N'ga (maquis on ultramafic substrate on Ile des Pins) is well supported. Relationships among all other members of the clade could not be resolved with the markers used, although most of them are morphologically and ecologically well defined. This phenomenon (morphological and ecological distinctiveness, but no resolution) is found, for example, in *D. labillardieri* (lanceolate leaves, hanging branches; river edges in middle elevation forests on schist, Fig 1D), *D. pancheri* (obcordate pubescent leaves, hanging branches, humid forests at low elevation on ultramafic soils, Fig. 1E) and many others. Due to the poor resolution of the phylogenetic trees, possible grouping of New Caledonian *Diospyros* species according to their ecological niches remains untested.

A greater than threefold variation within the genome size of *Diospyros* is observed, although the chromosome counts performed here and elsewhere indicates that they are diploid with $2n = 30$, and we hypothesize that the most recent common ancestor of *Diospyros* had a large genome because species belonging to earlier diverging clades (e.g. *E. divinorum* and *D. pentamera*) have large genomes. Developing firmer ideas about evolution of genome size in *Diospyros* would require many more measurements of species from throughout the phylogenetic tree, especially species from islands in the Indian and Pacific Ocean, which will be key to assessing evolution of genome size in NC clade III. The limited data available today suggest that polyploidy seems to be rare among wild *Diospyros* species. The diversification of species of the NC clade III remains an overall poorly understood subject, despite our extensive efforts to find variation relevant to addressing these questions. It seems that we can eliminate polyploidy as one feature of their evolution, but the question of the involvement of hybridization cannot be eliminated without the use of more variable markers. To address patterns of speciation and factors promoting divergence, we will have to turn more markers used in population genetic studies, such as AFLPs, microsatellites or fingerprinting methods based on next generation sequencing methods.

Acknowledgments

This work was funded by a grant from the Austrian Science Fund (FWF, Project-Number: P 22159-B16) awarded to R. Samuel. The authors thank the team of the Department of Systematic and Evolutionary Botany as well as the team of the Botanic Department of IRD Noumea for support with this study. Special thanks go to V. Klenja for the lab work. Thanks to the following persons for their help with lab work, field work and ideas to improve our manuscript: J.-P. Butin, C. Chambrey, G. Dagostini, E. Grasserbauer, V. Hequet, G. Kohl, D. & I. Létocart, F. Maghuly, W. Nigote, O. Paun, G. Schneeweiss, J. Schönenberger, H. Vandrot, Fam. Villegente, B. Wallnöfer and H. Weiss-Schneeweiss.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2013.07.002.

References

Alba, R., Kelmsnson, P.M., Cordonnier-Pratt, M.-M., Pratt, L.H., 2000. The phytochrome gene family in tomato and the rapid differential evolution of this family in angiosperms. Molecular Biology and Evolution 17, 362–373.

APG III, 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. Botanical Journal of the Linnean Society 161, 105–121.

Baldwin, B.G., Sanderson, M.J., 1998. Age and rate of diversification of the Hawaiian silversword alliance (Compositae). Proceedings of the National Academy of Sciences USA 95, 9402–9406.

Bartish, I.V., Antonelli, A., Richardson, J.E., Swenson, U., 2011. Vicariance or long-distance dispersal: historical biogeography of the pantropical subfamily Chrysophylloideae (Sapotaceae). Journal of Biogeography 38, 177–190.

Basinger, J.F., Christophel, D.C., 1985. Fossil flowers and leaves of the Ebenaceae from the Eocene of southern Australia. Canadian Journal of Botany 63, 1825–1843.

Bell, C.D., Soltis, D.E., Soltis, P.S., 2010. The age and diversification of the angiosperms re-revisited. American Journal of Botany 97, 1296–1303.

Bennett, M.D., Leitch, I.J., 2005. Genome size evolution in plants. In: Gregory, T.R. (Ed.), The Evolution of the Genome. Elsevier, San Diego, pp. 90–151.

Bennett, M.D., Leitch, I.J., 2010. Angiosperm DNA C-Values Database (Release 7.0, December 2010). <http://www.kew.org/cvalues/>.

Bennett, J.R., Mathews, S., 2006. Phylogeny of the parasitic plant family Orobanchaceae inferred from phytochrome A1. American Journal of Botany 93, 1039–1051.

Bremer, K., Friis, E.M., Bremer, B., 2004. Molecular phylogenetic dating of asterid flowering plants shows early Cretaceous diversification. Systematic Biology 53, 496–503.

Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. JModelTest 2: more models, new heuristics and parallel computing. Nature Methods 9, 772.

Drummond, A.J., Ho, S.Y.W., Phillips, M.J., Rambaut, A., 2006. Relaxed phylogenetics and dating with confidence. Molecular Biology and Evolution 29, 1969–1973 (PLoS Biology 4, 699–710).

Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution 29, 1969–1973.

Duangjai, S., Wallnöfer, B., Samuel, R., Munzinger, J., Chase, M.W., 2006. Generic delimitation and relationships in Ebenaceae sensu lato: evidence from six plastid DNA regions. American Journal of Botany 93, 1808–1827.

Duangjai, S., Samuel, R., Munzinger, J., Forest, F., Wallnöfer, B., Barfuss, M.H.J., Fischer, G., Chase, M.W., 2009. A multi-locus plastid phylogenetic analysis of the pantropical genus *Diospyros* (Ebenaceae), with an emphasis on the radiation

and biogeographic origins of the New Caledonian endemic species. Molecular Phylogenetics and Evolution 52, 602–620.
Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research 32, 1792–1797.
Emshwiller, E., Doyle, J.J., 1999. Chloroplast-expressed glutamine synthetase (*ncpGS*): potential utility for phylogenetic studies with an example from *Oxalis* (Oxalidaceae). Molecular Phylogenetics and Evolution 12, 310–319.
Espeland, M., Murienne, J., 2011. Diversity dynamics in New Caledonia: towards the end of the museum model? BMC Evolutionary Biology 11, 254.
Farris, J.S., Källersjö, M., Kluge, A.G., Bult, C., 1994. Testing significance of incongruence. Cladistics 10, 315–319.
Ferreira, M.A.R., Suchard, M.A., 2008. Bayesian analysis of elapsed times in continuous-time Markov chains. The Canadian Journal of Statistics 36, 355–368.
Forest, F., 2009. Calibrating the Tree of Life: fossils, molecules and evolutionary timescales. Annals of Botany 104, 789–794.
Galbraith, D.W., Harkins, K.R., Maddox, J.M., Ayres, N.M., Sharma, D.P., Firoozabady, E., 1983. Rapid flow cytometric analysis of the cell cycle in intact plant tissues. Science 220, 1049–1051.
Gaudeul, M., Rouhan, G., Gardner, M.F., Hollingsworth, P.M., 2012. AFLP markers provide insights into the evolutionary relationships and diversification of New Caledonian *Araucaria* species (Araucariaceae). American Journal of Botany 99, 68–81.
Geeraerts, A., Raeymaekers, J.A.M., Vinckier, S., Pletsers, A., Smets, E., Huysmans, S., 2009. Systematic palynology in Ebenaceae with focus on Ebenoideae: morphological diversity and character evolution. Review of Palaeobotany and Palynology 153, 336–353.
Gernhard, T., 2008. The conditioned reconstructed process. Journal of Theoretical Biology 253, 769–788.
Givnish, T.J., Millam, K.C., Berry, P.E., Sytsma, K.J., 2007. Phylogeny, adaptive radiation, and historical biogeography of Bromeliaceae inferred form *ndhF* sequence data. Aliso 23, 3–26.
Givnish, T.J., Millam, K.C., Mast, A.R., Paterson, T.B., Theim, T.J., Hipp, A.L., Henss, J.M., Smith, J.F., Wood, K.R., Sytsma, K.J., 2009. Origin, adaptive radiation and diversification of the Hawaiian lobeliads (Asterales: Campanulaceae). Proceedings of the Royal Society B 276, 407–146.
Givnish, T.J., Barfuss, M.H.J., Van Ee, B., Riina, R., Schulte, K., Horres, R., Gonsiska, P.A., Jabaily, R.S., Crayn, D.M., Smith, J.A.C., Winter, K., Brown, G.K., Evans, T.M., Holst, B.K., Luther, H., Till, W., Zizka, G., Berry, P.E., Sytsma, K.J., 2011. Phylogeny, adaptive radiation, and historical biogeography in Bromeliacae: insight from an eight-locus plastid phylogeny. American Journal of Botany 98, 872–895.
Grandcolas, P., Murienne, J., Robillard, T., Desutter-Grandcolas, L., Jourdan, H., Guilbert, E., Deharveng, L., 2008. New Caledonia: a very old Darwinian island? Philosophical Transactions of the Royal Society of London Series B, Biological Sciences 363, 3309–3317.
Green, A.F., Ramsey, T.S., Ramsey, J., 2011. Phylogeny and biogeography of ivies (*Hedera* spp., Araliaceae), a polyploid complex of woody vines. Systematic Botnay 36, 1114–1127.
Greilhuber, J., Ebert, I., 1994. Genome size variation in *Pisum sativum*. Genome 37, 646–655.
Guindon, S., Gascuel, O., 2003. A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. Systematic Biology 52, 696–704.
Guo, Y.-P., Wang, S.-Z., Vogl, C., Ehrendorfer, F., 2012. Nuclear and plastid haplotypes suggest rapid diploid and polyploid speciation in the N Hemisphere *Achillea millefolium* complex (Asteraceae). BMC Evolutionary Biology 12, 2.
Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. Nucleic Acids Symposium Series 41, 95–98.
Hasegawa, M., Kishino, H., Yano, T.-A., 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. Journal of Molecular Evolution 22, 160–174.
Hughes, C., Eastwood, R., 2006. Island radiation on a continental scale: exceptional rates of plant diversification after uplift of the Andes. PNAS 103, 10334–10339.
Knope, M.L., Morden, C.W., Funk, V.A., Fukami, T., 2012. Area and the rapid radiation of Hawaiian *Bidens* (Asteraceae). Journal of Biogeography 39, 1206–1216.
Ladiges, P.Y., Cantrill, D., 2007. New Caledonia–Australian connections: biogeographic patterns and geology. Australian Systematic Botany 20, 383–389.
Leitch, I.J., 2007. Genome size through the ages. Heredity 99, 121–122.
Lowry II, P.P., 1998. Diversity, endemism and extinction in the flora of New Caledonia: a review. In: Peng, C.F., Lowry, P.P., II (Eds.), Rare, Threatened, and Endangered Floras of Asia and the Pacific Rim. Institute of Botany, Taipei, Taiwan, pp. 181–206.
Magallón, S., 2010. Using fossils to break long branches in molecular dating: a comparison of relaxed clocks applied to the origin of angiosperms. Systematic Biology 59, 384–399.
Mathews, S., Donoghue, M.J., 1999. The root of angiosperm phylogeny inferred from duplicate phytochrome genes. Science 286, 947–950.
Mathews, S., Clements, M.D., Beilstein, M.A., 2010. A duplicate gene rooting of seed plants and the phylogenetic position of flowering plants. Philosophical Transaction of the Royal Society B Biological Sciences 365, 383–395.
McLoughlin, S., 2001. The breakup history of Gondwana and its impact on pre-Cenozoic floristic provincialism. Australian Journal of Botany 49, 271–300.
Millán-Martínez, M., 2010. Fossil record and age of the Asteridae. Botanical Reviews 76, 83–135.
Miller, M.A., Pfeiffer, W., Schwarz, T., 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: Proceedings of the Gateway Computing Environments Workshop (GCE), New Orleans, Louisiana, 14 November, 2010, pp. 1–8.
Mittermeier, R.A., Gil, P.R., Hoffmann, M., Pilgrim, J., Brooks, T., Mittermeier, C.G., Lamoreux, J., da Fonseca, G.A.B., 2004. Hotspots Revisited: Earth's Biologically Richest and Most Endangered Terrestrial Ecoregions. CEMEX, Mexico City.
Moody, M.L., Rieseberg, L.H., 2012. Sorting through the chaff, nDNA gene trees for phylogenetic inference and hybrid identification of annual sunflowers (*Helianthus* sect. *Helianthus*). Molecular Phylogenetics and Evolution 64, 145–155.
Morat, P., Jaffré, T., Tronchet, F., Munzinger, J., Pillon, Y., Veillon, J.-M., Chalopin, M., 2012. Le référentiel taxonomique Florical et les caractéristiques de la flore vasculaire indigène de la Nouvelle-Calédonie. Adansonia 34, 177–219.
Murienne, J., 2009. Testing biodiversity hypothesis in New Caledonia using phylogenetics. Journal of Biogeography 36, 1433–1434.
Murienne, J., Grandcolas, P., Piulachs, M., Bellés, X., D'Haese, C., Legendre, F., Pellens, R., Guilbert, E., 2005. Evolution on a shaky piece of Gondwana: is local endemism recent in New Caledonia? Cladistics 21, 2–7.
Murienne, J., Guilbert, E., Grandcolas, P., 2009. Species' diversity in the New Caledonian endemic genera *Cephalidiosus* and *Nobarnus* (Insecta: Heteroptera: Tingidae), an approach using phylogeny and species' distribution modelling. Biological Journal of the Linnean Society 97, 177–184.
Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., Kent, J., 2000. Biodiversity hotspots for conservation priorities. Nature 403, 853–858.
Nattier, R., Robillard, T., Desutter-Grandcolas, L., Couloux, A., Grandcolas, P., 2011. Older than New Caledonia emergence? A molecular phylogenetic study of the eneopterine crickets (Orthoptera: Grylloidea). Journal of Biogeography 38, 2195–2209.
Nie, Z.-L., Wen, J., Azuma, H., Qui, Y.-L., Sun, H., Meng, Y., Sun, W.-B., Zimmer, E.A., 2008. Phylogenetic and biogeographic complexity of Magnoliaceae in the Northern Hemisphere inferred from three nuclear data sets. Molecular Phylogenetics and Evolution 48, 1027–1040.
Nixon, K.C., Crepet, W.L., 1993. Late cretaceous fossil flowers with ericalean affinity. American Journal of Botany 80, 616–623.
Otto, F., Oldiges, H., Göhde, W., Jain, V.K., 1981. Flow cytometric measurement of nuclear DNA content variations as a potential in vivo mutagenicity test. Cytometry 2, 189–191.
Parisod, C., Alix, K., Just, J., Petit, M., Sarilar, V., Mhiri, C., Ainouche, M., Chalhoub, B., Grandbastien, M.A., 2009. Impact of transposable elements on the organization and function of allopolyploid genomes. New Phytologist 186, 37–45.
Pelletier, B., 2006. Geology of the New Caledonia region and its implications for the study of the New Caledonian biodiversity. In: Payri, C., Richer de Forges, B. (Eds.), Compendium of marine species from New Caledonia, Documents Scientifiques et Techniques II4. Institut de Recherche pour le Développement Nouméa, pp. 17–30.
Pellicer, J., Fay, M.F., Leitch, I.J., 2010. The largest eukaryotic genome of them all? Botanical Journal of the Linnean Society 164, 10–15.
Petrov, D.A., 2001. Evolution of genome size: new approaches to an old problem. Trends in Genetics 17, 23–28.
Pillon, Y., 2012. Time and tempo of diversification in the flora of New Caledonia. Botanical Journal of the Linnean Society 170, 288–298.
Pillon, Y., Hopkins, H.C., Munzinger, J., Amir, H., Chase, M.W., 2009a. Cryptic species, gene recombination and hybridization in the genus *Spiraeanthemum* (Cunoniaceae) from New Caledonia. Botanical Journal of the Linnean Society 161, 137–152.
Pillon, Y., Munzinger, J., Amir, H., Hopkins, H.C., Chase, M.W., 2009b. Reticulate evolution on a mosaic of soils: diversification of the New Caledonian endemic genus *Codia* (Cunoniaceae). Molecular Ecology 18, 2263–2275.
Pillon, Y., Munzinger, J., Amir, H., Lebrun, M., 2010. Ultramafic soils and species sorting in the flora of New Caledonia. Journal of Ecology 98, 1108–1116.
Pillon, Y., Johansen, J., Sakishima, T., Chamala, S., Barbazuk, W.B., Roalson, E.H., Price, D.K., Stacy, E.A., 2013. Potential use of low-copy nuclear genes in DNA barcoding: a comparison with plastid genes in two Hawaiian plant radiations. BMC Evolutionary Biology 13, 35.
Pirie, M.D., Doyle, J.A., 2012. Dating clades with fossils and molecules: the case of *Annoaceae*. Botanical Journal of the Linnean Society 169, 84–116.
Richardson, J.E., Pennington, R.T., Pennington, T.D., Hollingsworth, P.M., 2001. Rapid diversification of a species-rich genus of Neotropical rain forest trees. Science 293, 2242–2245.
Salard, M., Avias, J., 1968. Contribution à la connaissance de la flore fossile de la Nouvelle-Calédonie. Palaeontographica, Abteilung B, Paläophytologie 124, 1–44.
Silvestro, D., Schnitzler, J., Zizka, G., 2011. A Bayesian framework to estimate diversification rates and their variation through time and space. BMC Evolutionary Biology 11, 311–325.
Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688–2690.
Swenson, U., Hill, R.S., McLoughlin, S., 2001. Biogeography of *Nothofagus* supports the sequence of Gondwana break-up. Taxon 50, 1025–1041.
Swofford, D.L., 2003. PAUP*. Phylogenetic Analysis Using Parsimony (* and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Tamura, M., Tao, R., Yonemori, K., Utsunomiya, N., Sugiura, A., 1998. Ploidy level and genome size of several *Diospyros* species. Journal of the Japanese Horticultural Society 67, 306–312.

Tank, D.C., Sang, T., 2001. Phylogenetic utility of the glycerol-3-phosphate acyltransferase gene: evolution and implications in *Paeonia* (Paeoniaceae). Molecular Phylogenetics and Evolution 19, 421–429.

Tavaré, S., 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. Lectures on Mathematics in the Life Sciences 17, 57–86.

Tel-Zur, N., Abbo, S., Myslabodsky, D., Mizrahi, Y., 1999. Modified CTAB procedure for DNA isolation from ephiphytic cacti of the genera *Hylocereus* and *Selenicereus* (Cactaceae). Plant Molecular Biology Reporter 17, 249–254.

Temsch, E.M., Greilhuber, J., Krisai, R., 2010. Genome size in liverworts. Preslia 82, 63–80.

Weiss-Schneeweiss, H., Villaseñor, J.L., Stuessy, T.F., 2009. Chromosome numbers, karyotypes, and evolution in *Melampodium* (Asteraceae). International Journal of Plant Science 170, 1168–1182.

White, F., 1992. Twenty-two new and little known species of *Diospyros* (Ebenaceae) from New Caledonia with comments on section *Maba*. Bulletin du Muséum National d'Histoire naturelle 4ème série – Section B. Adansonia 2, 179–222.

White, F., 1993. Flore de la Nouvelle-Calédonie et Dépendances. 19. Ébénacées. Muséum National d'Histoire Naturelle, Paris.

Yockteng, R., Nadot, S., 2004. Phylogenetic relationships among *Passiflora* species based on the glutamine synthetase nuclear gene expressed in chloroplast (*ncpGS*). Molecular Phylogenetics and Evolution 31, 379–396.

Yule, G.U., 1925. A mathematical theory of evolution, based on the conclusions of Dr. J.C. Willis, FRS. Philosophical Transactions of the Royal Society of London Series B 213, 21–87.

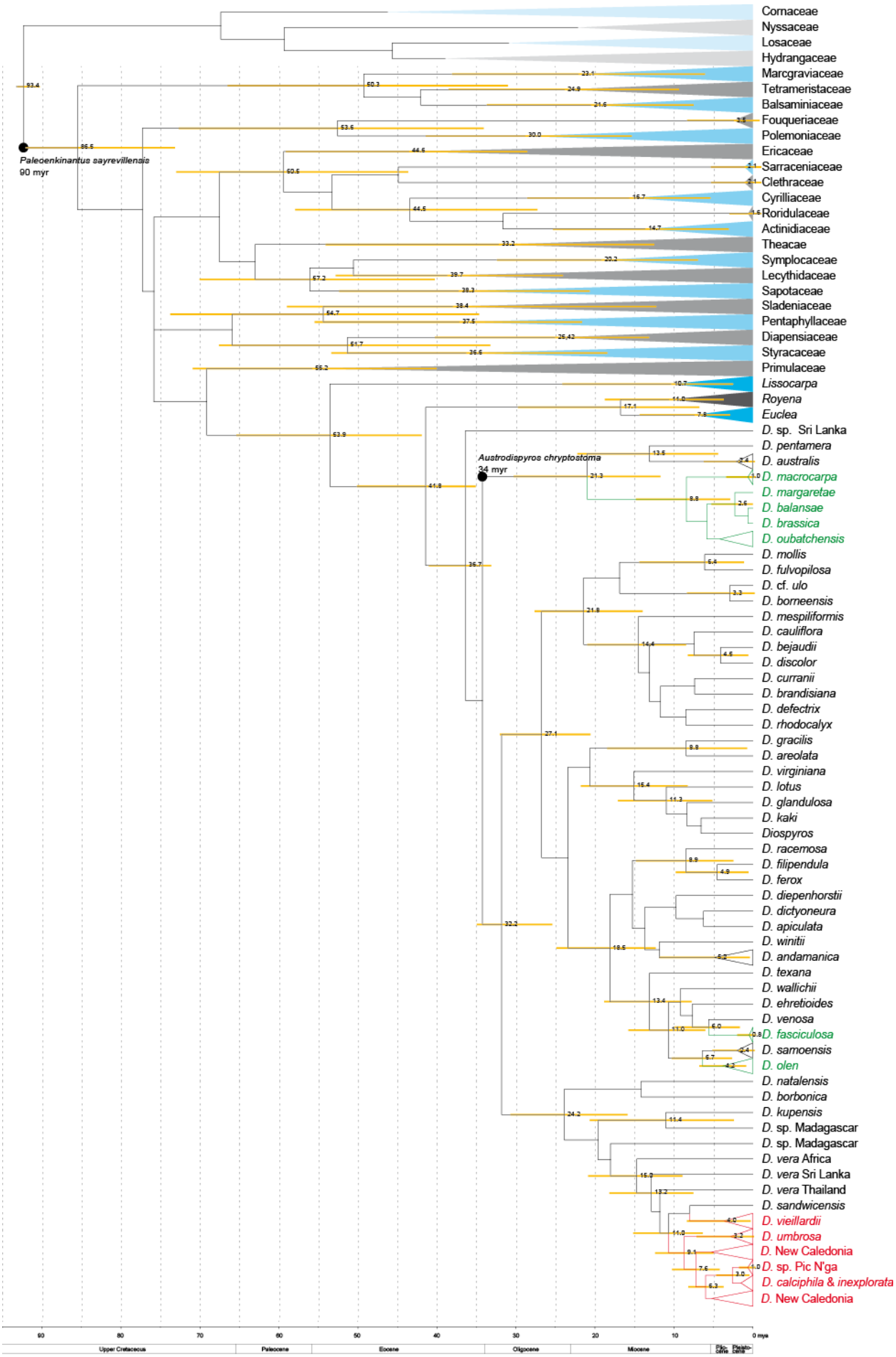
Zimmer, E.A., Wen, J., 2012. Using nuclear gene data for plant phylogenetics: progress and prospects. Molecular Phylogenetics and Evolution 65, 774–785.

Zuckerkandl, H., Pauling, L., 1965. Evolutionary divergence and convergence in proteins. In: Bryson, G., Vogel, H.J. (Eds.), Evolving Genes and Proteins. Academic Press, New York, pp. 97–166.

SUPPLEMANTARY MATERIAL

Supplemantay files S1 – S3 are not given here, because these are data files (BEAST input files for Bayesian analyses).

Supplementary Figure 4: Dated phylogeny of Ericales based on the a joint matrix the data set of Bell *et al*. (2010) together with our plastid sequences. Taxa from families other than Ebenaceae are collapsed to family level, taxa other than *Diospyros* are collapsed to generic level. Multiple accessions of a species are collapsed to species level. The NC clade III part of the tree is mostly collapsed due to lack of support of respective nodes. Nodes which were calibrated with fossils are marked with a black dot. Yellow bars represent the 95% highest posterior density interval. New Caledonian taxa are coloured, red represents clade III NC.


Paleoenkinantus sayrevillensis
90 myr
Austrodispyros chryptostoma
34 myr
Cornaceae
Nyssaceae
Losaceae
Hydrangeaceae
Marcgraviaceae
Tetrameristaceae
Balsaminiaceae
Fouqueriaceae
Polemoniaceae
Ericaceae
Sarraceniaceae
Clethraceae
Cyrilliaceae
Roridulaceae
Actinidiaceae
Theacae
Symplocaceae
Lecythidaceae
Sapotaceae
Sladeniaceae
Pentaphyllaceae
Diapensiaceae
Styracaceae
Primulaceae
Lissocarpa
Royena
Euclea
D. sp. Sri Lanka
D. pentamera
D. australis
D. macrocarpa
D. margaretae
D. balansae
D. brassica
D. oubatchensis
D. mollis
D. fulvopilosa
D. cf. ulo
D. borneensis
D. mespiliformis
D. cauliflora
D. bejaudii
D. discolor
D. curranii
D. brandisiana
D. defectrix
D. rhodocalyx
D. gracilis
D. areolata
D. virginiana
D. lotus
D. glandulosa
D. kaki
Diospyros
D. racemosa
D. filipendula
D. ferox
D. diepenhorstii
D. dictyoneura
D. apiculata
D. winitii
D. andamanica
D. texana
D. wallichii
D. ehretioides
D. venosa
D. fasciculosa
D. samoensis
D. olen
D. natalensis
D. borbonica
D. kupensis
D. sp. Madagascar
D. sp. Madagascar
D. vera Africa
D. vera Sri Lanka
D. vera Thailand
D. sandwicensis
D. vieillardii
D. umbrosa
D. New Caledonia
D. sp. Pic N'ga
D. calciphila & inexplorata
D. New Caledonia
90
80
70
60
50
40
30
20
10
0 mya
Upper Cretaceous
Paleocene
Eocene
Oligocene
Miocene
Plio-cene
Pleisto-cene

CHAPTER 2

Analyses of amplified fragment length polymorphisms (AFLP) indicate rapid radiation of *Diospyros* species (Ebenaceae) endemic to New Caledonia

Barbara Turner, Ovidiu Paun, Jérôme Munzinger, Sutee Duangjai, Mark W. Chase, Rosabelle Samuel

Status: published, BMC Evolutionary Biology 13 (2013) article 269, DOI: 10.1186/1471-2148-13-269

Contribution: Collection of material, collection of data, analysis of data, phylogenetic analysis, manuscript writing/editing

Turner et al. BMC Evolutionary Biology 2013, 13:269
http://www.biomedcentral.com/1471-2148/13/269

BMC Evolutionary Biology

RESEARCH ARTICLE Open Access

Analyses of amplified fragment length polymorphisms (AFLP) indicate rapid radiation of *Diospyros* species (Ebenaceae) endemic to New Caledonia

Barbara Turner[1*], Ovidiu Paun[1], Jérôme Munzinger[2], Sutee Duangjai[3], Mark W Chase[4,5] and Rosabelle Samuel[1]

Abstract

Background: Radiation in some plant groups has occurred on islands and due to the characteristic rapid pace of phenotypic evolution, standard molecular markers often provide insufficient variation for phylogenetic reconstruction. To resolve relationships within a clade of 21 closely related New Caledonian *Diospyros* species and evaluate species boundaries we analysed genome-wide DNA variation via amplified fragment length polymorphisms (AFLP).

Results: A neighbour-joining (NJ) dendrogram based on Dice distances shows all species except *D. minimifolia*, *D. parviflora* and *D. vieillardii* to form unique clusters of genetically similar accessions. However, there was little variation between these species clusters, resulting in unresolved species relationships and a star-like general NJ topology. Correspondingly, analyses of molecular variance showed more variation within species than between them. A Bayesian analysis with BEAST produced a similar result. Another Bayesian method, this time a clustering method, STRUCTURE, demonstrated the presence of two groups, highly congruent with those observed in a principal coordinate analysis (PCO). Molecular divergence between the two groups is low and does not correspond to any hypothesised taxonomic, ecological or geographical patterns.

Conclusions: We hypothesise that such a pattern could have been produced by rapid and complex evolution involving a widespread progenitor for which an initial split into two groups was followed by subsequent fragmentation into many diverging populations, which was followed by range expansion of then divergent entities. Overall, this process resulted in an opportunistic pattern of phenotypic diversification. The time since divergence was probably insufficient for some species to become genetically well-differentiated, resulting in progenitor/derivative relationships being exhibited in a few cases. In other cases, our analyses may have revealed evidence for the existence of cryptic species, for which more study of morphology and ecology are now required.

Keywords: Cryptic species, Island flora, Morphological diversification, Progenitor/derivative relationships, Species radiation, Woody plants

* Correspondence: barbara.turner@univie.ac.at
[1]Department of Systematic and Evolutionary Botany, Faculty of Life Sciences, University Vienna, Rennweg 14, 1030 Wien, Austria
Full list of author information is available at the end of the article

Background

Island floras are often characterized by high levels of endemism and groups of closely related but morphological and ecological divergent species that are mostly the result of single colonisation events followed by radiation e.g. [1,2]. New Caledonia was cited as one of the 34 biodiversity hotspots recognized by Conservation International [3,4]. Nearly 75% of the native flora is endemic [5], which is the fourth highest for an island [6]. While the continental part of New Caledonia (mainly Grande Terre) was entirely submerged during the Eocene (until 37 mya), a thick layer of heavy-metal-rich oceanic mantle accumulated [7]. Today, around one-third of the main island, Grande Terre, is still overlaid with ultramafic substrates. Generally, Grande Terre is a substrate mosaic [8], which is cited as one reason for the high level of endemism found there e.g. [9]. The climate in New Caledonia ranges from tropical to subtropical, and the main island is split by a mountain range into a humid eastern and a dry western part with prevailing winds and rain coming from the south-east. Taking climatic and geological factors together, Grande Terre has a wide range of environmental diversity. The main vegetation types in New Caledonia are evergreen humid forests, maquis, dry forests, littoral vegetation, and savannah [10].

One plant group that has taken advantage of many available habitats on New Caledonia is *Diospyros*, which is the largest genus (> 500 species in its broad circumscription [11]) of Ebenaceae, a pantropical family of woody plants. In New Caledonia *Diospyros* species range from sea level up to ca. 1250 m (the highest point New Caledonia is 1628 m), and species are found in all vegetation types except mangroves, with several species co-occurring in micro-sympatry (Table 1).

Diospyros colonised New Caledonia via long-distance dispersal at least four times [12]. In previous studies based on low-copy nuclear and/or multiple plastid markers [12,13], it was possible to resolve phylogenetic relationships for the majority of *Diospyros* species, except for one group of endemics from New Caledonia. Of the 31 New Caledonian *Diospyros* species, 24 belong to this clade of closely related endemics. In previous analyses, this strongly supported group is related to species found on islands throughout the Indian and Pacific Oceans as far east as Hawai'i [12,13]. However, due to extremely low levels of sequence divergence, it was not possible to tease apart relationships between these species (they formed a hard polytomy in most individual trees, and there was little informative variation that permitted clustering of pairs or groups of species). Most of these closely related species are morphologically and ecologically clearly differentiated (for examples see [13]), and several species are narrow endemics restricted to small areas.

Amplified fragment length polymorphism (AFLP; [14]) is a fingerprinting technique that has proven to be useful for revealing phylogenetic relationship among closely related taxa (e.g. *Hypochaeris*, [15]; *Lactuca*, [16]; *Phylica*, [17]; *Trollius*, [18]; *Ranunculus alpestris*, [19]; *Puya*, [20,21]; *Araucaria*, [22]). In contrast to standard phylogenetic markers, AFLP variation is spread across the whole genome, spanning both coding and non-coding DNA regions and may therefore be more representative of overall genetic patterns present as well as being highly informative for phylogenetic analyses at the low phylogenetic level [23,24]. Compared to other fingerprinting techniques AFLP shows increased reproducibility and does not require any prior knowledge of the analysed genomes. However, there are some detrimental issues to consider when working with AFLP data; these include potential non-homology and non-independence of fragments, asymmetry in the probability of loss/gain of fragments, and problems in distinguishing heterozygote from homozygote bands e.g. [23,25]. Despite these difficulties, several authors have used AFLPs to reveal phylogenetic relationships corroborated by analyses of other types of data, especially for species that have diverged recently or radiated within a short period of time e.g. [15,17,23,26].

Table 1 Occurrence of *Diospyros* species in different habitats in New Caledonia

		Substrate				
		Limestone	**Serpentine**	**Schist**	**Ultramafic rock**	**Volcanic rock**
Vegetation	Humid mountain forest			*D. parviflora*, *D. trisulca*		*D. flavocarpa*, *D. labillardierei*
	Humid low land forest				*D. glans*, *D. pancheri*, *D. parviflora*, *D. umbrosa*	*D. umbrosa*
	Mesophyll forest	*D. minimifolia*, *D. pustuala*, *D. tridentata*			*D. erudita*	*D. cherrieri*, *D. erudita*, *D. minimifolia*, *D. perplexa*, *D. pustulata*
	Maquis		*D. revolutissima*		*D. vieillardii*	
	Littoral forest	*D. calciphila*, *D. inexplorata*				*D. impolita*

Habitats are grouped according to vegetation type and substrate. Note that several species are co-occurring and that a few species are found in several habitats.

In this study we focus on this group of closely related species of *Diospyros* endemic to New Caledonia (Figure 1). Our aim was to clarify species boundaries as well as phylogenetic relationships between these New Caledonian *Diospyros* species. Integrated in a broader context, the outcome of our research should help us better understand the factors behind and mechanisms of speciation and radiation on islands.

Results

After excluding 186 replicates the final matrix used for analyses contained 192 individuals and 792 fragments. The AFLP profiles showed good reproducibility with a mean error-rate of 2.4% across all replicated samples. Because the focus of this study was on the phylogenetic relationships between species and species limits rather than intra-specific population genetics, we are presenting and discussing mostly the results of inter-specific relationships. We are presenting here only unrooted trees due to the low resolution of their backbone. We analysed the data using neighbour-joining (NJ) dendrograms and principal coordinate analysis (PCO) with different distance methods, and in both cases the Dice distance gave the highest resolution of relationships between species.

The NJ analysis resulted in a star-like dendrogram with a backbone of short branches lacking bootstrap support greater than 75%. All species except *D. minimifolia*, *D. parviflora* and *D. vieillardii* form single clusters in the NJ tree (Figure 2A). However, only eight (*D. calciphila*, *D. cherrieri*, *D. inexplorata*, *D. impolita*, *D. pustulata*, *D. trisulca*, *D. umbrosa* and *D. yahouensis*) of the 21 included species form clusters with bootstrap higher than 80%. The Bayesian inference (BI) produces a similar result. All species except *D. labillardierei*, *D. minimifolia*, *D. pancheri* and *D. parviflora* form single clusters in the BI tree (Figure 2B). Apart from *D. flavocarpa*, *D. revolutissima*, *D. tridentata* and *D. vieillardii*, all clustered species have high (> 0.95) posterior probabilities.

PCO separated accessions into two main groups (hereafter named "white" and "grey") that can be subdivided into six subgroups (Figure 3). Within the "white" group (defined in the STRUCTURE results below) subgroup one includes *D. vieillardii* (individuals indicated by squares in Figure 3), subgroup two *D. calciphila* (triangles) and subgroup three the rest of the individuals from this group (circles). In the "grey" group (more extensively described in the STRUCTURE results below) subgroup four included *D. flavocarpa*, *D. umbrosa* and *D. vieillardii* (indicated by squares in Figure 3), subgroup five *D. erudita* and *D. glans* (triangles) and subgroup six the remaining individuals (circles). A PCO of populations (not shown) based on the pair-wise F_{ST} distances obtained from the AMOVA resulted in similar groups and subgroups of populations as those obtained from the individual-based PCO. STRUCTURE analysis gave the highest value of ΔK for K = 2 plus few other sub-optimal K values (Figure 4A and B). However the latter contained clusters with negligible membership ("empty" clusters). Both K = 3 and K = 6 resulted in three visible clusters, with one cluster being only found in significantly admixed samples (Additional file 1). Visualisation of K =16 and K = 21 showed two clusters only and both analyses are highly similar to each other (Additional file 1). It has been argued the ad-hoc Evanno method [27] favours by default K =2 over K = 1 when searching for the correct number of clusters [28]. However, PCO separated individuals included in our analyses into two groups as well, and therefore we consider K = 2 as representative for our sample set. For K = 2, the allele-frequency divergence between the two groups was 0.0074. One group ("grey") includes the majority of accessions (Figure 4C). The other group ("white") includes *D. calciphila*, *D. labillardierei* (population 13 and accession BT179), *D. minimifolia* (majority of individuals), *D. pustulata*, *D.* sp. Pic N'ga, *D. tridentata* (accessions BT206 and BT207), *D. veillonii* (accession BT224) and *D. vieillardii* (population 37 [except accession BT017], population 39 [except accession BT100] and population 41). Seven individuals appear to be admixed (less than 90% identity with one of the groups); most of those are *D. vieillardii*. Several species (*D. labillardierei*, *D. minimifolia*, *D. tridentata*, *D. veillonii* and *D. vieillardii*) and even some populations comprise individuals belonging to each of the two groups.

In order to quantify the amount of genetic variation between species we have performed a non-hierarchical AMOVA with species assigned as "populations". This analysis showed as little as 30% of the variation to occur among the species. However, in the STRUCTURE, PCO, NJ and BI analyses several species seemed to be formed by genetically distinct populations assigned to different clusters and coming in distinct positions in the tree. To avoid mixing up of cryptic variation within a group, we run further AMOVAs with populations assigned as sample localities, despite the relatively low sample size per locality. Results of non-hierarchical AMOVA in this case indicate a higher level of differentiation between populations, resulting in an F_{ST} of 0.38. There was no visible difference in gene diversity between stands of co-occurring species and isolated populations. Several hierarchical AMOVAs (except one based on the STRUCTURE results) were not significantly more informative than the non-hierarchical AMOVAs (Table 2). Grouping populations according to geography or ecology, explains a surprisingly low amount of the variation (1.4 – 1.6%). Furthermore, allocating populations to the 21 included species assigns a relatively high percentage of variation at the between-species level (19.4%), but with a highly similar F_{ST} value to the non-hierarchical AMOVA results. When higher-level groupings paralleled the STRUCTURE results, we obtained the

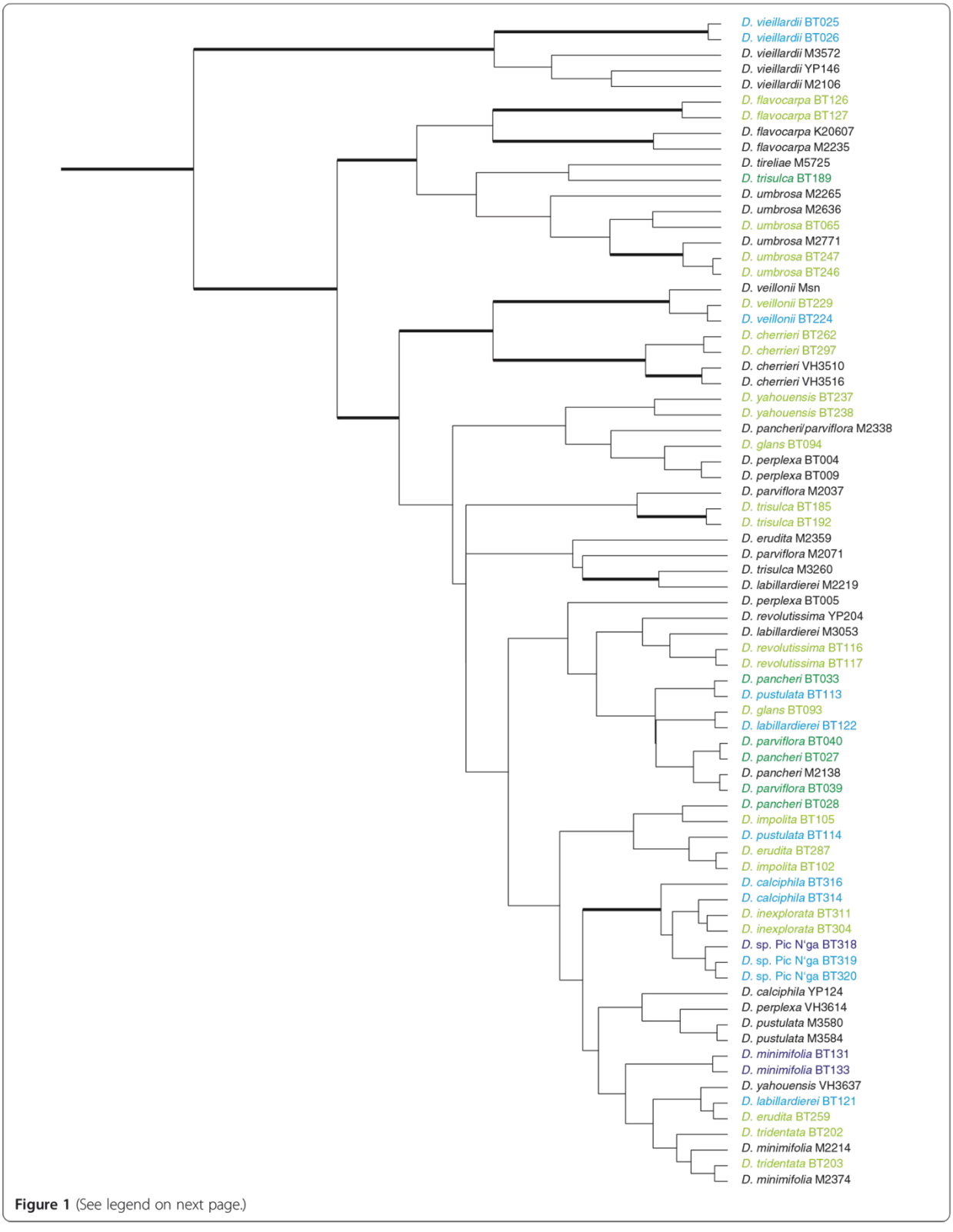


Figure 1 (See legend on next page.)

(See figure on previous page.)
Figure 1 Bayesian maximum clade credibility tree of New Caledonian *Diospyros* species based on plastid and nuclear DNA data (taken from Turner et al. [13]). Bold branches are supported (> 70% bootstrap and Bayesian posterior probability). Accessions in blue correspond to the white group found in STRUCTURE and PCO, green ones to the grey group (light blue/green accessions included in current data set, dark blue/green accessions failed in current analysis but colour indicates the group to which they most probably belong), accessions in black are not included in the present study.

highest F_{ST} value (0.4), albeit the percentage of variation between the two clusters as defined by STRUCTURE was only 9.5%, lower than the percentage of differentiation shown between species. Removing seven admixed samples (less than 90% membership form each of the two groups based on the STRUCTURE results) from the AMOVA gave nearly the same results as the analysis including them (Table 2).

The average gene diversity over loci within populations ranged from 0.03 in *D. erudita* (population 4) to 0.12 in *D. parviflora* (population 22). Contrary to predictions, the highest number of polymorphic sites, pair-wise differences and average gene diversity were not found in the admixed populations (according to the STRUCTURE results) but in *D. parviflora* (for details see Additional file 2).

Discussion

"Explosive" radiations featuring rapid opportunistic morphological and ecological diversification are phenomena previously reported for some islands (e.g. [29] and references therein). Extreme ancestral bottlenecks, together with

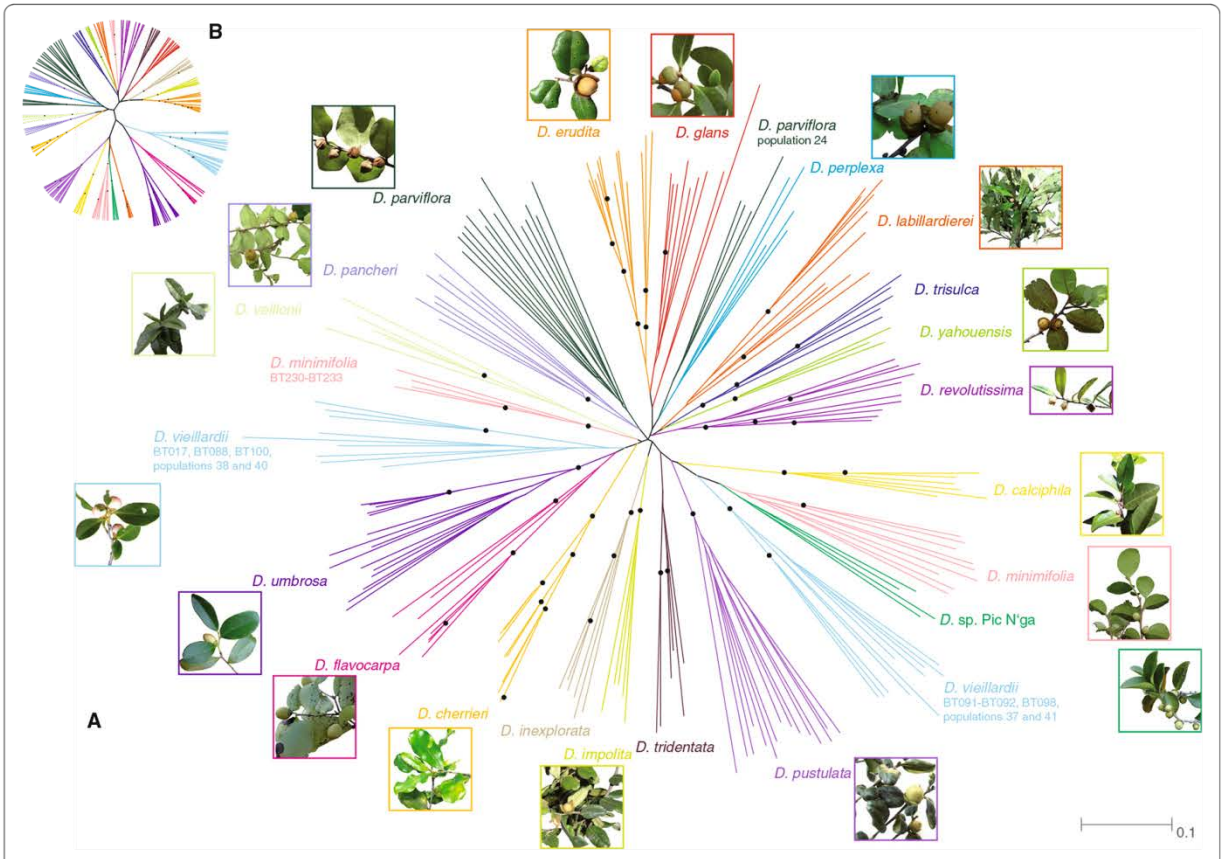


Figure 2 Phylogenetic dendrograms inferred from the data collected in this study. Each species is shown in a different colour. Colours were selected randomly and do not indicate any grouping. **A**: Neighbour joining dendrogram based on Dice distances. Black dots indicate nodes with > 80% bootstrap support. **B**: Bayesian maximum clade credibility dendrogram. Black dots indicate nodes with > 0.95 Bayesian posterior probability. Picture credits: *D. calciphila*: H. Benoît, www.endemia.nc; *D. cherrieri*: C. Chambrey; *D. erudita*, *D. pancheri*, *D. pustulata*, *D. umbrosa*, *D. vieillardii*: D. & I. Létocart, www.endemia.nc; *D. flavocarpa*, *D. minimifolia*, *D. revolutissima*, *D.* sp. Pic N'ga: J. Munzinger; *D. glans*, *D. parviflora*: J.-L. Ruiz, www.endemia.nc; *D. impolita*: J. Barrault, www.endemia.nc; *D. labillardierei*: B. Turner; *D. perplexa*, *D. yahouensis*: V. Hequet; *D. veillonii*: R. Amice, www.endemia.nc.

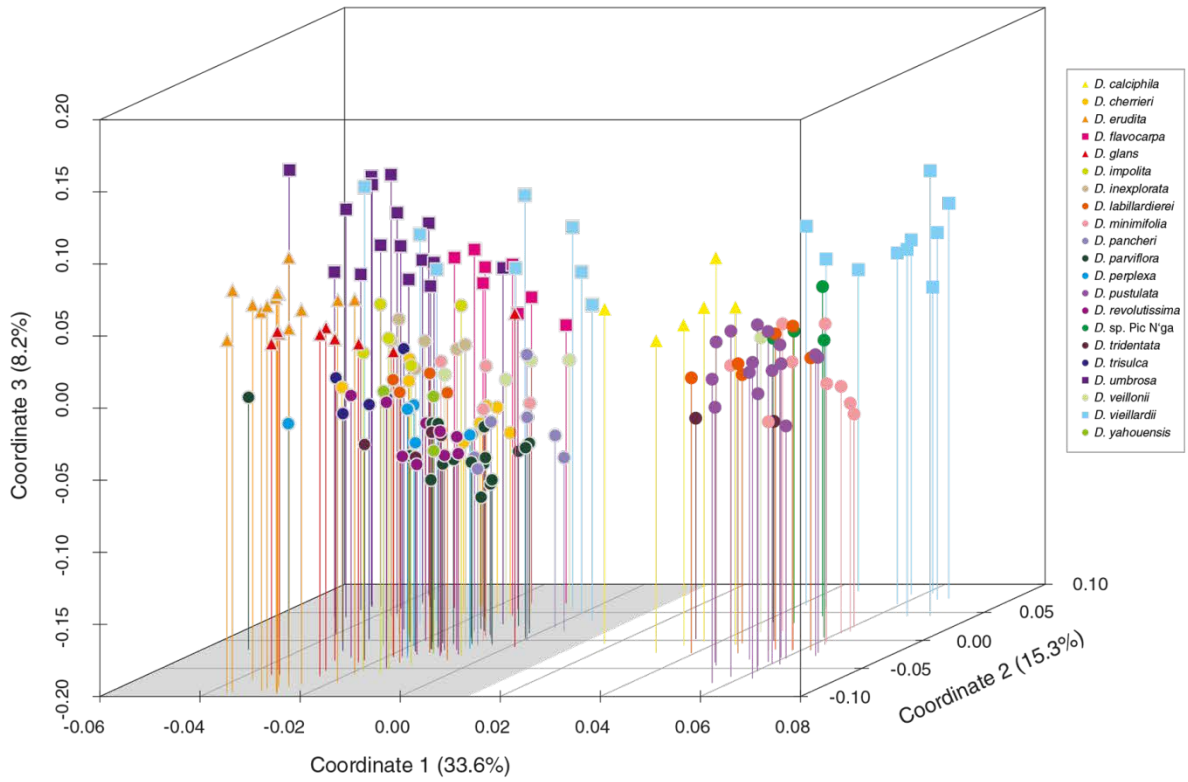


Figure 3 PCO of individual accessions based on Dice distances. Shading of the base-grid of the figure marks the two main groups inferred by STRUCTURE analysis – white and grey. Each species is shown in a different colour. Colours were selected randomly and do not indicate any grouping.

on-going hybridization and incomplete lineage sorting, can prevent phylogenetic reconstruction in cases of island radiations if they have been recent and produced many species [30]. However, a good understanding of phylogenetic relationships within radiating groups is key for further evolutionary studies into mechanisms and whether change is adaptive, due to drift in small populations or other phenomena [29].

For the endemic New Caledonian *Diospyros* species, previous studies, based on multiple plastid [12] and low-copy nuclear [13] markers, showed 21 species to be closely related (Figure 1) and were not able to clearly resolve phylogenetic relationships among them. In the combined data set (plastid and nuclear markers; [13]) only seven of the 21 species included were found to form highly supported groups of accessions from single species. Individuals belonging to each of the remaining 14 species failed to cluster according to their taxonomic circumscription. Dating analysis based on plastid and low-copy nuclear markers showed that the common ancestor of this clade of endemic New Caledonian *Diospyros* species has arrived in New Caledonia around nine million years ago [13]. *Diospyros vieillardii* has been shown to be sister to the rest of this endemic clade and separated from the rest of the species around 7.2 million years ago.

Results of the current study using genome-wide AFLP markers reveal that most species form unique groups paralleling recognised species. Around one-third (eight species, NJ dendrogram, Figure 2A) and one-half (11 species, Bayesian tree, Figure 2B) of the species, are genetically distinct with high support (Figure 2). However, the overall AFLP results prove unable to clearly resolve the backbone of trees, similar to previous results obtained from analyses of DNA sequence data [13]. Intra-specific variation was greater (~80%) than that found at inter-specific level (~20%). This low ratio of among- versus within-species divergence in the context of considerable morphological and ecological divergence is indicative of a recent diversification [22]. Such a process can explain why we were able to get clear species boundaries for most species but were unable to clearly resolve phylogenetic relationships between them.

Two species that did not form well-defined clades (*D. minimifolia* and *D. parviflora*) were previously considered by White [31] to show variability in leaf morphology that may indicate that they are in fact a collection of

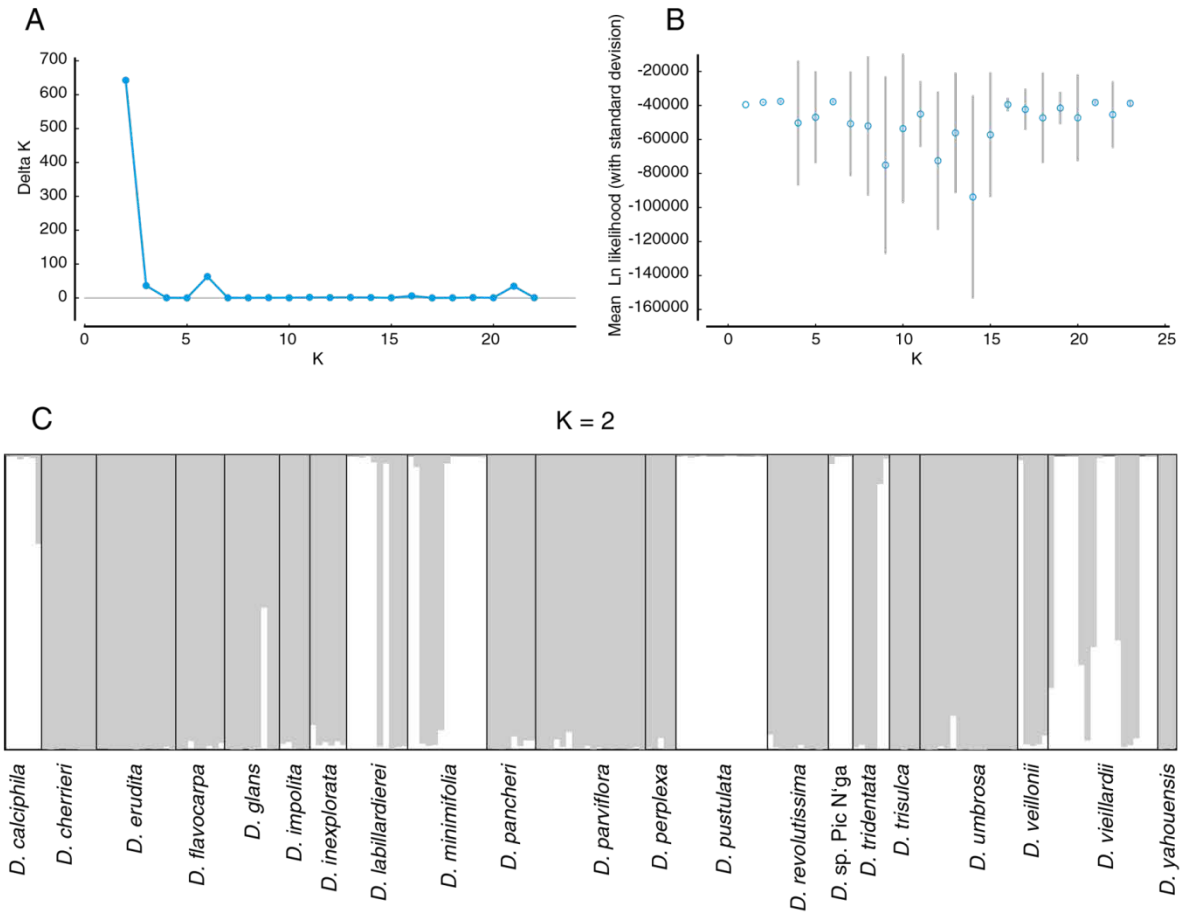


Figure 4 Results of the STRUCTURE analysis. A) Delta K values of the K values tested. **B)** Mean Ln likelihood of K values tested. **C)** Clustering of K = 2. The two groups are marked in different shades (white and grey).

several species. For *D. minimifolia* White [31] mentioned that the type population (close to population 15 of this study) has smaller leaves compared to other populations of this species. In our results this population clusters together with the majority of the *D. minimifolia* accessions; the population that is separated from the rest (population 16) is from Gaji. According to White [31] *D. parviflora* is a wide-spread species, showing considerable variability of leaf morphology even within populations, making it impossible to differentiate these into different species. Our results show all accessions of *D. parviflora*, except those from Plateau de Tango (population 24), to form a group. All included accessions from *D. parviflora* are from ultramafic localities.

To our surprise, the AFLP results do not show any significant grouping according to ecological (edaphic, climatic, elevational), geographical or morphological factors (Additional file 3). The two weakly differentiated groups revealed by STRUCTURE and PCO also do not correspond to any conspicuous phenotypic characteristics. The allele-frequency divergence between the two groups found by STRUCTURE is low, which explains why we did not observe the two groups in the Bayesian and NJ tree-building results. Taken together, these results indicate that positive selection has perhaps acted on few genomic regions [32] and has resulted in phenotypic diversification of New Caledonian *Diospyros*. Variation in copy number of specific genomic regions may be an additional aspect of molecular variation that, although invisible to AFLP markers, could form the basis of adaptation to different environmental conditions [33].

The individuals of *D. vieillardii*, *D. umbrosa* and *D. flavocarpa* form a minimally isolated group (squares in the grey group) in the PCO (Figure 3). Previous phylogenetic analyses (Figure 1) showed these three species to be sister to the rest of the taxa. Due to its morphological and ecological features *D.* sp. Pic N'ga from Île des Pins could be a hybrid between *D. calciphila* and *D. vieillardii*, but *D. vieillardii* is now not known from this island. In PCO, individuals of this putative species are located between

Table 2 Results of different AMOVAs conducted

	Percentage of variation				
Analysis	**Among groups**	**Among populations within groups**	**Within groups**	F_{ST}	***p* value**
Non-hierarchical	-	38.16	61.84	0.382	0.00
Species-wise	19.43	19.15	61.42	0.386	0.00
STRUCTURE	9.46	33.22	57.32	0.427	0.00
STRUCTURE no admixed	9.93	33.39	56.68	0.433	0.00
Geographic	1.58	36.97	61.45	0.385	0.00
Water	1.37	37.20	61.43	0.386	0.00
Soil	1.54	36.92	61.54	0.385	0.00

In the non-hierarchical analysis, no grouping was applied. In the species-wise analysis, samples were grouped according to taxonomic features (21 groups corresponding to the 21 species included).
In the STRUCTURE analysis, samples were grouped according to the results of STRUCTURE analysis (two groups corresponding to the two groups – white and grey – inferred by STRUCTURE); in the analysis without admixed samples seven samples with less than 90% identity to one of the two groups were removed. In the geography analysis the samples were grouped according to their origin (three areas – north, middle and south – of New Caledonia). The analysis based on water availability was structured into two groups – humid and dry. In the soil-type based analysis, species were grouped according to the substrate on which they were found (five groups – limestone, serpentine, schist, ultramafic rock and volcanic rock).

individuals of *D. calciphila* and *D. vieillardii* (Figure 3). The split between the two groups observed (Figures 3 and 4) could be relatively old, separating two lineages that developed in isolated regions. For instance, dry periods of the Pleistocene caused aridification in many areas, and some vegetation types persisted only in local refugia e.g. [34-36]. After climatic conditions became more favourable, the two groups probably expanded rapidly into newly suitable habitats where they overlapped; the time scale of these fluctuations (ca. 0.02 – 0.1 myr; [37]) was probably not enough to allow woody species with long generation time such as *Diospyros* to diverge and become permanently reproductively isolated [22]. There are a few admixed individuals in the STRUCTURE analysis (Figure 4), which implies that hybridization might have played a role in evolution of this group.

Accelerated rates of evolution at few genes as a result of positive selection could have resulted in the morphological and ecological diversification apparent today in this group of New Caledonian *Diospyros* species. Furthermore, in addition to retention of ancestral polymorphisms, frequent gene flow could have acted against genome-wide genetic differentiation between the species. Barriers to gene flow between these species may be highly porous, with only few genes responsible for ecological and morphological adaptations evolving on distinct trajectories under strong selection, which leaves the rest of their genomes open to gene flow [38]. Finding these few genes with AFLP is realistically improbable because they are a miniscule component in comparison the rest of these genomes. In the case of a recent and rapid radiation in plants, it could be argued that the bulk of regions sampled by AFLP have not evolved quickly enough to accumulate substitutions that could indicate species relationships. Our results are similar to those found in various other island genera (e.g. *Araucaria* in New Caledonia, [22]; *Ourisia* in New Zealand, [39]).

Diospyros vieillardii, which is sister to the rest of the taxa belonging to this group of New Caledonian endemics [12,13], is confined to ultramafic soils, which supports the hypothesis of this being an exaptation of the progenitor of this New Caledonian *Diospyros* clade to ultramafic soils when the whole island was still covered by heavy-metal-rich substrates; similar findings have been made in other plant groups in New Caledonia e.g. [9]. Later, erosion reduced the extent of this geological layer to one third of the island [7], and existing species began to move onto other substrates where they subsequently diverged, forming distinct species. Such observations have been made in various other New Caledonian groups (e.g. *Araucaria*, [22]; *Spiraeanthemum*, [35]; *Codia*, [40]). A few studies have examined the adaptive basis and processes involved in rapid radiations in New Caledonia e.g. [41] and Hawai'i (e.g. lobeliads, [42]; silverswords, [43]). Linking ecological parameters and/or phenotypic traits associated with speciation has to be done with caution because range alterations, subsequent evolution, and species extinctions might have erased initial signals found in only a few genes. Therefore, the associations observed today may be misleading, and the specific conditions/traits that were indeed linked to speciation, if any, may no longer be present [44].

Further work involving common garden experiments would provide insights into the effect of environmental conditions on morphological traits and therefore plasticity of genomes of the New Caledonian *Diospyros* species. Unfortunately, such experiments are time and cost intensive. It is difficult to obtain ripe fruits of all *Diospyros* species, and in addition it is difficult to germinate and grow them, which is a crucial aspect of conducting such experiments. Reciprocal transplantation of seedlings across environments are of course more easily conducted than common garden experiments, but they

are still time consuming and costly; in addition species adapted to one soil type often will not survive when transplanted to other soil types.

Conclusions

Although New Caledonian *Diospyros* are morphologically and ecologically diverse, they show little genetic divergence (based on DNA sequences and AFLP data). In this case of the endemic clade of New Caledonian *Diospyros*, AFLP data did not provide enough information to resolve phylogenetic relationships between the species, but it was sensitive enough for testing for the presence of genetic species boundaries. However, the AFLP results exhibit a good correlation with morphology-based species concepts. Further studies of this New Caledonian *Diospyros* group with deeper sampling of the genome using next generation sequencing methods are needed to get a clearer picture of the processes that formed this group.

Methods

Material

Material from New Caledonian *Diospyros* species was collected on the main island (Grande Terre) and on a smaller island, Île des Pins. When possible, we collected five individuals per population. Collecting population samples from tropical trees/shrubs is not always easy because the trees can be tall (and leaves therefore out of reach) and individuals are often far from each other. Collecting ten individuals in an area of ten square meters also does not make much sense for a study like this because these individuals are probably offspring from the same mother plant. As the focus of the present study is on the phylogenetic relationships between the species and not on population genetics within species, the authors consider the small size of the samples we collected to be sufficient. For widespread species, we collected populations throughout their range. For distribution of sampling sites, see Figure 5. From samples where fertile material was available, a voucher was made with several duplicates sent to the herbaria at Noumea (NOU), University of Montpellier II (MPU) and the University of Vienna (WU). When sterile, one voucher per population was taken; this was compared to already existing collections in Noumea Herbarium (NOU) from the same location and referred to that species if similar. In total we included in the present study 231 individuals of New Caledonian *Diospyros* species, which correspond to 20 identified and one unidentified species (due to absence of diagnostic reproductive organs at the time of collection), giving 47 populations in total. Details of the 192 individuals (43 populations) for which we were able to get useable results are given in Table 3. Silica-gel-dried material was used for DNA extraction.

DNA extraction

For DNA extraction, a modified sorbitol/high-salt CTAB method [46] was used (for details see [13]).

AFLP

Preparation and amplification of fragments followed the protocol of Vos et al. [14] with some modifications. Restriction of genomic DNA with two restriction enzymes and ligation of double-stranded adaptors to the resulting restricted fragments were performed in one step in a thermal cycler (Veriti, AB, Life Technologies; 37°C for 2 h followed by a 30 min hold at 17°C). Reactions comprised 1.1 μL 10x T4 DNA ligase buffer (Promega), 1.1 μL 0.5 M NaCl, 0.55 μL BSA (1 mg/ mL; New England BioLabs), 50 μM MseI adaptors (genXpress), 5 μM EcoRI adaptors (genXpress), 1 U MseI restriction endonuclease (New England BioLabs), 5 U EcoRI restriction endonuclease (New England BioLabs), 1 U T4 DNA ligase (Promega), and 0.5 μg DNA and were made up to a total volume of 11 μL with water. Ligated DNA fragments were diluted 10-fold with sterile water. Preselective amplification reactions contained 1.14 μL 10x RedTaq PCR reaction buffer (Sigma), 0.2 U RedTaq DNA polymerase (Sigma), 0.22 μL dNTPs (10 mM; AB, Life Technologies), 0.58 μL preselective primer pairs (EcoRI-A and MseI-CT, each 5 μM; Sigma), and 2 μL diluted restriction-ligation product, and were brought with water to a total volume of 10 μL. Amplification was carried out in the same machine used for restriction-ligation with the following profile: 2 min at 72°C, 20 cycles of 10 sec denaturing at 94°C, 30 sec annealing at 56°C, 2 min extension at 72°C, and a final extension step for 30 min at 60°C. The preselective PCR products were diluted 10-fold with sterile water. Reactions for selective amplification contained 0.5 μL 10x RedTaq PCR reaction buffer (Sigma), 0.1 U RedTaq DNA polymerase (Sigma), 0.11 μL dNTPs (10 mM; AB, Life Technologies), 0.27 μL MseI-primer (5 μM; Sigma), 0.27 μL EcoRI-primer (1 μM; Sigma), and 1 μL diluted preselective amplification product and were brought to a total volume of 5 μL with water. They were carried out in a GeneAmp PCR System 9700 (AB, Life Technologies) with the following profile: 1 min at 94°C, 9 cycles of 1 sec at 94°C, 30 sec at 65-57°C (reducing the temperature at 1°C per cycle), 2 min at 72°C, 25 cycles of 1 sec at 94°C, 30 sec at 56°C, 2 min at 72°C and a final extension for 30 min at 60°C. The selective PCR products were purified using Sephadex G-50 Superfine (GE Healthcare Bio-Sciences) applied to a MultiScreen-HV 96-Well Plate (Millipore) in three steps of 200 μL each and settled at 750 g (1, 1 and 5 min, respectively). The same speed was used for centrifugation of the samples (5 μL of each selective PCR product), again for 5 min. Two microliters of the eluate were combined with 10 μL HiDi and 0.1 μL GeneScan 500 ROX (AB, Life Technologies) and denatured for 3 min at 95°C

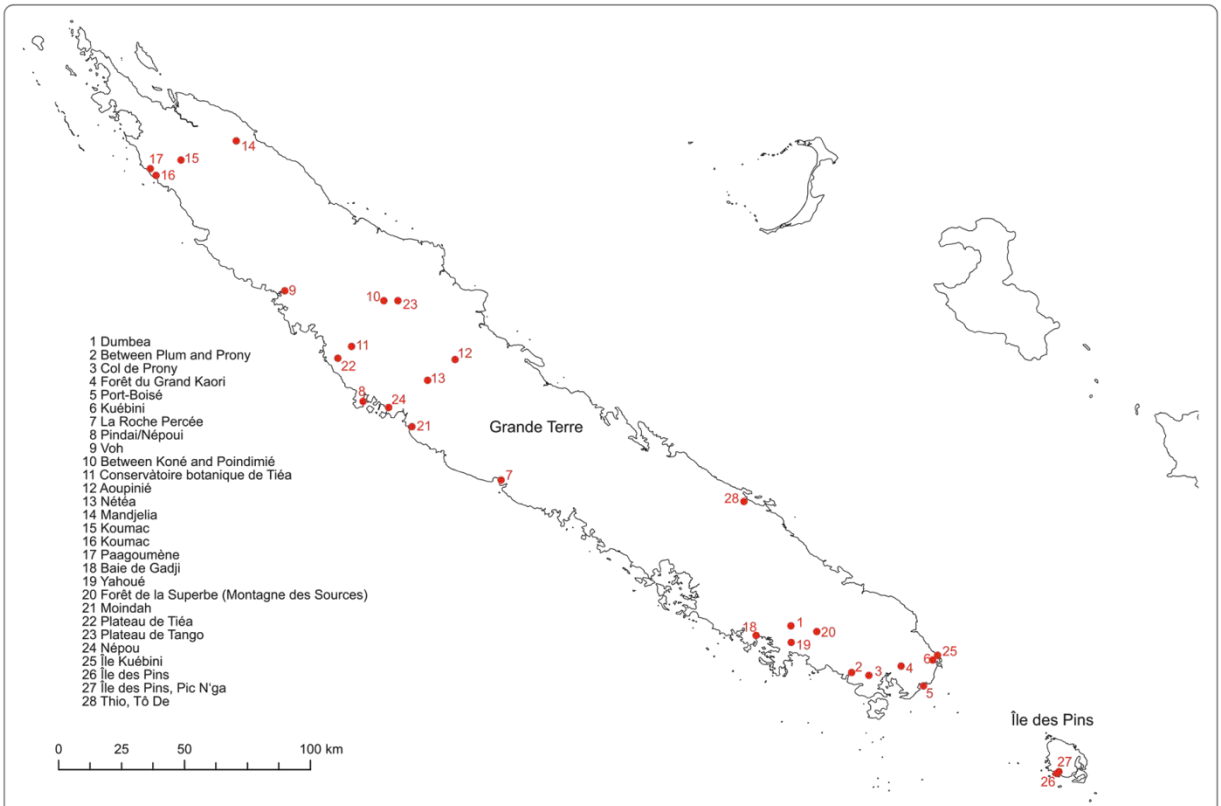


Figure 5 Map with sampling localities. Dots indicate sampling sites; the numbers associated with each dot refer to the list of sampling sites on this figure. Those numbers are used throughout the present study to characterize sampling sites.

before running them on a capillary sequencer (3130xl Genetic Analyzer, AB, Life Technologies).

The selective primer pairs (6Fam-EcoRI-AGC/MseI-CTGA, Vic-EcoRI-ATG/MseI-CTCG and Ned-EcoRI-ATC/MseI-CTGA) were chosen because they generated clear and not too many bands (thus decreasing the risk of fragments co-migrating by chance), with sufficient variability in preliminary tests. Although the genome size of the New Caledonian *Diospyros* species (1C-value: 1.5 – 2.3 pg; [13]) is smaller than the mean 1C-value of eudicots (2.7 pg, [47]), we found the AFLP profiles generated with Msel primers with four selective bases much clearer than those obtained from primers with just three selective bases.

Reproducibility was checked by repeating ca. 80% of the samples. This high number of repetitions was necessary because of initial difficulties with fragment sizing.

Scoring and phylogenetic analysis

Sizing and scoring of the data was performed with GeneMarker v2.2.0 (SoftGenetics). After pre-analysis using default settings, sizing profiles of all samples were checked and where necessary manually corrected. Most of these corrections concerned one of the following peaks of the size standard: 35 bp, 50 bp and 139 bp. These peaks were often not correctly recognized by the GeneMarker program. High-quality sizing-profiles (score > 90) were obtained for all samples. A panel of scorable fragments was established for each primer combination, and fragments between 65 – 510 bp were scored. The relative fluorescent unit (RFU) threshold was set at 40. Automatic scoring was conducted using Local Southern peak call, peak saturation, base line subtraction, spike removal, pull up correction, and a stutter peak filter of 5% (as described in [48]). The results were exported as presence/absence matrix. The outcome of the automatic scoring was manually checked and corrected for errors. These errors mostly concerned peaks for which shape was atypical. In total 486 samples corresponding to 231 individuals were scored. From 186 individuals replicate samples were performed (between two and five replicates per individual). The differences between the different samples (replicates) were counted and divided by the total number of phenotypic comparisons to get the error rate (calculated according to Bonin et al. [49]). After initial analysis (neighbour-joining, NJ) of the complete data set, replicates of samples and obviously failed samples were excluded from further analyses. As replicated samples of the corresponding individuals mostly clustered together,

Table 3 Table of accessions; showing all individuals used in this study

Taxon	Sample ID	Population	Sampling location	Voucher
D. calciphila F.White	BT312-BT317	1	26, littoral forest	JM6650, JM6653 (MPU, NOU, P)
D. cherrieri F.White	BT262, BT276-BT278	2	21, dry forest	NOU079551, WU062860 NOU054492, NOU054008
D. cherrieri	BT293-BT297	3	24, dry forest	NOU079547
D. erudita F.White	BT259-BT261, BT273-BT275	4	21, dry forest	WU062855, NOU079545 NOU079544, WU062870 NOU054010, WU062856 NOU054011, WU062857
D. erudita	BT280-BT285, BT287	5	22, dry forest	WU062858, Chambrey & Turner 20 (NOU)
D. flavocarpa (Vieill. ex P.Parm.) F.White	BT126-BT130	6	10, humid mountain forest	JM6625 (NOU)
D. flavocarpa	BT155, BT158-BT159	7	12, dense humid mountain forest	JM6632 (NOU)
D. glans F.White	BT020-BT022	8	1, forest near river	NOU053705, NOU030755, WU062846
D. glans	BT075	9	5, dense forest near road	NOU000819
D. glans	BT082, BT084, BT087, BT093-BT094	10	6, forest near river	NOU022860
D. impolita F.White	BT101-BT105	11	7, mesophyll forest near beach	NOU019538
D. inexplorata F.White	BT304, BT307-BT311	12	25, littoral forest	NOU005818
D. labillardierei F.White	BT121-BT125	13	10, river edge in mountain forest	JM6624 (NOU)
D. labillardierei	BT178-BT182	14	13, river edge	(NOU031346)
D. minimifolia F.White	BT134-BT135	15	11, dry forest	NOU019556
D. minimifolia	BT230-BT234	16	18, mesophyll forest near beach	NOU019554
D. minimifolia	BT263-BT264, BT266-267, BT269-BT270	17	21, dry forest	NOU079549, WU062872 NOU054493
D. pancheri Kosterm.	BT029-BT031, BT035	18	2, forest near road	JM6619, JM6620 (NOU)
D. pancheri	BT076-BT079	19	5, dense forest near road	
D. parviflora (Schltr.) Bakh.	BT042	20	3, wet forest	
D. parviflora	BT059, BT062-BT063, BT068	21	4, wet dense forest	NOU006656
D. parviflora	BT080, BT085, BT089-BT090	22	6, forest near river	JM6622 (NOU)
D. parviflora	BT248-BT250, BT252-BT253	23	20, humid forest at low elevation	tree no. 23109
D. parviflora	BT289-BT292	24	23, mountain forest	NOU079550
D. perplexa F.White	BT147-BT151	25	11, forest near river	JM6630 (NOU)
D. pustulata F.White	BT111-BT114	26	8, dry forest	
D. pustulata	BT136-BT140	27	11, dry forest	JM6629 (NOU)

Table 3 Table of accessions; showing all individuals used in this study *(Continued)*

D. pustulata	BT257-BT258, BT265, BT268, BT271-BT272	28	21, dry forest	NOU079548, WU062871 NOU053999
D. revolutissima F.White	BT116-BT120	29	9, maquis	NOU023189
D. revolutissima	BT218-BT222	30	17, maquis	JM6640 (NOU)
D. tridentata F.White	BT202-BT207	31	15, dry forest at low elevation	JM6639 (NOU)
D. trisulca F.White	BT185, BT192, BT197, BT199-BT201	32	14, mountain forest	NOU031344, JM6637 (NOU)
D. umbrosa F.White	BT061, BT065-BT066, BT071, BT073	33	4, wet dense forest	
D. umbrosa	BT170-BT171, BT175-BT177	34	13, dense humid forest	JM6635 (NOU)
D. umbrosa	BT246-BT247, BT251, BT254, BT256	35	20, humid forest at low elevation	NOU023234
D. veillonii F.White	BT224, BT226-BT229	36	18, mesophyll forest near beach	NOU019582
D. vieillardii (Hiern) Kosterm.	BT017, BT023-BT026	37	1, forest near river	JM6618 (NOU)
D. vieillardii	BT055, BT057-BT058	38	4, dry open forest	
D. vieillardii	BT088, BT091-BT092, BT098, BT100	39	6, forest near river	
D. vieillardii	BT215-BT217	40	16, maquis	NOU023242
D. vieillardii	BT324-BT325, BT328	41	28, forest near river	
D. yahouensis (Schltr.) Kosterm.	BT237-BT239	42	19, mesophyll forest	P00057340
D. sp. Pic N'ga	BT319, BT321-BT323	43	27, maquis	JM6065 (NOU)

The numbers of sampling localities are the same as in Figure 2. Voucher-Codes: JMXXXX: collection number J. Munzinger; Tree N° XXX: Tree of New Caledonian Plant Inventory and Permanent Plot Network (NC-PIPPN, [45]); NOUXXXXXXX: Herbarium accession number of Noumea herbarium (NOU); WUXXXXXX: Herbarium accession number of the Herbarium of the University Vienna (WU); P: Herbarium of the Natural History Museum Paris; MPU: Herbarium of the University of Montpellier II.

selection of samples from each individual for further analyses was random and not according to any pattern or protocol. For the final analyses we ended up with 192 individuals.

All three primer-combinations were combined in a single matrix and analysed together. Different distance measures were tested for their power to resolve relationships with our data set. Distance matrixes were calculated in PAUP* v4b10 ([50]; Nei-Li distance) and SplitsTree v4.12.6 ([51]; uncorrected P, Dice, corrected and uncorrected Hamming). Phylogenetic relationships based on previously mentioned distance matrices were reconstructed using SplitsTree v4.12.6 [51] to create unrooted NJ dendrograms. To assess robustness of branches NJ-bootstrap (NJ-BS) analyses were performed using SplitsTree v4.12.6 [51] and PAUP* v4b10 [50]. Bayesian inference (BI) was conducted with BEAST v1.7.5 [52], with two runs each 20 million generations, sampling every 1,000th generation and removal of the first 30% of trees as burn in.

To visualise the pattern of genetic clustering of individuals and populations, we plotted principal coordinate analysis (PCO) using the R-package scatterplot3d [53] based on an individual Dice distance matrix, and respectively, on AMOVA-derived pair-wise F_{ST} distances calculated with Arlequin v3.5.1.2 [54]. To investigate further significant groupings of the included individuals we used the program STRUCTURE v2.3.3 [55,56] on the Bioportal computing cluster of the University Oslo [57]. We ran STRUCTURE for K = 1–23 with 10 replicates each and a model based on admixture and independent allelic frequencies, without taking into account information regarding sampling localities. Each run had 3 million iterations with 10% additional burn in. The calculation of deltaK (ΔK; [27]) and preparation of the input file for Clumpp was done with Harvester [58]. Production of a combined file from the ten replicates of the best K was perfomed using Clumpp v1.1.2 [59] with the full search algorithm. The graphical representation of STRUCTURE results was prepared with Distruct v1.1 [60].

Both non-hierarchical and hierarchical analyses of molecular variance (AMOVA) and calculations of population statistics were conducted using Arlequin v3.5.1.2 [54]. For hierarchical AMOVAs groups have been defined based on different possible clusterings (Additional file 4) according to STRUCTURE results, taxonomy, distribution patterns and ecological traits.

Availability of supporting data

AFLP presence/absence matrix and phylogenetic analyses are deposited in treeBASE under study 14798 (http://purl.org/phylo/treebase/phylows/study/TB2:S14798).

Additional files

Additional file 1: STRUCTURE results of suboptimal K values (3, 6, 16 and 21) in comparison with K =2. Delta K likelihoods are given for each K.

Additional file 2: Table showing the population statistics inferred from non-hierarchical AMOVA based on STRUCTURE results. Populations marked bold differ in this analysis from the general population grouping given in Table 3.

Additional file 3: Figure of the neighbour joining dendrogram coloured according to soil type (colour of the branches) and water availability (colour of taxa names). This dendrogram is the same as Figure 3A, but coloured according to ecological features.

Additional file 4: Table giving the details of the different AMOVAs conducted. The numbers in the populations column are the same as given in Table 3, respectively, in Additional file 1 for the STRUCTURE based AMOVA.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

BT carried out the acquisition and analysis of the data, drafted the manuscript and assisted collecting the plant material. OP helped with data analysis. JM collect and identified the plant material and helped to interpret the results. Previous studies of SD helped to design this project. MWC helped to design this project and gave linguistic support to the manuscript. RS, designed and coordinated the study and helped to draft the manuscript. All authors read, commented, corrected and approved the final manuscript.

Acknowledgements

This work was funded by a grant from the Austrian Science Fund (FWF, Project-Number: P 22159-B16) awarded to R. Samuel. The authors thank the team of the Department of Systematic and Evolutionary Botany of the University of Vienna as well as the team of the IRD AMAP Noumea for support with this study. Special thank goes to V Klenja for the lab work. Thanks to the following persons for their help with lab work, field work and ideas to improve this work: F Balao Robles, J-P Butin, C Chambrey, G Dagostini, E Grasserbauer, V Hequet, G Kohl, D & I Létocart, W Nigote and H Vandrot. Specimens are deposited in the herbaria of Noumea (NOU), University of Montpellier II (MPU) and the University of Vienna (WU).

Author details

[1]Department of Systematic and Evolutionary Botany, Faculty of Life Sciences, University Vienna, Rennweg 14, 1030 Wien, Austria. [2]IRD, UMR AMAP, TA A51/PS2, 34398 Montpellier Cedex 5, France. [3]Department of Forest Biology, Faculty of Forestry, Kasetsart University, Bangkok, Thailand. [4]Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3DS, UK. [5]School of Plant Biology, The University of Western Australia, Crawley, WA 6009, Australia.

Received: 14 August 2013 Accepted: 9 December 2013
Published: 12 December 2013

References

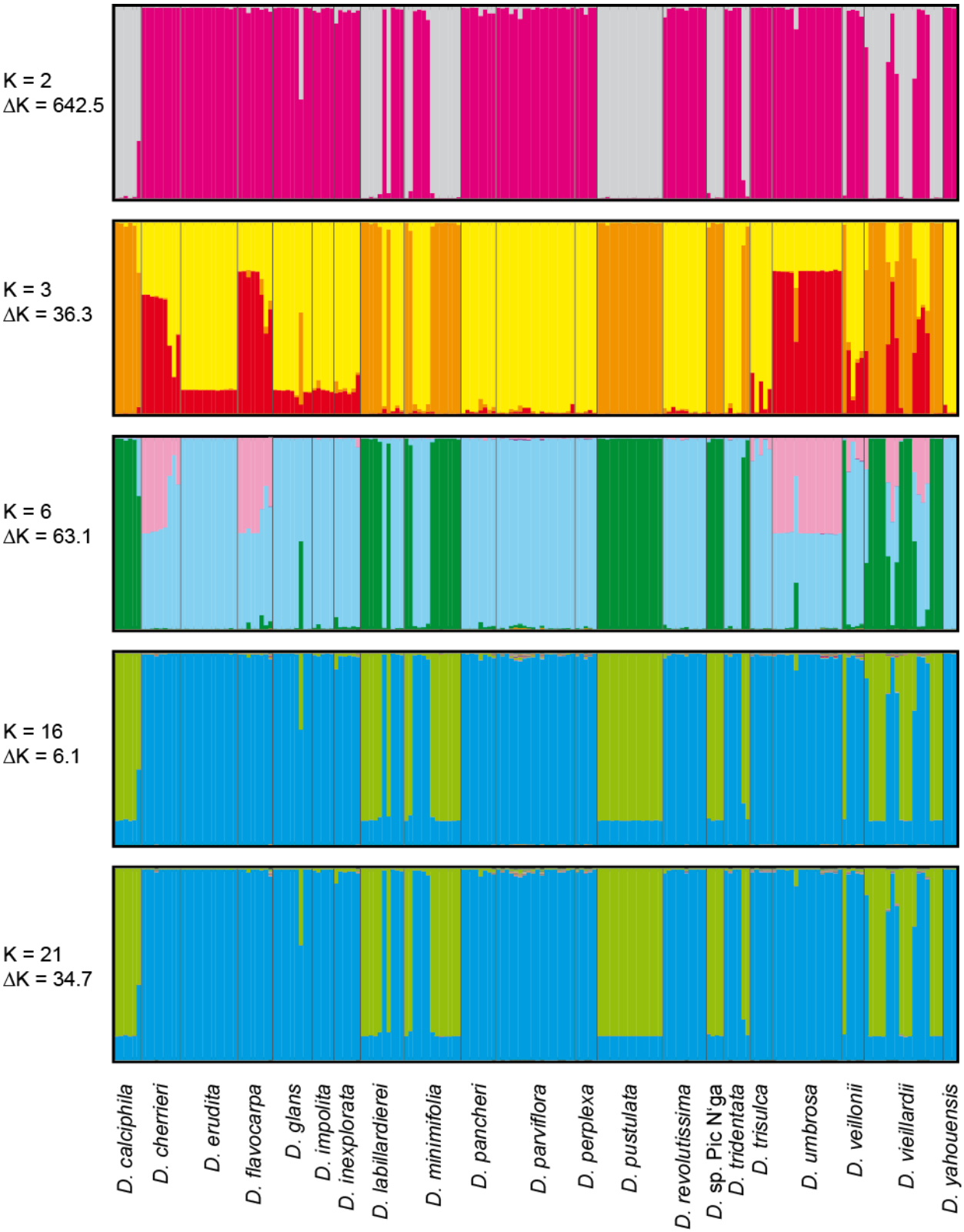
1. Givnish TJ, Millam KC, Mast AR, Paterson TB, Theim TJ, Hipp AL, Henss JM, Smith JF, Wood KR, Sytsma KJ: **Origin, adaptive radiation and diversification of the Hawaiian lobeliads (Asterales: Campanulaceae).** *Proc R Soc B* 2009, **276:**407–146.
2. Knope ML, Morden CW, Funk VA, Fukami T: **Area and the rapid radiation of Hawaiian *Bidens* (Asteraceae).** *J Biogeogr* 2012, **39:**1206–1216.
3. Mittermeier RA, Gil PR, Hoffmann M, Pilgrim J, Brooks T, Mittermeier CG, Lamoreux J, da Fonseca GAB: *Hotspots revisited: Earth's biologically richest and most endangered terrestrial ecoregions.* Mexico City: CEMEX; 2004.
4. Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J: **Biodiversity hotspots for conservation priorities.** *Nature* 2000, **403:**853–858.
5. Morat P, Jaffré T, Tronchet F, Munzinger J, Pillon Y, Veillon J-M, Chalopin M: **Le référentiel taxonomique Florical et les caractéristiques de la flore vasculaire indigène de la Nouvelle-Calédonie.** *Adansonia* 2012, **34:**177–219.

6. Lowry PP II: **Diversity, endemism and extinction in the flora of New Caledonia: a review.** In *Rare, threatened, and endangered floras of Asia and the Pacific rim*. Edited by Peng CF, Lowry PP II. Taiwan: Institute of Botany, Taipei; 1998:181–206.
7. Pelletier B, Payri C, Richer De Forges B: **Geology of the New Caledonia region and its implications for the study of the New Caledonian biodiversity.** In *Compendium of marine species from New Caledonia, Documents Scientifiques et Techniques II4*. New Caledonia: Institut de Recherche pour le Développement Nouméa; 2006:17–30.
8. Maurizot P, Vendé-Leclerc M: **New Caledonia geological map, scale 1/500000.** *Direction de l'Industrie, des Mines et de l'Energie - Service de la Géologie de Nouvelle-Calédonie, Bureau de Recherches Géologiques et Minières* 2009.
9. Pillon Y, Munzinger J, Amir H, Lebrun M: **Ultramafic soils and species sorting in the flora of New Caledonia.** *J Ecol* 2010, **98**:1108–1116.
10. Jaffré T, Rigault F, Munzinger J: **La végétation.** In *Atlas de la Nouvelle-Calédonie*. Edited by Bonvallot J, Gay J-C, Habert E. Nouméa: IRD Editions; 2012:77–80.
11. Duangjai S, Wallnöfer B, Samuel R, Munzinger J, Chase MW: **Generic delimitation and relationships in Ebenaceae sensu lato: evidence from six plastid DNA regions.** *Am J Bot* 2006, **93**:1808–1827.
12. Duangjai S, Samuel R, Munzinger J, Forest F, Wallnöfer B, Barfuss MHJ, Fischer G, Chase MW: **A multi-locus plastid phylogenetic analysis of the pantropical genus *Diospyros* (Ebenaceae), with an emphasis on the radiation and biogeographic origins of the New Caledonian endemic species.** *Mol Phylogenet Evol* 2009, **52**:602–620.
13. Turner B, Munzinger J, Duangjai S, Temsch EM, Stockenhuber R, Barfuss MHJ, Chase MW, Samuel R: **Molecular phylogenetic of New Caledonian *Diospyros* (Ebenaceae) using plastid and nuclear markers.** *Mol Phylogenet Evol* 2013, **69**:740–763.
14. Vos P, Hogers R, Bleeker M, Reijans M, Van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M: **AFLP: a new technique for DNA fingerprinting.** *Nucleic Acids Res* 1995, **23**:4407–4414.
15. Tremetsberger K, Stuessy TF, Kadlec G, Urtubey E, Baeza CM, Beck SG, Valdebenito HA, Ruas CF, Matzenbacher NI: **AFLP phylogeny of South American species of *Hypochaeris* (Asteraceae, Lactuceae).** *Syst Bot* 2006, **31**:610–626.
16. Koopman WJM, Zevenbergen MJ, van den Berg RG: **Species relationships in *Lactuca s.l.* (Lactuceae, Asteraceae) inferred from AFLP fingerprints.** *Am J Bot* 2001, **88**:1881–1887.
17. Richardson JE, Fay MF, Cronk QCB, Chase MW: **Species delimitation and the origin of populations in island representatives of *Phylica* (Rhamnaceae).** *Evolution* 2003, **57**:816–827.
18. Despré L, Giells L, Redoutet B, Taberlet P: **Using AFLP to resolve phylogenetic relationships in a morphologically diversified plant species complex when nuclear and chloroplast sequences fail to reveal variability.** *Mol Phylogenet Evol* 2003, **27**:185–196.
19. Paun O, Schönswetter P, Winkler M, IntraBioDiv Consortium, Tribsch A: **Historical divergence vs. contemporary gene flow: evolutionary history of the calcicole Ranunculus alpestris group (Ranunculaceae) in the European Alps and the Carpathians.** *Mol Ecol* 2008, **17**:4263–4275.
20. Schulte K, Silvestro D, Kiehlmann E, Vesely S, Novoa P, Zizka G: **Detection of recent hybridization between sympatric Chilean Puya species (Bromeliaceae) using AFLP markers and reconstruction of complex relationships.** *Mol Phylogenet Evol* 2010, **57**:1105–1119.
21. Jabaily RS, Sytsma KJ: **Historical biogeography and life-history evolution of Andean Puya (Bromeliaceae).** *Bot J Linn Soc* 2012, **171**:201–224.
22. Gaudeul M, Rouhan G, Gardner MF, Hollingsworth PM: **AFLP markers provide insights into the evolutionary relationships and diversification of New Caledonian Araucaria species (Araucariaceae).** *Am J Bot* 2012, **99**:68–81.
23. Koopman WJM: **Phylogenetic signal in AFLP data sets.** *Syst Biol* 2005, **54**:197–217.
24. Degnan JH, Rosenberg NA: **Discordance of species trees with their most likely gene trees.** *PLoS Genet* 2006, **2**:762–768.
25. Meudt HM, Clarke AC: **Almost forgotten or latest practice? AFLP applications, analyses and advances.** *Trends Plant Sci* 2007, **12**:106–117.
26. Bussell JD, Waycott M, Chappill JA: **Arbitrarily amplified DNA markers as characters for phylogenetic inference.** *Perspect Plant Ecol, Evol Systematics* 2005, **7**:3–26.
27. Evanno G, Regnaut S, Goudet J: **Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study.** *Mol Ecol* 2005, **14**:2611–2620.
28. Vigouroux Y, Glaubitz JC, Matsouka Y, Goddman MM, Sánchez GJ, Doebley J: **Population structure and genetic diversity of New World maize races assessed by DNA microsatellites.** *Am J Bot* 2008, **94**:1240–1253.
29. Glor RE: **Phylogenetic insights on adaptive radiation.** *Ann Rev Ecol, Evol Systematics* 2010, **41**:251–270.
30. Lerner HRL, Meyer M, James HF, Hofreiter M, Fleischer RC: **Multilocus resolution of phylogeny and timescale in the extant adaptive radiation of Hawaiian honeycreepers.** *Curr Biol* 2011, **21**:1838–1844.
31. White F: *Flore de la Nouvelle-Calédonie et Dépendances. 19. Ébénacées*. Paris: Muséum National d'Histoire Naturelle; 1993.
32. Kapralov MV, Votintseva AA, Filatov DA: **Molecular adaptation during a rapid adaptive radiation.** *Mol Biol Evol* 2013, **30**:1051–1059.
33. Schmidt JM, Good RT, Appleton B, Sherrard J, Raymant GC: **Copy number variation and transposable elements feature in recent ongoing adaptation at the Cyp6g1 locus.** *PLoS Genet* 2010, **6**:e1000998.
34. Pintaud J-C, Tanguy J, Puig H: **Chorology of New Caledonian palms and possible evidence of Pleistocene rain forest refugia.** *C R Acad Sci* 2011, **324**:453–463.
35. Pillon Y, Hopkins HC, Munzinger J, Amir H, Chase MW: **Cryptic species, gene recombination and hybridization in the genus Spiraeanthemum (Cunoniaceae) from New Caledonia.** *Bot J Linn Soc* 2009, **161**:137–152.
36. Poncet V, Munoz F, Munzinger J, Pillon Y, Gomez C, Couderc M, Tranchant-Dubreuil C, Hamon S, de Kochko A: **Phylogeography and niche modelling of the relict plant Amborella trichopoda (Amborellaceae) reveal multiple Pleistocene refugia in New Caledonia.** *Mol Ecol* 2013. doi:10.1111/mec.12554.
37. Bennett KD: **Milankovitch cycles and their effects on species in ecological and evolutionary time.** *Paleobiology* 1990, **16**:11–21.
38. Kane NC, King MG, Barker MS, Raduski A, Karrenberg S, Yatabe Y, Knapp SJ, Rieseberg LH: **Comparative genomic and population genetic analyses indicate highly porous genomes and high levels of gene flow between divergent Helianthus species.** *Evolution* 2009, **63**:2061–2075.
39. Meudt HM, Lockhart PJ, Bryant D: **Species delimitation and phylogeny of a New Zealand plant species radiation.** *BMC Evol Biol* 2009, **9**:111.
40. Pillon Y, Munzinger J, Amir H, Hopkins HC, Chase MW: **Reticulate evolution on a mosaic of soils: diversification of the New Caledonian endemic genus Codia (Cunoniaceae).** *Mol Ecol* 2009, **18**:2263–2275.
41. Murienne J, Guilbert E, Grandcolas P: **Species diversity in the New Caledonian endemic genera Cephalidiosus and Nobarnus (Insecta: Heteroptera: Tingidae), an approach using phylogeny and species distribution modelling.** *Biol J Linn Soc* 2009, **97**:177–184.
42. Givnish TJ, Montgomery RA, Goldstein G: **Adaptive radiation of photosynthetic physiology in the Hawaiian lobeliads: light regimes, static light responses, and whole-plant compensation points.** *Am J Bot* 2004, **91**:228–246.
43. Baldwin BG, Sanderson MJ: **Age and rate of diversification of the Hawaiian silversword alliance (Compositae).** *Proc Natl Acad Sci USA* 1998, **95**:9402–9406.
44. Barraclough TG: **What can phylogenetics tell us about speciation in the Cape flora?** *Divers Distrib* 2006, **12**:21–26.
45. Ibanez T, Munzinger J, Dagostini G, Hequet V, Rigault F, Jaffré T, Birnbaum P: **Structural and floristic diversity of mixed tropical rainforest in New Caledonia: New data from the New Caledonian Plant Inventory and Permanent Plot Network (NC-PIPPN).** *Appl Veg Sci*. doi:10.1111/avsc.1270.
46. Tel-Zur N, Abbo S, Myslabodski D, Mizrahi Y: **Modified CTAB procedure for DNA isolation from epiphytic cacti of genera Hylocereus and Selenicereus (Cactaceae).** *Plant Mol Biol Report* 1999, **17**:249–254.
47. Bennett MD, Leitch IJ: **Angiosperm DNA *C*-values database (release 8.0, Dec. 2012).** http://www.kew.org/cvalues/.
48. Safer S, Tremetsberger K, Guo Y-P, Kohl G, Samuel MR, Stuessy TF, Stuppner H: **Phylogenetic relationships in the genus Leontopodium (Asteraceae: Gnaphalieae) based on AFLP data.** *Bot J Linn Soc* 2011, **165**:364–377.
49. Bonin A, Bellemain E, Bronken Eidesen P, Pompanon F, Brochmann C, Taberlet P: **How to track and assess genotyping errors in population genetic studies.** *Mol Ecol* 2004, **13**:3261–3273.
50. Swofford DL: *PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4*. Sunderland, Massachusetts: Sinauer Associates; 2003.
51. Huson DH, Bryant D: **Application of phylogenetic networks in evolutionary studies.** *Mol Biol Evol* 2006, **23**:254–267.
52. Drummond AJ, Suchard MA, Xie D, Rambaut A: **Bayesian phylogenetics with BEAUti and the BEAST 1.7.** *Mol Biol Evol* 2012, **29**:1969–1973.

53. Ligges U, Mächler M: **Scatterplot3d - an R package for visualizing multivariate data.** *J Stat Softw* 2003, **8**:1–20.
54. Excoffier L, Laval G, Schneider S: **Arlequin (version 3.0): an integrated software package for population genetics data analysis.** *Evol Bioinformatics Online* 2005, **1**:47–50.
55. Pritchard JK, Stephens M, Donnely P: **Inference of population structure using multilocus genotype data.** *Genetics* 2000, **155**:945–959.
56. Hubisz MJ, Falush D, Stephens M, Pritchard JK: **Inferring weak population structure with the assistance of sample group information.** *Mol Ecol Resour* 2009, **9**:1322–1332.
57. **Lifeportal.** http://www.uio.no/english/services/it/research/hpc/lifeportal/.
58. Earl DA, Von Holdt BM: **STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method.** *Conserv Genet Resour* 2012, **4**:359–361.
59. Jakobsson M, Rosenberg NA: **CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure.** *Bioinformatics* 2007, **23**:1801–1806.
60. Rosenberg NA: **Distruct: a program for the graphical display of population structure.** *Mol Ecol Notes* 2004, **4**:137–138.

doi:10.1186/1471-2148-13-269
Cite this article as: Turner *et al.*: Analyses of amplified fragment length polymorphisms (AFLP) indicate rapid radiation of *Diospyros* species (Ebenaceae) endemic to New Caledonia. *BMC Evolutionary Biology* 2013 **13**:269.

ADDITIONAL FILES



Additional File 1: STRUCTURE results of suboptimal K values (3, 6, 16 and 21) in comparison with K =2. Delta K likelihoods are given for each K.

Additional File 2: Table showing the population statistics inferred from non-hierarchical AMOVA based on STRUCTURE results.

Population	Taxon	Sample ID	Number of individuals	Number of polymorphic sites	Pairwise difference	Average gene diversity	Pairwise difference	Average gene diversity
01	*D. calciphila*	BT312-BT317	6	123	52.3	0.067	51.4	0.066
02	*D. cherrieri*	BT262, BT276-BT278	4	52	29.5	0.037		
03	*D. cherrieri*	BT293-BT297	5	69	33.4	0.042		
04	*D. erudita*	BT259-BT261, BT273-BT275	6	48	22.1	0.028		
05	*D. erudita*	BT280-BT285, BT287	7	110	44.5	0.056		
06	*D. flavocarpa*	BT126-BT130	5	66	32.6	0.041		
07	*D. flavocarpa*	BT155, BT158-BT159	3	109	72.7	0.092		
08	*D. glans*	BT020-BT022	3	45	30.0	0.038		
09	*D. glans*	BT075	1	-	-	-		
10	*D. glans*	BT082, BT084, BT087, BT093-BT094	5	121	55.8	0.070	51.5	0.065
11	*D. impolita*	BT101-BT105	5	105	50.2	0.063		
12	*D. inexplorata*	BT304, BT307-BT311	6	137	60.5	0.076		
13	***D. labillardierei***	**BT121-BT125, BT179**	**6**	**91**	**39.4**	**0.050**		
14	***D. labillardierei***	**BT178, BT180-BT182**	**4**	**91**	**48.7**	**0.061**		
15	***D. minimifolia***	**BT134-BT135, BT234**	**3**	**78**	**52.0**	**0.066**		
16	***D. minimifolia***	**BT230-BT233**	**4**	**87**	**46.7**	**0.059**		
17	*D. minimifolia*	BT263-BT264, BT266-267, BT269-BT270	6	166	71.5	0.090		
18	*D. pancheri*	BT029-BT031, BT035	4	118	64.0	0.081		
19	*D. pancheri*	BT076-BT079	4	131	71.5	0.090		
20	*D. parviflora*	BT042	1	-	-	-		
21	*D. parviflora*	BT059, BT062-BT063, BT068	4	154	85.0	0.107		
22	*D. parviflora*	BT080, BT085, BT089-BT090	4	176	94.8	0.120		
23	*D. parviflora*	BT248-BT250, BT252-BT253	5	170	79.6	0.100		
24	*D. parviflora*	BT289-BT292	4	125	68.2	0.086		
25	*D. perplexa*	BT147-BT151	5	161	75.4	0.095		
26	*D. pustulata*	BT111-BT114	4	82	43.5	0.055		
27	*D. pustulata*	BT136-BT140	5	79	37.4	0.047		
28	*D. pustulata*	BT257-BT258, BT265, BT268, BT271-BT272	6	105	45.2	0.057		
29	*D. revolutissima*	BT116-BT120	5	143	67.2	0.085		
30	*D. revolutissima*	BT218-BT222	5	124	59.2	0.075		
31a	***D. tridentata***	**BT202-BT205**	**4**	**110**	**58.7**	**0.074**		
31b	***D. tridentata***	**BT206-BT207**	**2**	**45**	**45.0**	**0.057**		
32	*D. trisulca*	BT185, BT192, BT197, BT199-BT201	5	155	75.4	0.095		
33	*D. umbrosa*	BT061, BT065-BT066, BT071, BT073	5	104	46.6	0.059		
34	*D. umbrosa*	BT170-BT171, BT175-BT177	5	59	28.4	0.036		
35	*D. umbrosa*	BT246-BT247, BT251, BT254, BT256	5	149	70.2	0.088		
36a	***D. veillonii***	**BT224**	**1**	**-**	**-**	**-**		
36b	***D. veillonii***	**BT226-BT229**	**4**	**106**	**57.2**	**0.072**		
37a	***D. vieillardii***	**BT017, BT088, BT100**	**3**	**86**	**57.3**	**0.072**	**-**	**-**

37b	***D. vieillardii***	**BT023-BT026**	**4**	**90**	**48.8**	**0.062**		
38	*D. vieillardii*	BT055, BT057-BT058	3	120	80.0	0.101	-	-
39	***D. vieillardii***	**BT091-BT092, BT098**	**3**	**108**	**72.0**	**0.091**		
40	*D. vieillardii*	BT215-BT217	3	82	54.7	0.069		
41	*D. vieillardii*	BT324-BT325, BT328	3	74	49.3	0.062		
42	*D. yahouensis*	BT237-BT239	3	72	48.0	0.061		
43	*D.* sp. Pic N'ga	BT319, BT321-BT323	4	110	60.3	0.076		

Populations marked bold differ in this analysis from the general population grouping given in Table 3.

Additional File 4: Table giving the details of the different AMOVAs conducted.

Analysis	No of groups	Groups	Populations within groups
non-hierarchical	1	all	01 - 43
species wise	21	calciphila	01
		cherrieri	02, 03
		erudita	04, 05
		flavocarpa	06, 07
		glans	08, 09, 10
		impolita	11
		inexplorata	12
		labillardierei	13, 14
		minimifolia	15, 16, 17
		pancheri	18, 19
		parviflora	20, 21, 22, 23, 24
		perplexa	25
		pustulata	26, 27, 28
		revolutissima	29, 30
		tridentata	31
		trisulca	32
		umbrosa	33, 34, 35
		veillonii	36
		vieillardii	37, 38, 39, 40, 41
		yahouensis	42
		sp Pic N'ga	43
STRUCTURE	2	White	01, 13, 15, 17, 26 - 28, 31b, 36a, 37b, 39, 41, 43
		Grey	02 - 12, 14, 16, 18 - 25, 29 – 30, 31a, 32 – 34, 36b, 37a, 38, 40, 42
geographic	3	north	30 - 32, 40
		middle	02 - 07, 11, 13 - 15, 17, 24 - 29, 34
		south	01, 08 - 10, 12, 16, 18 - 23, 33, 35 - 39, 41 - 43
water	2	dry	01 - 05, 11, 12, 15 - 17, 26 - 31, 36, 38, 40, 42, 43
		humid	06 - 10, 13, 14, 18 - 25, 32 - 35, 37, 39, 41
soil	5	ultramafic	05, 08 - 10, 18 - 24, 33, 35, 37 - 39, 41, 43
		limestone	01, 12, 26, 31
		volcanic	02 - 04, 06, 07, 11, 13 - 17, 25, 27, 28, 34, 36, 42
		serpentine	29, 30, 40
		schist	32

The numbers in the population's column are the same as given in Table 3, respectively, in Additional file 1 for the STRUCTURE based AMOVA.

Additional File 3: Figure of the neighbour joining dendrogram coloured according to soil type (colour of the branches) and water availability (colour of taxa names). This dendrogram is the same as Fig. 3A, but coloured according to ecological features.

CHAPTER 3

Genome wide RADseq resolves adaptive radiation of *Diospyros* species in New Caledonia

Barbara Turner, Ovidiu Paun, Jérôme Munzinger, Mark W. Chase, Rosabelle Samuel

Status: to be submitted to Molecular Biology and Evolution

Contribution: Collection of material, collection of data, phylogenetic analysis of data, manuscript writing/editing

INTRODUCTION

Diospyros (Ebenaceae) is a large genus (roughly 500 species) of woody plants found world-wide in the tropics and subtropics. Among these, 31 *Diospyros* species are found in New Caledonia, an archipelago in the southern pacific. Previous studies based on plastid markers (Duangjai *et al*. 2009) showed that *Diospyros* colonised New Caledonia at least four times via long distance dispersal. Of these colonisations, two dispersal events resulted in one species each, a third dispersal produced a small clade comprising five species, and the forth event has generated a clade of 24 species. These 24 species are endemic to New Caledonia and have been shown to be closely related, with low-copy nuclear and plastid markers (Turner *et al*. 2013a; Duangjai *et al*. 2009), but their inter-relationships have been difficult to resolve due to the low levels of variation detected in these relative standard phylogenetic markers. Even the application of genome wide AFLP marks did not clarify phylogenetic relationships between these species (Turner *et al*. 2013b). Most of these closely related species are morphologically and ecologically clearly differentiated, and species delimitations have been confirmed by AFLP data (Turner *et al*. 2013b). Due to its special geological history, New Caledonia is a mosaic of soil-types (Pelletier 2006; Maurizot and Vendé-Leclerc 2009), and in combination with the climatic factors this results in a heterogeneous environment across a fairly small geographic range. *Diospyros* species are found in many of these habitats and in some localities, several species occur in microsympatry. Dating analysis using combined plastid and nuclear DNA sequence data showed that lineages forming this group of *Diospyros* species arrived in New Caledonia around nine million years ago (mya; Turner *et al*. 2013a), with a more recent radiation that produced the 24 endemic species. Taken into consideration that these are woody plants with generation time of at least several years (~7 years; Verdú 2002), it becomes obvious that these are young species. Phenotypic changes and adaptation to environmental conditions do not necessarily depend on large genetic alterations (Kane *et al*. 2009); they can be due to mutations at few loci. Finding such relatively small differences within the genomes of the species is challenging.

Restriction-site associated DNA (RAD) sequencing is a Next-generation sequencing application proposed by Miller *et al*. (2007). DNA fragments obtained from digestion with restriction enzymes are sequenced and single nucleotide polymorphisms (SNPs) are identified in these sequences. Like other restriction-site based methods (e.g. AFLP; Vos *et al*. 1995), this technique is useful at low taxonomic levels (e.g. intra- and interspecific level; Rubin *et al*. 2012; Cariou *et al*. 2013) because distantly related taxa have fewer restriction sites in common than closely related taxa and therefore fewer homologous fragments will be obtained. RAD has been used to reveal differences in the genomes between varieties of a species (e.g. *Solanum melongena*, Barchi *et al*. 2011) or individuals of a population (e.g. *Sisymbrium austriacum*,

Vandepitte *et al*. 2013; *Wyeomyia smithii*, Emerson *et al*. 2010). A few studies have made use of RADseq to resolve phylogenetic relationships between species (e.g. *Pedicularis*, Eaton and Ree 2013; cichlid fish, Wagner *et al*. 2013; *Drosophila*, Rubin *et al*. 2012).

This study focuses on the species-rich group of closely related New Caledonian *Diospyros* species. Here we aim to clarify the phylogenetic relationships between these *Diospyros* species using RAD sequencing.

MATERIALS AND METHODS

Taxon Sampling and DNA isolation

Leaf material from New Caledonian *Diospyros* species was collected on Grande Terre and Île des Pins (Fig. 1) and stored in silica gel. Herbarium voucher are deposited in the herbaria of Noumea (NOU), University of Montpellier II (MPU) and University of Vienna (WU). Details of collection and vouchers are given in Turner *et al*. (2013b). We included in this study 84 individuals from 26 localities, representing 21 species of New Caledonian *Diospyros* (Tab. 1) that have been previously shown to have radiated rapidly after a single long-distance dispersal event (Turner *et al*. 2013a). Whenever possible, we aimed to investigate at least two individuals per locality and a minimum of three individuals per species. One of the studied species (collected at Pic N'ga on Île des Pins) could not be unambiguously identified, due to the absence of diagnostic reproductive organs at collection, and will be hereafter referred to as *D.* sp. Pic N'ga.

Leaf DNA extractions performed with a modified sorbitol/high-salt CTAB method (Tel-Zur *et al*. 1999) were already available (Turner *et al*. 2013a). As we observed significant differences between standard Nanodrop (Thermo Scientific) and Quant-It Pico-Green (Life Technologies) quantifications of the DNA samples, these have been purified using the NucleoSpin gDNA clean-up kit (Macherey-Nagel), according to the manufacturer's protocol.

RADseq Library Preparation

By using an average genome size in the target group of 1C = 1.9 pg (Turner *et al*. 2013a) and the RAD counter available from https://www.wiki.ed.ac.uk/display/RAD Sequencing we have estimated that 60 individually barcoded samples can be pooled together to investigate *ca*. 18,000 restriction sites per genome, if using the *SbfI* high-fidelity restriction enzyme (New England Biolabs). A second RAD library was later prepared in order to increase the coverage of selected samples to a minimum 1 mill high-quality read pairs per individual and to add 24 new individuals. The RAD libraries were prepared using a protocol adapted from Baird *et al*. (2008) with modifications as described below. We started with 300 ng DNA per individual and used

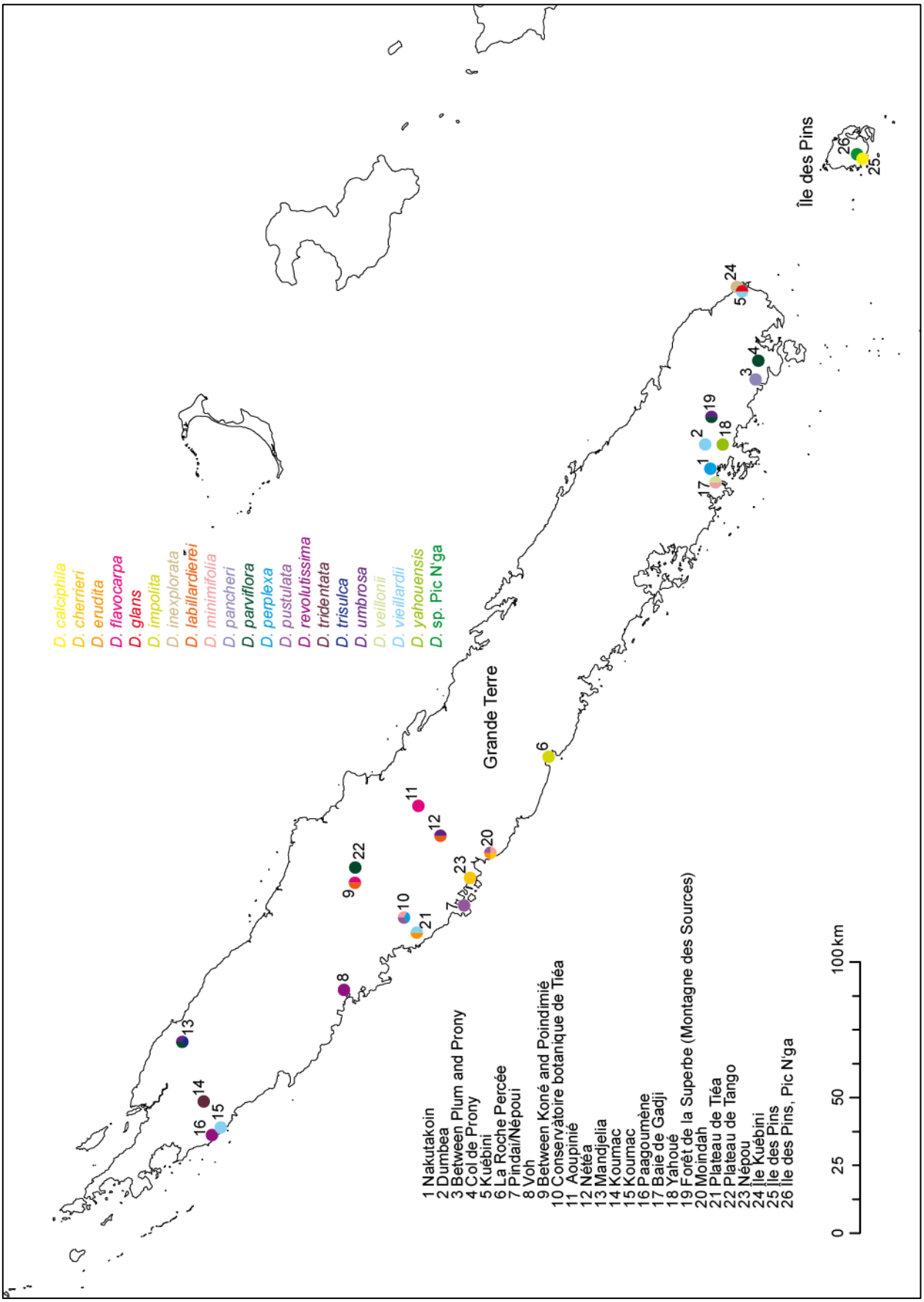


Figure1: Map of New Caledonia with sampling localities. Dots indicate sapling sites; the numbers associated with each dot refer to the list of sampling sites on this figure. The numbers are used throughout the text to characterize sampling sites. The colour of the dots indicates the species collected at this locality. The size of “pie slices” is not necessarily proportional to species abundance.

Table 1: Table of accessions, showing all individuals used in this study. The numbers of sampling localities are the same as in Fig. 1.
Voucher-Codes:
JMXXXX: collection number J. Munzinger;
Tree N° XXX: Tree of New Caledonian Plant Inventory and Permanent Plot Network (NC-PIPPN, Ibanez *et al*. 2013);
NOUXXXXXXX: Herbarium accession number of Noumea herbarium (NOU);
WUXXXXXX: Herbarium accession number of the Herbarium of the University Vienna (WU);
PXXXXXXXX: Herbarium accession number of the Herbarium of the Natural History Museum Paris;
MPU: Herbarium of the University of Montpellier II

Taxon	Sample ID	Sampling location	Voucher
D. calciphila F.White	BT313	25, littoral forest	JM6650, JM6653
	BT314		(MPU, NOU, P)
	BT317		
D. cherrieri F.White	BT276	20, dry forest	NOU054492
	BT278		NOU054008
D. cherrieri	BT293	23, dry forest	NOU079547
	BT294		
D. erudita F.White	BT259	20, dry forest	WU062855
	BT261		NOU079544, WU062870
D. erudita	BT280	21, dry forest	WU062858
	BT281		Chambrey & Turner 20 (NOU)
D. flavocarpa (Vieill. ex P.Parm.) F.White	BT129	9, humid mountain forest	JM6625 (NOU)
	BT130		
D. flavocarpa	BT156	11, dense humid mountain forest	JM6632 (NOU)
	BT157		
D. glans F.White	BT093	5, forest near river	NOU022860
	BT094		
D. impolita F.White	BT102	6, mesophyll forest near beach	NOU019538
	BT103		
	BT105		
D. inexplorata F.White	BT308	24, littoral forest	NOU005818
	BT310		
	BT311		
D. labillardierei F.White	BT122	9, river edge in mountain forest	JM6624 (NOU)
	BT125		
D. labillardierei	BT178	12, river edge	(NOU031346)
	BT182		
D. minimifolia F.White	BT131	10, dry forest	NOU019556
	BT135		
D. minimifolia	BT232	17, mesophyll forest near beach	NOU019554
	BT233		
D. minimifolia	BT263	20, dry forest	NOU079549, WU062872
	BT269		NOU054493
D. pancheri Kosterm.	BT028	3, forest near road	JM6619, JM6620 (NOU)
	BT031		
	BT035		
D. parviflora (Schltr.) Bakh.	BT038	4, wet forest	
	BT041		
	BT042		
D. parviflora	BT187	13, mountain forest	JM6636 (NOU)
D. parviflora	BT250	19, humid forest at low elevation	tree no. 23109
D. parviflora	BT289	22, mountain forest	NOU079550
	BT290		
	BT291		
D. perplexa F.White	BT004	1, mesophyll forest	JM6611, JM6613 (NOU)
D. perplexa	BT147	10, forest near river	JM6630 (NOU)
	BT148		

Table 1 continued

Taxon	Sample ID	Sampling location	Voucher
D. pustulata F.White	BT111	7, dry forest	
	BT112		
D. pustulata	BT137	10, dry forest	JM6629 (NOU)
	BT140		
D. pustulata	BT265	20, dry forest	NOU079548, WU062871
	BT268		NOU053999
D. revolutissima F.White	BT117	8, maquis	NOU023189
	BT120		
D. revolutissima	BT219	16, maquis	JM6640 (NOU)
	BT221		
D. tridentata F.White	BT203	14, dry forest at low	JM6639 (NOU)
	BT206	elevation	
	BT207		
D. trisulca F.White	BT185	13, mountain forest	NOU031344
	BT192		JM6637 (NOU)
	BT199		
	BT201		
D. umbrosa F.White	BT176	12, dense humid forest	JM6635 (NOU)
	BT177		
D. umbrosa	BT197	13, mountain forest	
D. umbrosa	BT246	19, humid forest at low	NOU023234
	BT247	elevation	
D. veillonii F.White	BT224	17, mesophyll forest	NOU019582
	BT226	near beach	
	BT227		
D. vieillardii (Hiern) Kosterm.	BT025	2, forest near river	JM6618 (NOU)
	BT026		
D. vieillardii	BT088	5, forest near river	
	BT100		
D. vieillardii	BT215	15, maquis	NOU023242
	BT217		
D. vieillardii	BT286	21, dry forest	
D. yahouensis (Schltr.) Kosterm.	BT238	18, mesophyll forest	P00057340
	BT239		
D. sp. Pic N'ga	BT318	26, maquis	JM6065 (NOU)
	BT320		
	BT323		

double barcoding with six base-pair barcodes within P1 Illumina adapters and, respectively, four base-pair barcodes within P2 adapters. P1 and P2 barcodes were chosen to differ by at least three base pairs from each other. We ligated 200 mM P1 adapters to the restricted samples overnight at 16 °C. Samples containing differently barcoded P1 adapters were pooled and sheared by sonication in a Bioruptor Pico (Diagenode) to an average size of ca. 400 bp using two cycles of 55s “on” and 55s “off” at 6 °C. We have further performed a left- and right-size selection with SPRIselect (Beckman Coulter) by using 0.7x and 0.55x volume of SPRI reagent to sample, according to the manufacturer’s protocol. After ligating P2 adapters, samples (at this stage barcoded with different P1-P2 combinations) were pooled in one library so that each sample would be equally represented. Two size selections on the left side with 0.65x volume

SPRI reagent were finally performed: one before the 18 cycles PCR amplification with the Phusion Master Mix (Thermo Fischer Scientific) and once after. Libraries were sequenced on an Illumina HiSeq at CSF Vienna (http://csf.ac.at/ngs/) as 100-bp paired-end reads.

Filtering SNPs from RADseq Data

As the first step of the bioinformatic analyses, libraries were demultiplexed into individual samples according to the respective barcode combinations using the RADPOOLS module of the RADTOOLS v. 1.2.4 package (Baxter *et al*. 2011). During this process, we have allowed for single errors at the barcode sites, as the reads could still be unambiguously allocated to individuals. Disqualified reads have been discarded from further analyses. The 84 individual files were then imported in the CLC GENOMIC WORKBENCH v. 6.5 (Qiagen) and trimmed/filtered to retain only full length (i.e., 94 bp after barcode trimming) reads, free of any adaptor sequence, with all bases of a Phred quality score over or equal to 30 (i.e., accuracy ≥ 99.9%). The final high-quality, filtered and demultiplexed data set contained close to 161 million read-pairs.

The forward reads were then used for running the DENOVO_MAP.PL script of STACKS v. 1.05 (Catchen *et al*. 2011). To find the best settings for STACKS, we first varied the value of the minimum number of identical reads required for a stack (i.e., the setting “m”) from five to 15, by allowing one base-pair difference between loci when processing one individual (i.e., the setting “M”) and when building the catalogue (i.e., setting “n”). We have chosen the value of m = 13 as best for our data because it delivered the most polymorphic stacks with less than 10 SNP positions that are covered by data in at least 90% of individuals (Tab. 2). Further, for the value of m = 13 we have run additional tests by varying the value of “M” from one to four and the value of “n” from zero to six (Tab. 2). The final combination of settings chosen was m = 13, M = 1 and n = 1.

The deleveraging algorithm of ustacks has been left on, in order to split loci merged incorrectly and remove highly repetitive sequences from further analyses. To avoid retention of any merged paralogs, the loci having > 10 SNPs have been blacklisted in further analyses by filtering them out using the EXPORT_SQL.PL script from STACKS. Finally, we have retained for phylogenetic analyses only SNPs from the loci with data present for at least 75 individuals. The SNP data have been extracted by using the POPULATIONS script of STACKS. Exploration of data matrices including loci with more missing data (maximum 20 individuals) and allowing more SNPs per locus resulted in less resolved phylogenetic trees and have been discarded.

Phylogenomic Analyses

Phylogenetic analyses were run using both maximum parsimony and Bayesian inference. Parsimony analyses were run with PAUP* v4b10 (Swofford 2003) with gaps traded as missing,

stepwise addition and tree-bisection-reconnection. To estimate clade support, bootstrapping with 1,000 replicates was performed. We report here a strict consensus tree, rooted with *D. vieillardii,* according to earlier results (Turner *et al*. 2013a). For Bayesian inference and molecular dating, the program BEAST v1.7.5 (Drummond *et al*. 2012) was run on CIPRS Science Gateway (http://www.phylo.org/portal2/; Miller et al. 2010). Estimation of evolutionary models was conducted with jModeltest v2.1.4 (Darriba *et al*. 2012; Guindon and Gascuel 2003). For Bayesian analysis, the transversional model (TVMef; Posada 2003) with equal frequencies modelled with a gamma distribution and a proportion of invariable sites (TVMef+Γ+I) has been indicated as the best. We used a relaxed uncorrelated log-normal clock model (Drummond *et al*. 2006) and a Yule speciation model (Gernhard 2008; Yule 1925). Substitution rates between bases (gamma shape 10), alpha (gamma shape 10), and p-inv (uniform) were inferred by Modeltest. Two independent Metropolis-coupled Markov chain Monte Carlo (MCMC) analyses each with 20 million generations were run sampling each 1,000th generation. The initial 10% of trees obtained from each MCMC run were removed as burn in; the remaining trees of both runs were used to calculate a maximum clade credibility tree. Dating was obtained by taking into account the age of the split between *D. vieillardii* and the rest of the group (7.2 million years ago) conforming to a previous dating of New Caledonian *Diospyros* (Turner *et al*. 2013a). This age was taken as minimum (i.e., no fixed upper limit) for the shared node of *D. vieillardii* and the rest of the endemic NC clade.

Table 2: The number of stacks obtained with STACKS by varying the value of the minimum number of identical reads required for a stack (m), number of nucleotides different between loci when processing one individual (M), and when building the catalogue (n).

Settings (m, M, n)	Total stacks	Stacks present in at least 2 inds	Stacks with 1-10 SNPs, covered in at least 70 inds	Stacks with over 11 SNPs
5, 1, 1	2,377,725	163,982	1,066	1,024
7, 1, 1	1,308,369	92,923	1,622	1,130
9, 1, 1	830,081	62,560	1,711	1,117
10, 1, 1	674,766	53,443	1,722	1,111
11, 1, 1	552,461	46,326	1,723	1,114
12, 1, 1	455,927	41,211	1,720	1,103
13, 1, 1	379,550	37,336	1,725	1,093
14, 1, 1	320,107	34,291	1,719	1,096
15, 1, 1	272,903	31,905	1,719	1,088
13, 2, 1	371,779	35,652	1,479	1,347
13, 3, 1	366,112	34,978	1,388	1,431
13, 1, 0	411,564	44,390	1,553	112
13, 1, 2	364,437	35,505	1,227	1,765
11, 1, 2	532,568	45,475	1,174	1,836

Clustering and tracking admixture

Within groups of closely related species, hybridization can shape evolutionary patterns long after initial divergence events (Mallet 2005; Lexer and Widmer 2008; Rieseberg 2009; Martin *et al*. 2013). Hence representation of relationships as networks rather bifurcating trees could better reflect real situations (Huson and Scornavacca 2011). Homoplasy can of course be derived by independent substitutions as well as hybridization, so network-like results are not necessarily the product of the latter phenomenon. We used SPLITSTREE v. 4.12.6 (Huson and Bryant 2006) to draw a phylogenetic network based on the Hamming distance (Hamming 1950). The simple calculation method of Hamming distance was considered appropriate for the RAD-seq derived SNP dataset that completely lacks indels.

To investigate higher-level clustering of the included individuals and potential hybridization between different groups we have used the program STRUCTURE v2.3.3 (Pritchard *et al*. 2000; Hubisz *et al*. 2009). We ran STRUCTURE at the Bioportal of the University of Oslo for K = 1 to 10, each with ten replicates and a model based on admixture and independent allelic frequencies, without taking into account information regarding sampling localities. Each run had five million iterations with 10% additional burn in. The calculation of deltaK (ΔK; Evanno *et al*. 2005) and preparation of the input files for CLUMPP were performed with HARVESTER (Earl *et al*. 2012). To avoid any stochastic aspect of the process, we have produced a permuted matrix from ten replicates for selected K values with CLUMPP v1.1.2 (Jakobsson and Rosenberg 2007) in the full search algorithm for K ≤ 3, and the greedy algorithm and 1,000 repeats for K values above 3. The graphical display of STRUCTURE results was prepared with DISTRUCT v1.1 (Rosenberg 2004).

In addition to STRUCTURE analysis we plotted a principal coordinate (PCO) using the R-package SCATTERPLOT3D (Ligges and Mächler 2003) based on the Hamming distance matrix to visualise genetic clustering of individuals.

Searching for Adaptive Signals

The phylogenomic trees obtained suggest that ecologically driven isolation (particularly thorough soil type and climatic heterogeneity on New Caledonia) could have made a major contribution in shaping speciation events across the radiating group. In particular, both the MP and BI results provide evidence of several sister clades with divergent preference for ultramafic versus volcanic soils: *D. flavocarpa* vs. *D. umbrosa*; *D. perplexa* (L10) vs. *D. parviflora* (L22); *D. yahouensis* + *D. perplexa* BT004 vs. *D. pancheri*; *D. minimifolia* (L17) vs. *D.* sp. Pic N'ga. In order to test if any particular genomic regions have systematically been affected as a result of positive selection or genetic hitchhiking, we searched for RAD regions that contained SNPs with pairwise F_{ST} values > 0.5 at least for two pairs of sister species with divergent soil preferences.

We have then made use of the paired-end RADseq data (Etter *et al.* 2011) and assembled mini-contigs of the candidate stacks by extracting a list of reads for each locus with SORT_READ_PAIRS.PL from STACKS, sorting reads from FastQ files with FASTQ.FILTER.PL (Luis M. Rodriguez unpublished, available from http://enveomics.blogspot.co.at/2013/04/fastqfilterpl.html) and assembling each set in the CLC GENOMIC WORKBENCH (Qiagen), with automatic optimization of the word and bubble sizes and updating the contigs after mapping back the reads. We finally performed functional annotation analyses for the obtained contigs using BLAST2GO (BioBam; Götz *et al.* 2008) with default settings and integrating GO (www.geneontology.org), KEGG (www.genome.jp/kegg) and InterProScan (www.ebi.ac.uk/Tools/InterProScan) information in our results. The biological meaning of the set of sequences has been investigated with the combined graph option of BLAST2GO.

RESULTS

Filtering SNPs from RADseq data

After demultiplexing, trimming and filtering raw reads, we retained on average 1.9 mill +/- 0.7 mill high-quality pairs of reads per individual. Under the final parameters, the *de novo* assembly pipeline of STACKS produced 37,336 loci (excluding any stacks identified in only one individual), which corresponds to the number of RAD loci expected based on the genome size of these species (i.e., twice the number of predicted restriction sites). By retaining SNPs from loci covered in minimum 75 individuals with maximum ten polymorphic nucleotide positions, we obtained a data matrix containing 8,488 SNPs, which has been used for phylogenomic analyses. We further filtered out any apomorphic SNPs, distinguishing single individuals from the rest, to obtain a reduced matrix of 2,832 SNPs for the STRUCTURE analyses.

Phylogenomic analyses

Since MP, BI and distance based methods resulted in similar topologies for convenience reasons we will refer hereafter to two groups of species (1 and 2), each with two sub-sets of taxa (A and B). Two of these (sub) groups are monophyletic and well supported in both MP and BI. The set of taxa marked on the tree as 2A is potentially a grade; the subset marked as 1A is clearly a grade. Subset one includes *D. vieillardii*, *D. umbrosa* and *D. flavocarpa* (1A) as well as *D. cherrieri* and *D. veillonii* (group 1B). The second subset comprises *D. erudita*, *D. glans*, *D. impolita*, *D. labillardierei*, *D. pancheri*, *D. parviflora*, *D. perplexa*, *D. pustulata*, *D. revolutissima*, *D. trisulca* and *D. yahouensis* (2A) as well as *D. minimifolia*, *D. tridentata*, *D.* sp. Pic N'ga, *D. calciphila* and *D. inexplorata* (group 2B). These groupings will be used throughout the text and are marked in all trees.

The MP analysis resulted in 31 equally parsimonious trees. One of these most parsimonious trees is given in Figure 2, indicating on it branches present in the strict consensus tree. All species included except for *D. erudita*, *D. minimifolia*, *D. parviflora*, and *D. perplexa* form unique genetic clusters. Although individuals from the same population of each of these four species group together, the different population samples for these morphologically defined species do not cluster together. Furthermore, all populations of *D. pustulata* form a unique cluster, but it lacks bootstrap percentage (BP) greater than 50. *Diospyros vieillardii* is isolated in a highly supported and internally structured cluster (BP 100). The next group, clearly separated from the rest of the species, is a clade (BP 100) formed by *D. flavocarpa* and *D. umbrosa*, followed by the strongly supported (BP 100) group 1B, including just *D. cherrieri* and *D. veillonii*. The rest of the species forms groups present in the strict consensus tree of the parsimony analysis, but having low bootstrap support. The species of group 2B are forming a medium supported group (BP 89). Individuals of *D. calciphila* and *D. inexplorata* could not be separated from each other. We do not observe any major grouping related to ecological factors like soil type or water availability, but sister species often show divergent ecological preferences.

To make the BI tree clearer we have collapsed its structure to the highest possible level (either species or population level, depending on what was possible). The general topology of the BI tree (Fig. 3) is similar to that of the MP (Fig. 2). However relationships between some of the clades within group 2A differ between MP and BI. The backbone of the BI tree is slightly better supported than the MP tree. The phylogenetic relationships between the earlier diverged lineages (group 1) are the same as in the MP. However, the sister clade relationship between *D.flavocarpa*/*D. umbrosa* and the remaining group, excluding *D. vieillardii* is not supported (i.e., Bayesian posterior probability [PP] lower than 0.95), in contrast to the MP tree. Apart from *D. erudita*, *D. minimifolia*, *D. parviflora*, *D. perplexa* and *D. pustulata,* all other species form highly supported clusters (PP 1.00). However, as in the MP analysis, individuals of the same population always group together and are well supported (PP 1.00) except for one population of *D. pustulata* (location 20). Higher-level relationships between *D. revolutissima*, the group of *D. erudita*, *D. impolita*, *D. pustulata*, and group 2B are not supported. Similarly as in the MP tree, in the BI results *D. calciphila* does not form a unique cluster and could not be clearly separated from *D. inexplorata.*

The molecular clock analysis resulted in a slightly older age for the split of *D. vieillardii* from the rest of the group, estimated at 7.4 mya, with a broad 95% confidence interval of 2.7 myr. The next divergence (i.e., *D. flavocarpa*/*D. umbrosa* from the rest of the species) took place around 6.6 mya. The lineage forming *D. cherrieri* and *D. veillonii* separated from the rest around 5.6 mya. The other lineages started to diversify around 4 mya. Clade 2B is a young group, about 2.7 myr old. Most speciation events seem to have taken place between 3.5 and 1.5 mya.

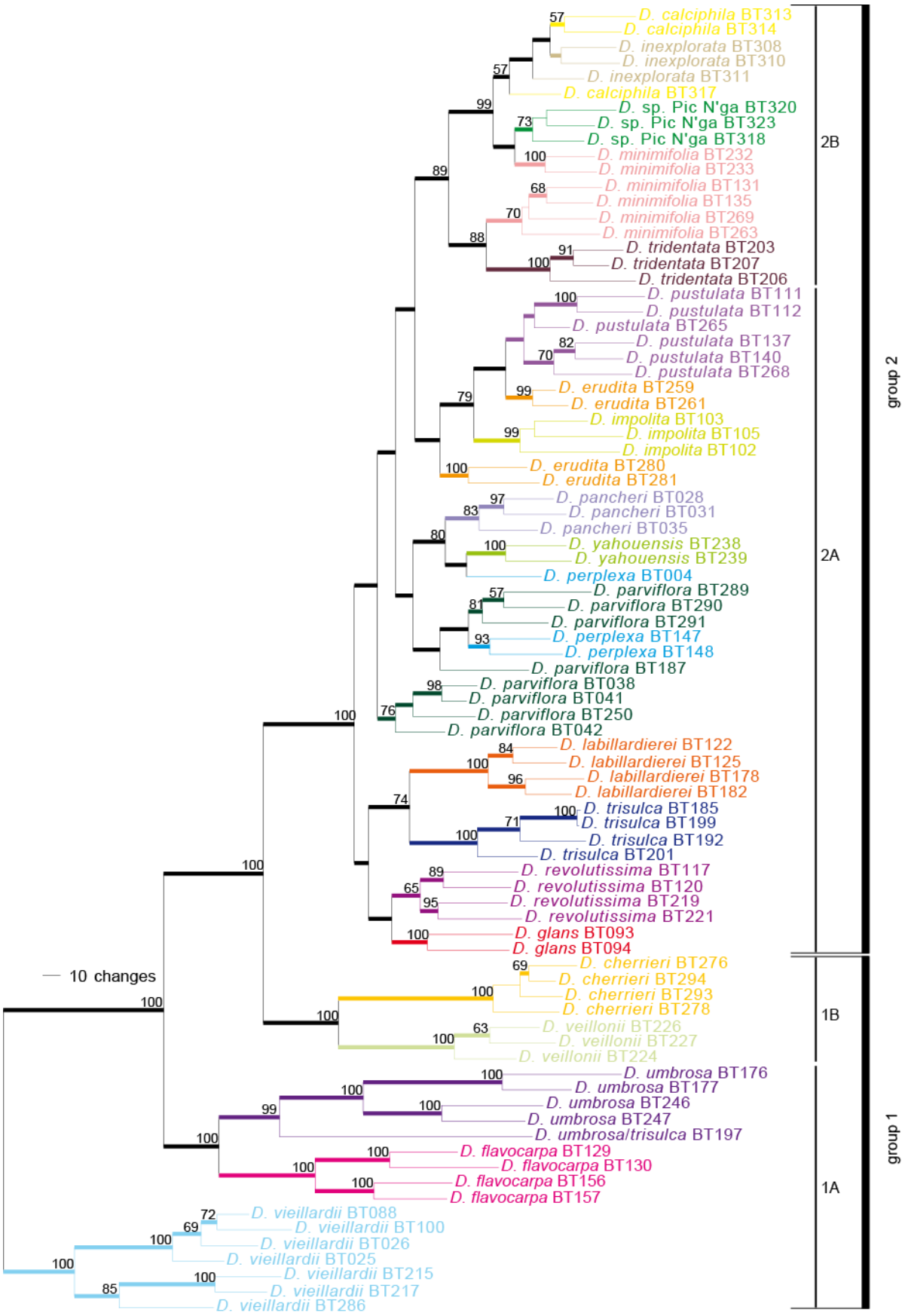


Figure 2: One of 31 equally parsimonious trees of 8,488 SNP data set derived from RAD-seq in the radiating *Diospyros* group from New Caledonia. Numbers on branches indicate bootstrap support over 50%. Branches in bold are present in the strict consensus of the 31 trees. For improved visibility, each species was randomly coloured differently.

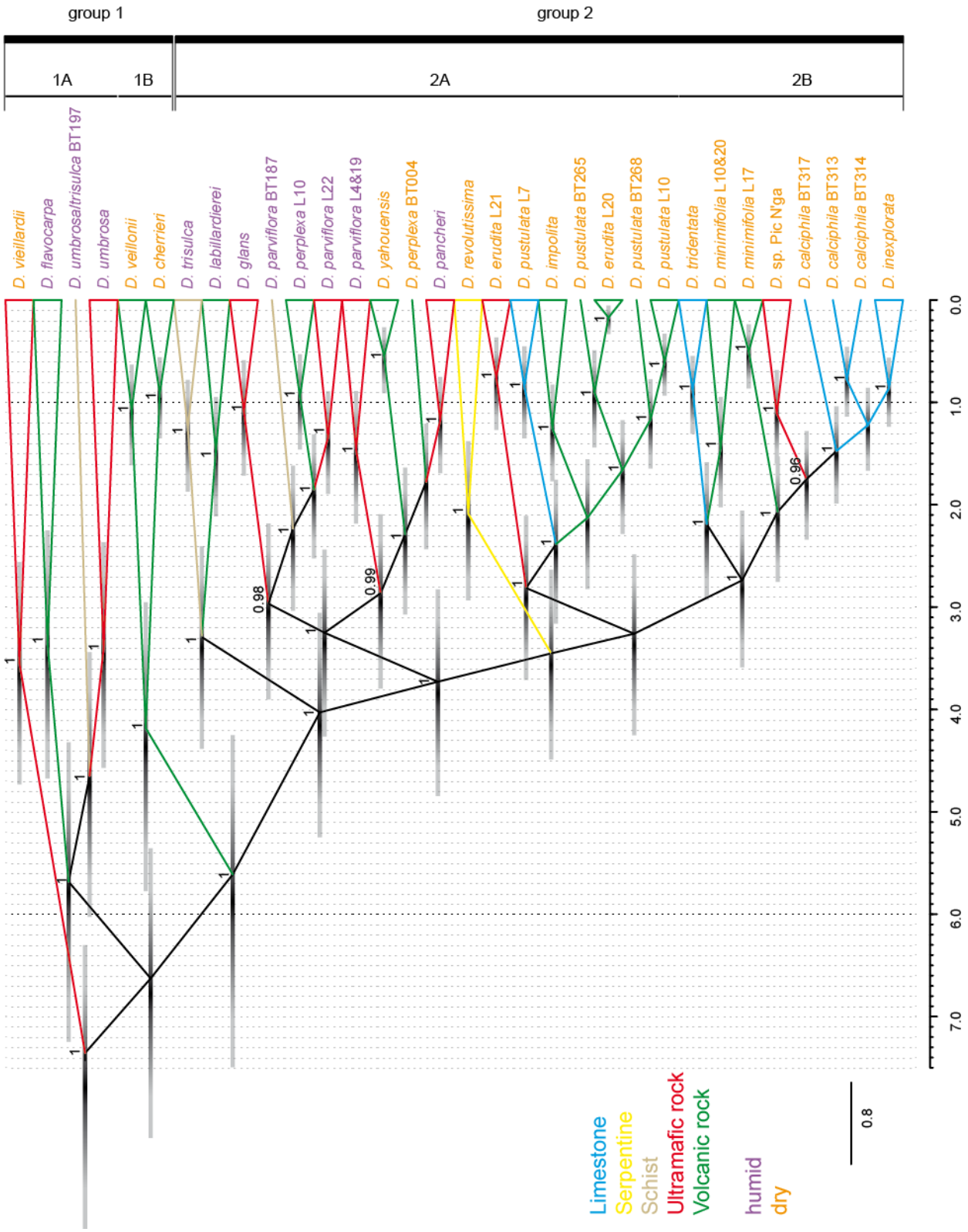


Figure 3: Phylogenetic tree of the radiating *Diospyros* group on New Caledonia derived from Bayesian inference. For simplicity, individuals are collapsed to species/population level wherever possible. Branch-colours indicate soil type preference; colour of taxa names gives the preference for climate type. Node bars indicate the 95% confidence interval for the age of the corresponding node; posterior probabilities are given for nodes with PP > 0.95. A time scale is given at the bottom of the figure. Abbreviations: L: location; refers to location number given in figure 1.

Genetic clustering and patterns of reticulation

The SPLITSTREE network (Fig. 4) follows the general pattern shown by the MP and BI trees. The branches within group 1 and between group 1 and 2 are significantly longer than those within group 2. Group 2 shows a reticulate and putatively hybridogenic history. Conflicting information, indicative of hybridization or incomplete lineage sorting, is also apparent at the level of deeper relationships within group 1.

STRUCTURE gave the highest delta K value (ΔK) for K = 2 followed by sub-optimal ΔK peaks for K = 4 and K = 6 (Fig. 5). The analysis considering two groups (K = 2) separated the species into the older groups (1A, red, Fig. 5C) and younger groups (2, blue, Fig. 5C) as well as an "admixed" group 1B between the other two. The four groups defined for K = 4 (Fig. 5D), reassemble the two groups (1 and 2) with their two sub-groups (A and B) described earlier. In this analysis *D. flavocarpa* and *D. umbrosa* seem to be slightly admixed between *D. veillardii* (red cluster in Fig. 5D) and group 1B (orange cluster in Fig. 5D). Within group 2A (blue cluster in Fig. 5D), individuals of *D. erudita* L20, *D. labillardierei*, *D. parviflora* L22, *D. perplexa* L10 and *D. trisulca* are "pure", and the rest of the individuals is partly admixed with group 2B (purple in Fig. 5D). The analysis considering six groups (K = 6) is only weakly supported and two of the groups (yellow and green, Fig. 5E) found in this analysis do not contain any "pure" individual.

The first coordinate of the PCO (Fig. 6), explaining 70% of the variation in the dataset, separated *D. vieillardii* as well as *D. flavocarpa* and *D. umbrosa* from the rest of the species. *Diospyros cherrieri* and *D. veillonii* are separated from the group containing the majority of the species (group 2) along the third coordinate, summarizing 6% of variation in the data.

Patterns of convergent adaptive divergence

Tests for particular genomic regions, which have been systematically involved in divergences between sister species with distinct preference for ultramafic versus volcanic substrates, resulted in 50 regions with pairwise F_{ST} values over 0.5 for at least two such pairs. Four regions have been found to be significantly different in three of the four pairs of species. Functional annotations have been successful for only 15 of the regions. The combined graph analysis of Blast2Go indicates enrichment for regions with molecular functions localized at the membrane level, related to intracellular transport, molecular binding and catalytic activity (Fig. 7).

DISCUSSION

To resolve shallow phylogenetic relationships within a rapidly radiating *Diospyros* group on New Caledonia we employed the RADseq technique, because it combines the advantages

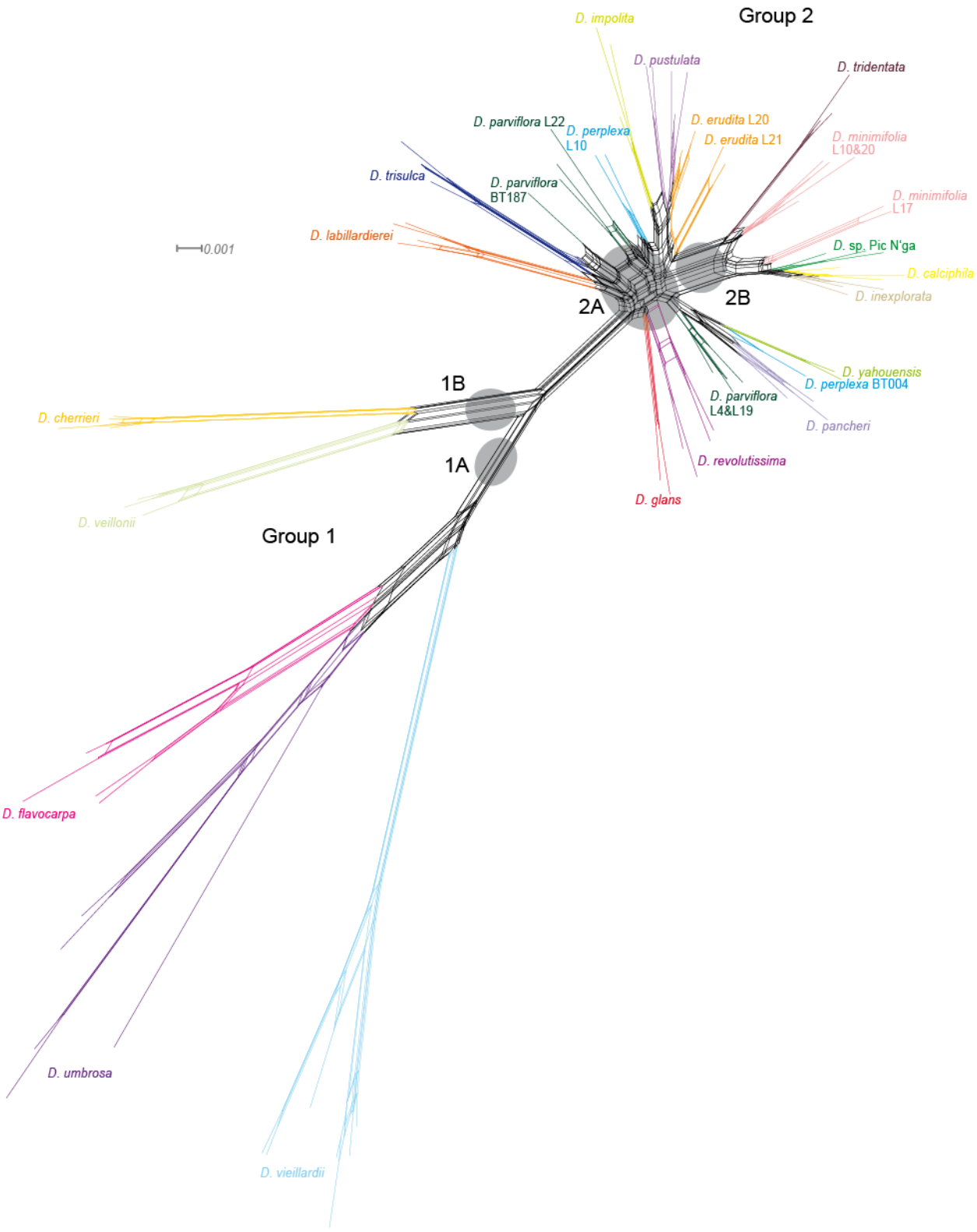


Figure 4: Neighbour-joining network based on Hamming distance. Each species is shown in a different colour.

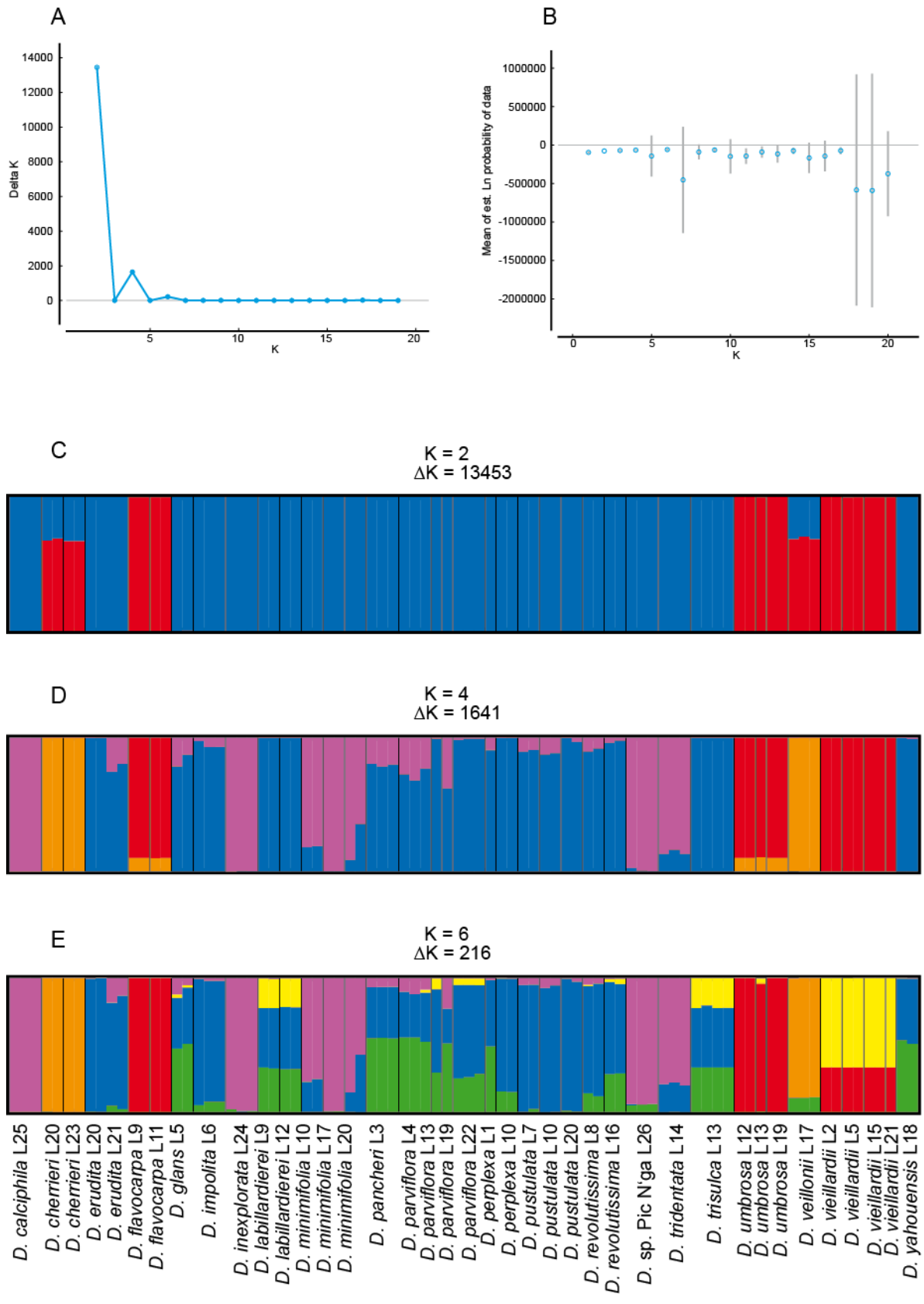


Figure 5: Structure results. Abbreviations: L: location; refers to location number given in figure 1.
A) Delta K values of the K values tested.
B) Mean Ln likelihood of K values tested.
C) Clustering of K = 2
D) Clustering of K = 4
E) Clustering of K = 6

of classical sequencing (i.e., confidence of homology), AFLP (i.e., genome-wide sampling of information), and next generation sequencing (i.e. high-throughput). Using thousands of SNPs derived from over 30,000 RAD loci assembled *de novo* from Illumina reads, we could infer much better resolved trees than previous trees based on multiple gene sequences (Duangjai et al. 2009; Turner et al. 2013a) and genome-wide fingerprinting analyses (Turner *et al*. 2013b). An increase in phylogenetic resolution when using RADseq in comparison with more traditional methods has been already shown for some organisms, for example, the adaptive radiation of cichlid fishes in Lake Victoria (Keller *et al*. 2013; Wagner *et al*. 2013), *Pedicularis* section *Cyathophora* (Eaton and Ree 2013), and the pitcher plant mosquito, *Wyeomyia smithii* (Emerson *et al*. 2010). Despite the high number of informative loci investigated, relationships inferred here for *Diospyros* are not always well supported, indicating that processes blurring phylogenetic signals, such as interspecific hybridization and/or incomplete lineage sorting, may have been common during some episodes of speciation in this group. Because it requires the presence or accumulation in time of a rich ancestral genetic pool (van Oppen *et al*. 2001; Maddison and Knowles 2006; Glor 2010; Lerner *et al*. 2011), we consider rather improbable that incomplete lineage sorting has significantly affected, on a genome-wide scale, phylogenetic patterns within this group, which radiated rapidly after a single and fairly recent long-distance dispersal event (Duangjai *et al*. 2009), their early history most probably associated with an extreme genetic bottleneck.

Genetic structure and gene flow

The SPLITSTREE and STRUCTURE analyses provide evidence for a fair amount of admixture between species (Fig. 5C). In K = 2, seven out of 84 individuals are admixed (considered hereafter individual with between 5 and 95% membership to one cluster). These admixed individuals are members of *D. cherierri* and *D. veillonii* (1B). In the phylogenetic network, these two species are positioned between species of 1A and 2 (Fig. 4), whereas in the PCO (Fig.6) they cluster outside of but close to the species cluster 2. Both BI and MP cluster group 1B together with group 2, and this relationship receives maximum support. In a further STRUCTURE analysis (K = 4) 36 individuals are admixed, especially those from group 2A. In the phylogenetic trees (MP and BI, Figs. 2 and 3), most relationships between populations of group 2A receive no support. In the network analysis, populations of group 2 exhibit a major reticulation (Fig. 4), confirming the presence of some gene flow. The possibility of speciation in the face of gene flow has been previously reported, for example, in the case of *Heliconius* butterflies (Martin *et al*. 2013), whereas a creative role of hybridization during rapid radiation is generally accepted (e.g., Seehausen 2004; Gavrilets and Losos 2009; Losos 2010; Glor 2010).

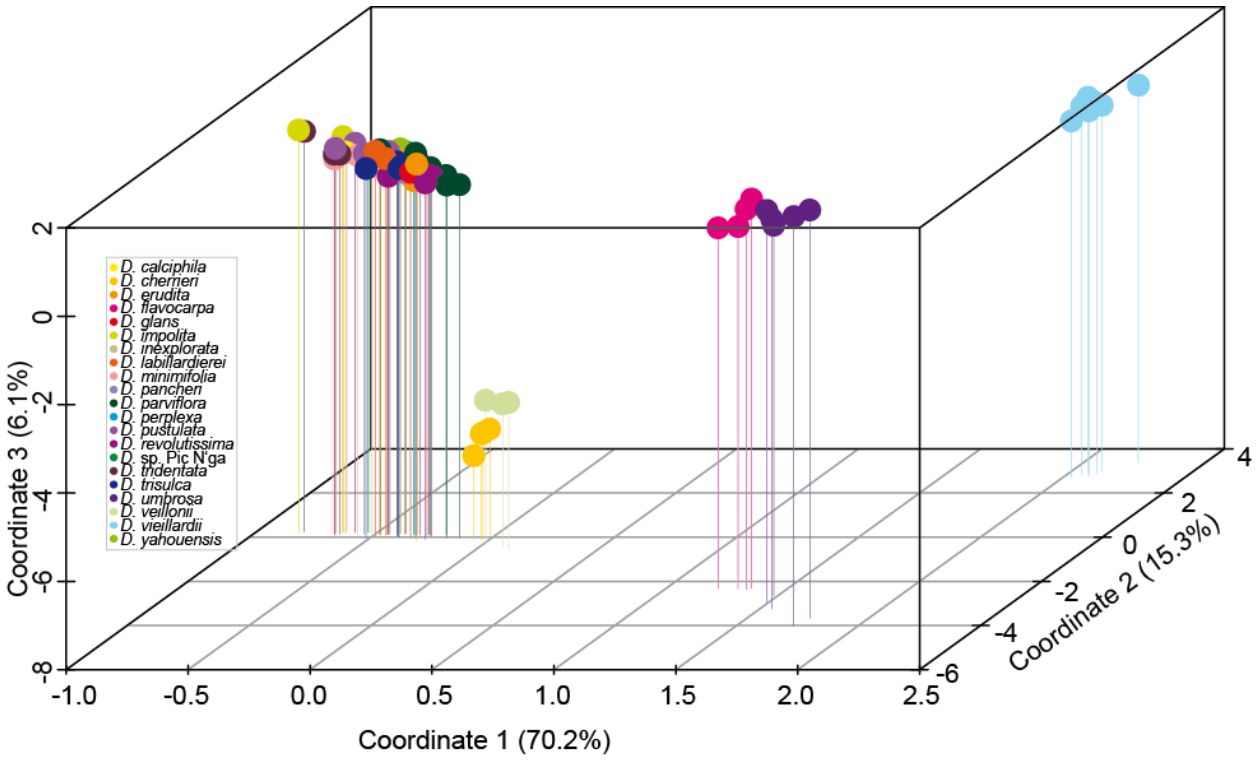


Figure 6: Principal coordinates analysis based on Hamming distances. Each species is shown in a different colour.

Clustering patterns of the species on the island

The inferred phylogenetic relationships in *Diospyros* point to some regional clustering among populations and species. In particular, groupings of species within group 2A follow mainly geographic patterns. The phylogenetic relationships between *D. erudita*, *D. impolita* and *D. pustulata* remain unclear, but they form a supported group (PP 1) of species found in dry, non-dense forests in the middle western part of Grande Terre. A similar geographic pattern is observed for *D. pancheri*, *D. yahouensis* and the accession of *D. perplexa* from L1; they have been all collected in southern Grande Terre (Fig. 1). We observe that the grouping of different populations of non-clustering species like *D. minimifolia*, *D. perplexa* and *D. parviflora* has some relationship to the region of the island from where they came. The population of *D. minimifolia* from Gadji (L17, Fig. 1) is genetically different from the rest of the individuals of this species found in the middle of Grande Terre. This population from Gadji clusters with species from Île des Pins (L25 and L26, Fig. 1) and Île Kuebini (L24, Fig. 1), which are all in the south of New Caledonia. Accessions of *D. parviflora* and *D. perplexa* collected around the central region of New Caledonia (Fig. 1) form a highly supported group, whereas the southern populations of *D. parviflora* fall in a unique cluster. This phenomenon of individuals grouping with co-occurring species rather than with populations of the same species but from different localities is also found in other organisms (e.g. *Heliconius*, Martin *et al*. 2013) and may be indicative of ongoing local gene flow.

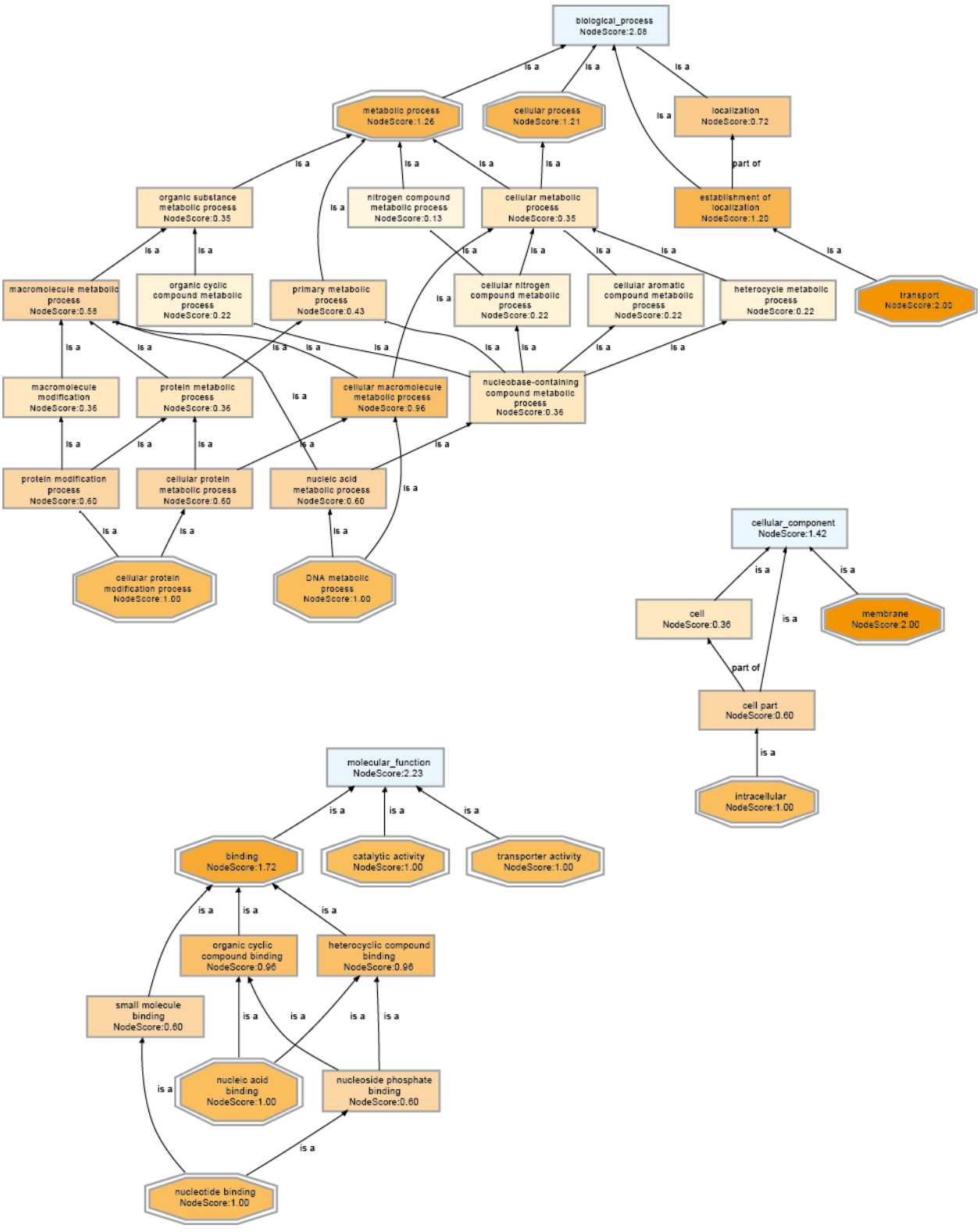


Figure 7: Blast2Go-derived combined graph for the annotated genomic regions hypothesized to be involved in divergent adaptation to substrate type.

Individuals of *D. calciphila* and *D. inexplorata* could not be clearly separated in any of the analyses. Both species occur in regions with similar ecological conditions (forests on calcareous substrate along the coast) and they are morphologically similar. *Diospyros calciphila* is described from the islands surrounding the main island and *D. inexplorata* is found only in one locality the south of the main island. We consider it to be likely that these individuals represent the same species.

Adaptive radiation and age

The species-rich New Caledonian *Diospyros* clade is the result of rapid radiation (Turner *et al*. 2013a, 2013b) resulting in more than 20 morphologically and ecologically diversified species with low genetic divergence. Our results suggest that both sympatric ecological divergence and allopatric diversification (i.e., resulting in regional patterns of diversity) shaped successive rounds of speciation in the *Diospyros* radiation. To further investigate the molecular targets of natural selection during parallel divergence (Stern 2013) in substrate preference, we searched for loci that are divergent (high F_{ST} values) between sister taxa occurring on different soil-types. Most of the annotated15 loci correspond to genes involved in transporting and binding through/to the cell membrane. As the New Caledonian soil-types are different in heavy-metal content and availability of mineral nutrients, these specific adaptations in binding and transporting substances to/through the cell membrane appear meaningful. It is, however, difficult to argue that this differentiation is responsible for particular speciation events or if it has evolved later. A similarly limited number of genomic regions on which positive selection might have acted has also been found in Hawaiian species of *Schiedea* (Kapralov *et al*. 2013), which exhibit, like the New Caledonian *Diospyros* species, great morphological and ecological variation.

Not much information is available about generation time of *Diospyros*, but the literature (Verdú 2002) suggests something like seven years, which seems to be a reasonable time for the New Caledonian *Diospyros* species. Taken this generation time and the age of the closely related group (group 2, around 4 mya) together, we can estimate that since the divergence of this group maximum 500-600 k generations have been present up to now. This low number of generations is one reason for the low sequence divergence observed among the New Caledonian *Diospyros* species and was probably not long enough for them to become permanently reproductively isolated (Gaudeul *et al*. 2012).

ACKNOWLEDGEMENTS

This work was funded by a grant from the Austrian Science Fund (FWF, Project-Number: P 22159-B16) awarded to R. Samuel. Thanks to F Balao Robles for his help with bioinformatics

issues. Specimens are deposited in the herbaria of Noumea (NOU), University of Montpellier II (MPU) and the University of Vienna (WU).

REFERENCES

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson EA. 2008. Rapid SNP discovery and genetic mapping using sequences RAD markers. PLoS ONE 3: e3376.

Barchi L, Lanteri S, Portis E, Aacquadro A, Valè G, Toppino L, Rotino GL. 2011. Identification of SNP and SSR markers in eggplant sing RAD tag sequencing. BMC Genomics 12: 304.

Baxter SW, Davey JW, Johnston JS, Shelton AM, Heckel DG, Jiggins CD, Blaxter ML. 2011. Linkage mapping and comparative genomics using next-generation RAD sequencing of a non-model organism. PLoS ONE 6: e19315.

Cariou M, Duret L, Charlat S. 2013. Is RAD-seq suitable for phylogenetic inference? An in silico assessment and optimization. Ecology and Evolution 3: 846-852.

Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. Nature Methods 9: 772.

Drummond AJ, Ho SYW, Phillips MJ, Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. PLoS Biology 4: 699-710.

Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution 29: 1969-1973.

Duangjai S, Samuel R, Munzinger J, Forest F, Wallnöfer B, Barfuss MHJ, Fischer G, Chase MW. 2009. A multi-locus plastid phylogenetic analysis of the pantropical genus *Diospyros* (Ebenaceae), with an emphasis on the radiation and biogeographic origins of the New Caledonian endemic species. Molecular Phylogenetics and Evolution 52: 602-620.

Earl DA, vonHoldt BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4: 359-361.

Eaton DAR, Ree RH. 2013. Inferring phylogeny and introgression using RADseq data: an example from flowering plants (*Pedicularis*: Orobanchaceae). Systematic Biology 62: 789-706.

Emerson KJ, Merz CR, Catchen JM, Hohenlohe PA, Cresko WA, Bradshaw WE, Holzapfel CM. 2010. Resolving postglacial phylogeography using high-throughput sequencing. PNAS 107: 16196-16200.

Etter PD, Preston JL, Bassham S, Cresko WA, Johnson EA. 2011. Local *de novo* assembly of RAD paired-end contigs using short sequencing reads. PLoS ONE 6: e18561.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14: 2611-2620.

Gaudeul M, Rouhan G, Gardner MF, Hollingsworth PM. 2012. AFLP markers provide insights into the evolutionary relationships and diversification of New Caledonian *Araucaria* species (Araucariaceae). American Journal of Botany 99: 68-81.

Gavrilets S, Losos JB. 2009. Adaptive radiation: contrasting theory with data. Science 323: 732-737.

Gernhard T. 2008. The conditioned reconstructed process. Journal of Theoretical Biology 253: 769-788.

Glor RE. 2010. Phylogenetic insights on adaptive radiation. Annual Review on Ecology, Evolution and Systematics 41: 251-270.

Götz, S., García-Gómez, J.M., Terol, J., Williams, T.D., Nueda, M.J., Robles, M., Talón, M., Dopazo, J., and Conesa, A. (2008) High-throughput functional annotation and data mining with the Blast2GO suite Nucleic Acids Research 36: 3420–3435.

Guindon S, Gascuel O. 2003. A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. Systematic Biology 52: 696-704.

Hamming RW. 1950. Error detecting and error correcting codes. The Bell System Technical Journal 29: 147-160.

Hubisz MJ, Falush D, Stpehens M, Pritchard JK. 2009. Inferring weak population structure with the assistance of sample group information. Molecular Ecology Resources 9: 1322-1332.

Huson DH, Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. Molecular Biology and Evolution 23: 254-267.

Huson DH, Scornavacca C. 2011. A survey of combinatorial methods for phylogenetic networks Genome Biology and Evolution 3: 23-35.

Ibanez T, Munzinger J, Dagostini G, Hequet V, Rigault F, Jaffré T, Birnbaum P. 2013. Structural and floristic diversity of mixed tropical rainforest in New Caledonia: New data from the New Caledonian Plant Inventory and Permanent Plot Network (NC-PIPPN). Applied Vegetation Science: DOI: 10.1111/avsc.12070.

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

Kane NC, King MG, Barker MS, Raduski A, Karrenberg S, Yatabe Y, Knapp SJ, Rieseberg LH. 2009. Comparative genomic and population genetic analyses indicate highly porous genomes and high levels of gene flow between divergent *Helianthus* species. Evolution 63: 2061-2075.

Kapralov MV, Votintseva AA, Filatov DA. 2013. Molecular adaptation during a rapid adaptive radiation. Molecular Biology and Evolution 30: 1051-1059.

Keller I, Wagner CE, Greuter L, Mwaiko S, Selz OM, Sivasundar A, Wittwer S, Seehausen O. 2013. Population genomic signatures of divergent adaptation, gene flow and hybrid speciation in the rapid radiation of Lake Victoria cichlid fishes. Molecular Ecology 22: 2848–2863.

Lerner HRL, Meyer M, James HF, Hofreiter M, Fleischer RC. 2011. Multilocus resolution of phylogeny and timescale in the extant adaptive radiation of Hawaiian honeycreepers. Current Biology 21: 1838-1844.

Lexer C, Widmer A. 2008. The genic view of plant speciation: recent progress and emerging questions. Philosophical Transaction of the Royal Society of London, series B, Biological Sciences 363: 3023-3036

Ligges U, Mächler M. 2003. Scatterplot3d - an R Package for visualizing multivariate data. Journal of Statistical Software 8: 1-20.

Maddison WP, Knowles LL. 2006. Inferring phylogeny despite incomplete lineage sorting. Systematic Biology 55: 21-30.

Mallet J. 2005. Hybridization as an invasion of the genome. Trends in Ecology and Evolution 20: 229-237.

Martin SH, Dasmahapatra KH, Nadeau NJ, Slazar C, Walters JR, Simpson F, Blaxter M, Manica A, Mallet J, Jiggins CD. 2013. Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. Genome Research 23: 1817-1828.

Maurizot P, Vendé-Leclerc M. 2009. New Caledonia geological map, scale 1/500000. Direction de l'Industrie, des Mines et de l'Energie - Service de la Géologie de Nouvelle-Calédonie, Bureau de Recherches Géologiques et Minières.

Miller MA, Pfeiffer W, Schwarz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees, in: Proceedings of the Gateway Computing Environments Workshop (GCE), 14 Nov. 2010, New Orleans, Louisiana, pp. 1-8.

Miller MR, Dunham JP, Amores A, Cresko WA, Johnson EA. 2007. Rapid and cost effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. Genome Research 17: 240-248.

Pelletier B. 2006. Geology of the New Caledonia region and its implications for the study of the New Caledonian biodiversity. In: Payri C, Richer de Forges B. (Eds.), Compendium of marine species from New Caledonia, Documents Scientifiques et Techniques II4, Institut de Recherche pour le Développement Nouméa, pp. 17-30.

Posada D. 2003. Using MODELTEST and PAUP* to select a model of nucleotide substitution. Current Protocols in Bioinformatics: DOI: 10.1002/0471250953.bi0605s00.

Pritchard JK, Stephens M, Donnely P. 2000. Inference of population structure using multilocus genotype data. Genetics 155: 945-959.

Rieseberg LH. 2009. Evolution: replacing genes and traits through hybridization. Current Biology 19: R119-R122.

Rosenberg NA. 2004. Distruct: a program for the graphical display of population structure. Molecular Ecology Notes 4: 137-138.

Rubin BER, Ree HE, Moreau CS. 2012. Inferring phylogenies from RAD sequence data. PLoS ONE 7: e33394.

Seehausen O. 2004. Hybridization and adaptive radiation. Trends in Ecology and Evolution 19: 198-207.

Stern DL. 2013. The genetic causes of convergent evolution. Nature Reviews Genetics 14: 751-764.

Swofford DL. 2003. PAUP*. Phylogenetic analysis using parsimony (*and other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Tel-Zur N, Abbo S, Myslabodski D, Mizrahi Y. 1999. Modified CTAB procedure for DNA isolation from epiphytic cacti of genera *Hylocereus* and *Selenicereus* (Cactaceae). Plant Molecular Biology Reporter 17: 249-254.

Turner B, Munzinger J, Duangjai S, Temsch EM, Stockenhuber R, Barfuss MHJ, Chase MW, Samuel R. 2013a. Molecular phylogenetic of New Caledonian *Diospyros* (Ebenaceae) using plastid and nuclear markers. Molecular Phylogenetics and Evolution 69: 740-763.

Turner B, Paun O, Munzinger J, Duangjai S, Chase MW, Samuel R. 2013b. Amplified fragment length polymorphism (AFLP) data suggest rapid radiation of *Diospyros* species (Ebenaceae) endemic to New Caledonia. BMC Evolutionary Biology 13: 269.

Van Oppen MJH, McDonald BJ, Willis B, Miller DJ. 2001. The evolutionary history of the coral genus *Acropora* (Scleractinia, Cnidaria) based on a mitochondrial and a nuclear marker: reticulation, incomplete lineage sorting, or morphological convergence? Molecular Biology and Evolution 18: 1315-1329.

Vandepitte K, Honnay O, Mergeay J, Breyne P, Roldán-Ruiz I, De Meyer T. 2013. SNP discovery using paired-end RAD-tag sequencing on pooled genomic DNA of *Sisymbrium austriacum* (Brassicaceae). Molecular Ecology Resources 13: 269-275.

Verdú M. 2002. Age at maturity and diversification in woody angiosperms. Evolution 56: 1352-1361.

Vos P, Hogers R, Bleeker M, Reijans M, Van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research 23: 4407-4414.

Wagner CE, Keller I, Wittwer S, Selz OM, Mwaiko S, Greuter L, Sivasundar A, Seehausen O. 2013. Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. Molecular Ecology 22: 7878-798.

Yule GU. 1925. A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F.R.S. Philosophical Transactions of the Royal Society of London Series B 213: 21-87.

CHAPTER 4

Characterization of nuclear and plastid genomes of *Diospyros* species endemic to New Caledonia by low-coverage next generation sequencing

Barbara Turner, Hanna Schneeweiss, Ovidiu Paun, Jérôme Munzinger, Mark W. Chase, Rosabelle Samuel

Status: in prep for BMC Genomics

Contribution: Collection of material, collection of data, analysis of data, phylogenetic analysis, manuscript writing/editing

INTRODUCTION

Genome size and polyploidy

Genome size varies nearly 2,400-fold across angiosperms (Leitch and Leitch 2013), and most variation in DNA amount is caused by different amounts of noncoding, repetitive DNA, mostly retrotransposons and tandem repeats of satellite DNA (Leitch 2007; Bennett and Leitch 2005; Parisod *et al*. 2009; Petrov 2001). Previous studies showed that polyploidy is altogether rare in *Diospyros* (White 1992; Tamura *et al*. 1998; Turner *et al*. 2013a), although some cultivated species are polyploids (e.g. *D. rhombifolia* 4*x*, *D. ebenum* 6*x*, *D. kaki* 6*x* and 9*x*, *D. virginiana* 6*x* and 9*x*; Tamura *et al*. 1998). The basic chromosome number in *Diospyros* is *x* = 15. Investigations of New Caledonian *Diospyros* species revealed a continuous variation in genome size, and chromosome counts showed that the investigated species are diploid (2n = 30, Fig. 1; Turner *et al*. 2013a).

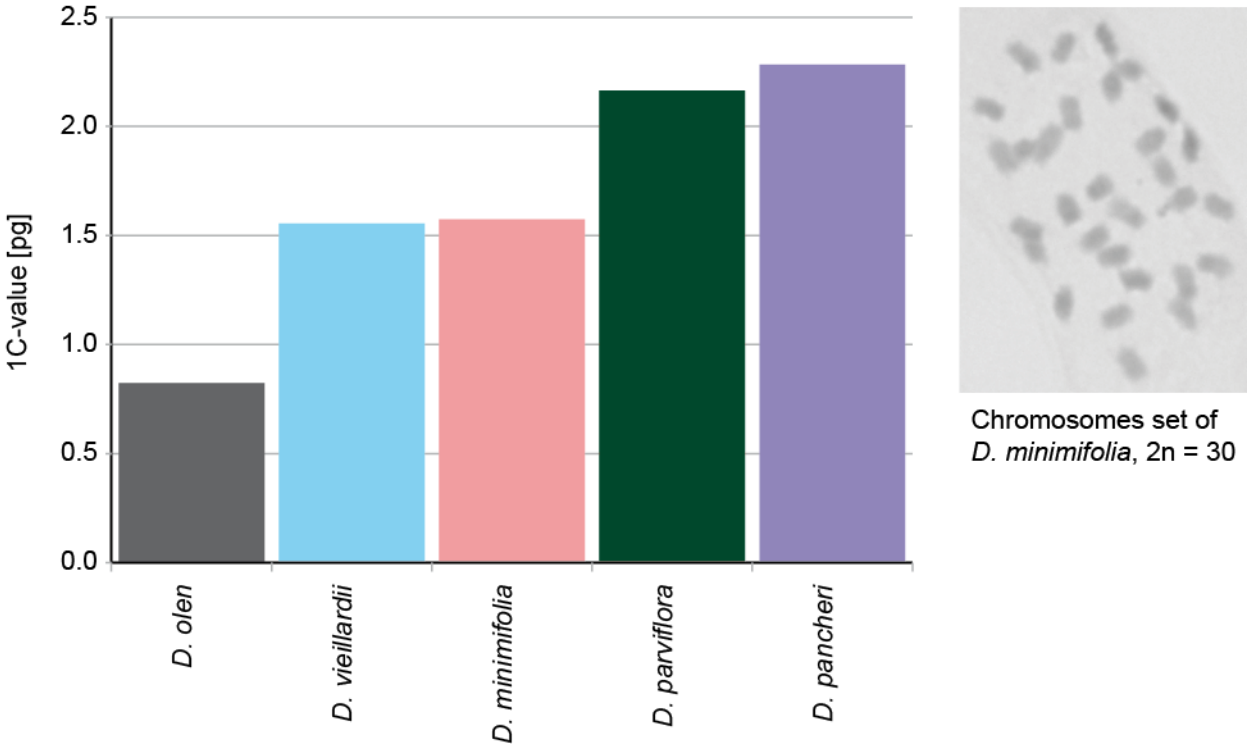


Figure 1: Genome sizes of *Diospyros* species investigated and chromosomes set of *Diospyros minimifolia*.

Repetitive elements

The repetitive fraction of a genome has for long been seen as junk or parasitic DNA (e.g. Doolittle and Sapienza 1980; Orgel and Crick 1980; Kubis *et al*. 1998; Volff 2006; Kejnovsky *et al*. 2012), although the repetitive DNA can make up to over 80 % of a genome (Keith *et al*. 2013). Generally the repetitive DNA elements can be classified to two main groups

distinguished by their genomic organisation (Kubis *et al*. 1998). Satellite DNAs are arranged in tandem arrays of monomers. Transposable elements (TEs) are major components of eukaryotic genomes (Wicker *et al*. 2007). TEs are present in many copies in the nuclear genome and can constitute a significant portion of the host genome, especially in plants (Oliver *et al*. 2013). Their genetic structure and replication/transposition mode, allows distinction of two main classes of mobile elements (Wicker *et al*. 2007). Class I elements (retrotransposons, REs) transpose via an RNA intermediate being transcribed from a genomic copy and reverse-transcribed into DNA by a TE-encoded reverse transcriptase (Wicker *et al*. 2007). New DNA copies insert at new locations of the host genome. Each replication cycle produces large numbers of new copies, making retrotransposons major contributors of the repetitive fraction of plant genomes (Kelly *et al*. 2012). One of their types, long terminal repeat REs (LTR-REs; Ty1/*copia* and Ty3/*gypsy*) are particularly abundant in plant genomes. Class II elements (DNA transposons) move within genomes using a “cut and paste” mechanism (mediated by transposase) and are less abundant in plants.

The activity of TEs has been inferred to fluctuate across evolutionary time. Higher activity of TEs has been observed to correlate with elevated levels of stress, including environmental changes (e.g. temperature and humidity), presence of toxins or other chemicals and interactions with other organisms (Oliver and Greene 2009; Casacuberta and González 2013). An increase in TE activity has also been observed in the first generations after polyploidization in allo- and autopolyploid species (e.g. Paun *et al*. 2007; Parisod *et al*. 2010; Renny-Byfield *et al*. 2011), which are frequent in angiosperms. These data suggest that TEs might play a role in adaptation to new environmental conditions and might participate in large-scale genomic alterations.

Next generation sequencing

Next generation sequencing methods offer the possibility to generate large amounts of data at low cost (Glenn 2011). Thus far, there are several NGS methods available ranging from whole-genome sequencing to amplicon sequencing of PCR generated fragments (i.e. fragments enriched by PCR amplification; for a general review of NGS applications see McCormack *et al*. 2013). Illumina technology is a frequently used NGS method because it is the cheapest (cost per bp; Glenn 2011). Illumina platforms can be also used to sequence the repetitive fractions of a genome at a low coverage. Such data allow characterisation of DNA sequences present in the target genomes in high copy numbers, such as, transposable elements and tandem repeats. Transposable element content inferred from NGS data have been previously used to infer genome evolution in phylogenetic questions context (e.g. Piednoël *et al*. 2012), to test phylogenetic relationships (Dodsworth *et al*. submitted) and characterize the evolutionary dynamics of genomes (e.g. Natali *et al*. 2013; Sveinsson *et al*. 2013; Renny-Byfield *et al*. 2011, 2012, 2013).

Organelle genomes

Whole-genome sequencing generates not only sequences from the nuclear genome of the individual investigated (unless only nuclear DNA has been subjected to sequencing), but also from organellar genomes (plastids and mitochondria). Sequences from the plastid genome have been extensively used to infer phylogenetic relationships among plants (e.g. Barfuss *et al*. 2005; Duangjai *et al*. 2009; Russell *et al*. 2010). Recently, whole plastid genome sequencing has become affordable and is used to generate phylogenies based on whole plastid genomes (e.g. Yang *et al*. 2013; Ku *et al*. 2013; Barrett *et al*. 2013, 2014). Sequences derived from mitochondria are not commonly used for phylogenetic reconstructions in plants, because of low sequence divergence, extensive recombination and mitochondrial genomic rearrangements (Hiesel *et al*. 1994; Kress *et al*. 2005).

Diospyros in New Caledonia

New Caledonia comprises an archipelago in the southern Pacific known for its characteristic, endemic flora (Lowry 1998). Due to a complex geological history, New Caledonia features a mosaic of different soil-types (Pelletier 2006; Maurizot and Vendé-Leclerc 2009), which in combination with its climatic heterogeneity results in many different habitats. *Diospyros* (Ebenaceae) is a large genus of woody plants found world-wide in the tropics and subtropics, including 31 species in New Caledonia. Previous studies based on plastid markers (Duangjai *et al*. 2009) showed that *Diospyros* colonised New Caledonia at least four times via long-distance dispersal. Two of the successful dispersal events resulted in one species each still surviving at present, a third led to a small clade comprising five species, and the fourth event gave rise to a group of 24 species. These 24 species are all endemic to New Caledonia and have been shown to be closely related using low-copy nuclear and plastid markers (Duangjai *et al*. 2009; Turner *et al*. 2013a). Data obtained from genome-wide RAD-sequencing, proved to be helpful to resolve phylogenetic relationships among the species (Chapter 3). Most of these closely related species are morphologically and ecologically clearly differentiated, and species delimitations were confirmed by analyses of AFLP (Turner *et al*. 2013b) and RAD sequencing data (Chapter 3). *Diospyros* species are found in many habitats, and in some localities several species co-occur in sympatry. Dating analysis based on combined plastid and nuclear DNA sequence data showed that the lineages forming this group of New Caledonian *Diospyros* species arrived in New Caledonia around nine million years ago (mya; Turner *et al*. 2013a). Taken into consideration that these are woody plants with generation times of several years, it becomes obvious that these are relatively recent evolved/radiated species.

Aims

The differences in genome size observed among four closely related endemic New Caledonian *Diospyros* species (1C-values: 1.6 – 2.3 pg) compared to *Diospyros olen* (also endemic to New Caledonia but from a different clade; 1C-value: 0.8 pg) despite their likely identical ploidy level make this system attractive to analyse the composition of the r repetitive DNA fraction in these genomes using NGS. Here we present first results, based on analyses of six different *Diospyros* species (Table 1) including five endemics from New Caledonia and a wide spread species found in the south and west Pacific (*D. vera*). Sequencing of 17 further species is in progress. The primary aim of this study is to characterize the repetitive fraction of the nuclear genomes and to identify elements that potentially are involved in genome size changes. Since low-coverage NGS delivers also whole-plastid genome sequences, plastid genomes are investigated to determine their information content for phylogenetic analyses.

MATERIAL AND METHODS

DNA was extracted from silica gel-dried leaf material using a modified sorbitol/high-salt CTAB method (Tel-Zur *et al*. 1999). Extracts were purified using the NucleoSpin gDNA clean-up kit (Marcherey-Nagel, Germany), according to manufacturer's protocol.

From each of the 6 samples, 200 ng DNA was sheared for 55 sec. using an ultra-sonicator (Covaris, Massachusetts, USA), resulting in mean fragment size of 400 bp. All 6 individuals were barcoded, pooled to reach an equal representation of each individual in the final library and paired end sequenced in one Illumina lane. Libraries were sequenced on an Illumina HiSeq as 100-bp paired-end reads. Library preparation, sequencing, and de-multiplexing of the raw data were performed by CSF (Campus Science Support Facilities, Vienna, Austria; http://www.csf.ac.at/facilities/ngs/). For further analysis, data sets corresponding to single individuals were used.

Analysis of repeated elements

Analyses were conducted with RepeatExplorer (Novák *et al*. 2013) on the online platform (http://repeatexplorer.umbr.cas.cz/). All reads were subjected to quality control. Only reads with at least 90% of all bases having a quality score of at least 10 were further processed. From these quality-filtered reads, only those for which both reads met the quality criteria were used for further analyses. Analyses of repeats were performed for each individual separately as well as for a combined data set, with a minimal overlap length of 55 bp for clustering and a minimal overlap of 40 bp for assembly. Reads were checked against the Viridiplantae RepeatMasker database to facilitate cluster annotation. In the combined data set equal genomic amounts (*ca*. 2% of the whole genome) were used from each individual. Clusters which could not be

annotated by RepeatExplorer were manually annotated using DOTTER (Sonnhammer and Durbin 1996) and BLASTN searches (Altschul *et al*. 1990).

Genomic proportions for the repeat-types observed were calculated by dividing the total number of reads found for each repeat type by the total number of reads obtained for the corresponding individual. To be able to compare the results obtained from the investigated species, the genomic proportions were multiplied with a correction factor considering the genome size of the corresponding species. As *D. pancheri* has the highest genome size, this value was used as reference (factor 1.00) and for the other species the correction factor was calculated by dividing the genome size of the respective species by the genome size of *D. pancheri*. For *D. vera* it was not possible establish genome size and therefore no such correction could be made.

Assembling and annotating plastid genomes

Reads originating from the plastid genome were filtered using a multistep and iterative in-house established pipeline (Paun, personal communication). First, the individual raw files were imported in the CLC GENOMIC WORKBENCH v. 6.5 (Qiagen) and trimmed them by quality at $p < 0.05$, retaining reads of at least 30 bp. Further the reads of *D. vera* were mapped on the plastid genome of *Camellia sinensis* (Theaceae, Ericales) retrieved from GenBank (GenBank accession number: KC143082.1). For this initial mapping settings with mismatch cost of 2, and insertion and deletion cost of 3, requiring at least 80% of a read is 90% similar to the target for each successful map were used. With these settings 403,980 reads mapped to the *Camellia* plastome. Those reads were filtered using FastQ.filter.pl (Rodriguez LM unpublished, available from http://enveomics.blogspot.co.at/2013/04/fastqfilterpl.html) and assembled in CLC GENOMIC WORKBENCH, with automatic optimization of the word and bubble sizes and updating the contigs after mapping back the reads. We obtained three contigs, which have been concatenated by aligning them to the *Camellia sinensis* reference sequence manually in the program BioEdit v7.1.11 (Hall 1999). Both inverted repeats (IR) were reconstructed together and duplicated to obtain a complete plastid genome.

The plastid genomes of the rest of the *Diospyros* species were obtained in a similar way, but mapping has been performed against the assembled *D. vera* genome. Finally, annotation of coding regions was performed using DOGMA (Wyman *et al*. 2004) using only *D. vieillardii* plastid genome. The circular plastid genome map was visualized with OGDRAW (Lohse *et al*. 2007).

Characteristics of reads like amount of reads obtained from each individual and GC-content were evaluated with FastQC (available form http://www.bioinformatics.babraham.ac.uk/projects/fastqc/).

Phylogenetic analyses

The alignment of the plastid genomes of *Diospyros* was used to construct a phylogenetic tree using parsimony algorithm. Statistical support for the topology was obtained by bootstrapping with 1,000 replicates. Parsimony analysis and bootstrapping were both performed using PAUP* v4b10 (Swofford 2003).

Table 1: Accessions used for the analyses and their characteristics.

Species	Ploidy level	1C-value [pg]	1C-value [Mb]	Total length of raw reads [Mb]	Genome coverage	GC content of raeds	No. of reads used for clustering analyses	No. of reads in annotated clusters*	No. of annotated clusters*
D. olen Hiern	2*n*	0.8	805.7	2771.0	3.4 x	38%	2,648,368	1,228,667	264
D. vera (Lour.) A.Chev.	-	-	-	2924.3	-	39%	824,360	434,403	204
D. vieillardii (Hiern) Kosterm.	2*n*	1.6	1519.8	2659.4	1.7 x	41%	4,036,514	2,735,319	320
D. minimifolia F.White	2*n*	1.6	1538.2	2657.4	1.7 x	41%	2,639,862	1,713,459	282
D. parviflora (Schltr.) Bakh.	2*n*	2.2	2116.4	2759.0	1.3 x	41%	1,981,364	1,224,806	257
D. pancheri Kosterm.	-	2.3	2232.7	2494.7	1.1 x	41%	1,160,854	720,874	170

* Clusters containing minimum 0.01% of the reads.

RESULTS AND DISCUSSION

Transposable elements

The number of raw reads obtained ranged from 24,966,688 (*D. pancheri*) to 29,242,716 (*D. vera*) (Tab. 1). Since the same amount of DNA (200 ng) was used for each individual for library preparation, the genome coverage varied from 1.1 x in *D. pancheri* (1C-value: 2.3 pg) up to 3.4 x in *D. olen* (1C-value: 0.8 pg). The number of clusters ranged from 170 (*D. pancheri*) to 320 (*D. vieillardii*). Thus 43.7% (*D. olen*) to 67.1% (*D. vieillardii*) of the reads could be annotated.

In *Diospyros*, the most frequently observed repetitive DNA elements are LTR-retroelements (Fig. 2). Among these, Ty3/*gypsy* elements are most abundant, as observed in other plant groups (e.g. tobacco, Renny-Byfield *et al*. 2013; sunflower, Natali *et al*. 2013; pea, Macas *et al*. 2007). In other analysed plant genomes either the Ty1/*copia* or the Ty3/*gypsy* elements prevail, however there is no clear trend that the dominating type of LTR elements has any correlation with phylogenetic grouping among the taxa, where they are observed. Among the Ty3/*gypsy* elements chromovirus- and Ogre-elements are the most common found elements (supplementary figure S1). The number of TEs is higher in the genomes of the endemic New Caledonian species (mean 52.3%), compared to *D. olen* (34.4%). Within the endemic New Caledonian *Diospyros* species, the species with larger genomes have proportionally more TEs than those with smaller genome sizes. *Diospyros vieillardii* and *D. minimifolia* (both with 1C values = 1.6 pg) show similar total amounts of repetitive DNA, but differ in the amounts of the different repeat types. *Diospyros parviflora* and *D. pancheri* (1C values = 2.2 - 2.3 pg) differ slightly in the amount of repetitive elements, with the satellite-elements contributing to most of the variation between these two species (Fig. 2). Due to the lack of genome size data from *D. vera* results for this species are only given as genome proportions.

Beside the differences in transposable elements content, differences in the amounts of reads annotated as tandem repeats (satDNAs) were observed (Tables 2 and 3). The highest proportions of satellites were found in *D. pancheri* and *D. vera* (Fig. 3B). Differences in amounts of satellite-repeats were also found in other plant groups between genera (e.g. Orobanchaeeae, Piednoël *et al*. 2012), species (e.g. *Nicotiana*, Renny-Byfield *et al*. 2013; *Musa*, Čížková *et al*. 2013) as well as between male and female individuals in dioeceious plants (e.g. *Rumex acetosa*; Steflova *et al*. 2013).

Altogether there is an increase in the amount of repeated elements among the *Diospyros* species from the closely related group in comparison to *D. olen*. This increase in REs and their potential activity within the genomes could be associated with the rapid radiation of this group of New Caledonian *Diospyros* species.

Figure 2: Repat type content [Gbp] (corrected for genome size)

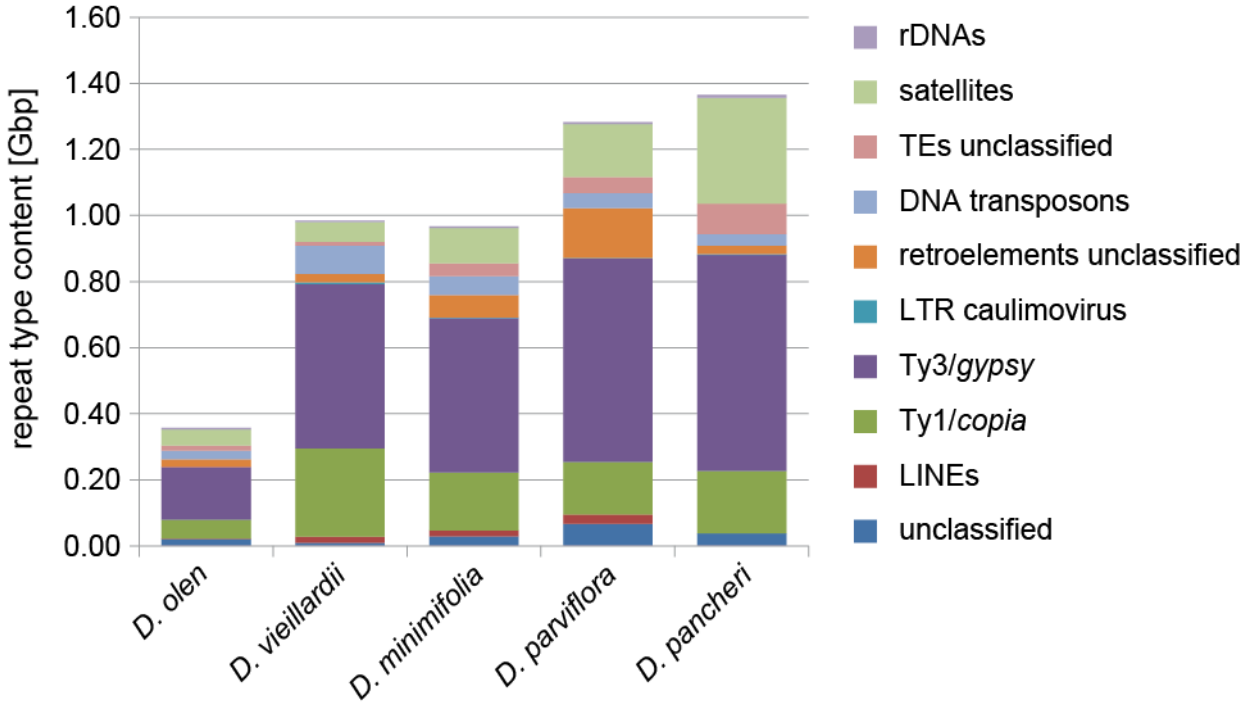


Figure 3: Phylogenetic tree of *Diospyros* species based on whole plastid genome sequences (A) and genomic proportions of repeated elements (B).

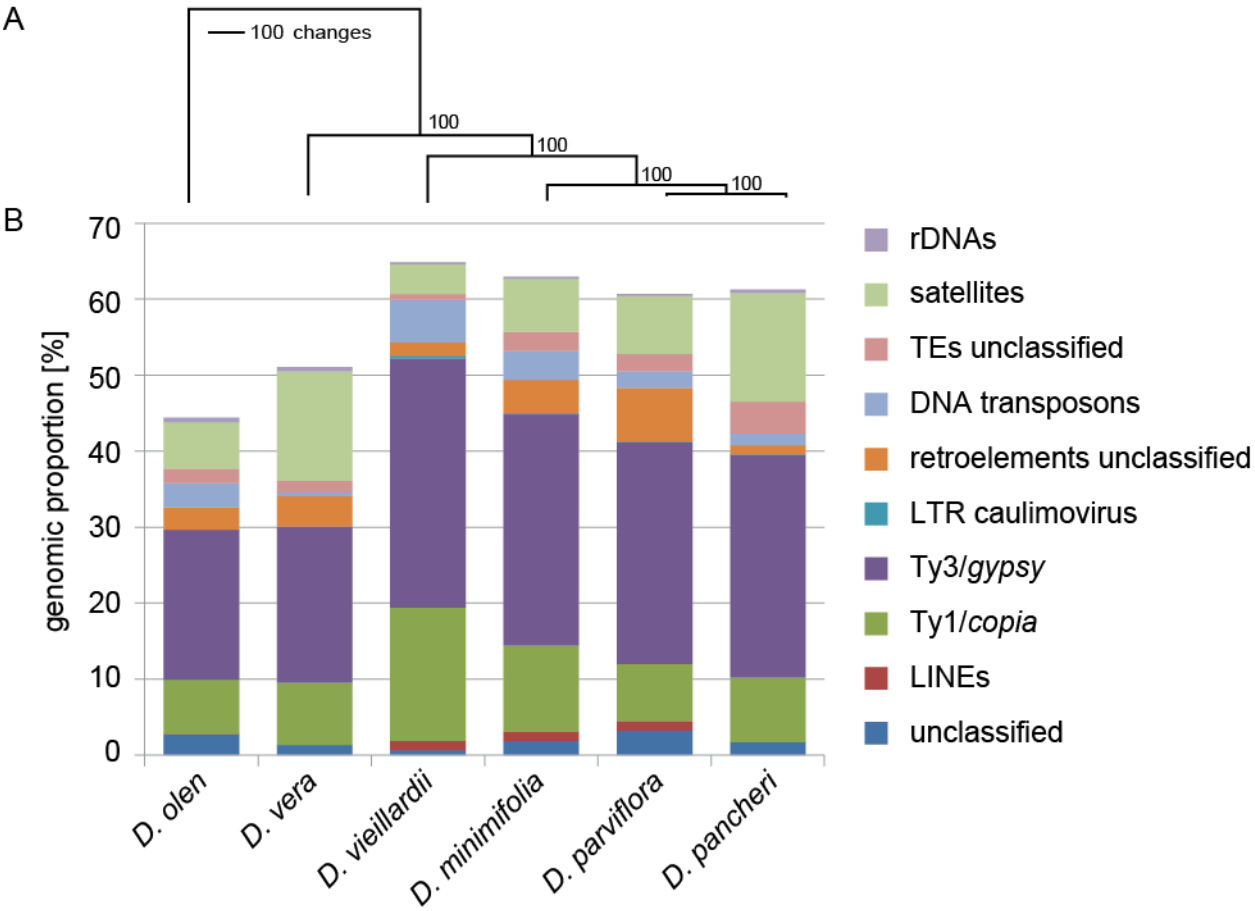


Table 2: Genomic proportions [%] of different repeats detected in analyzed species of *Diospyros*.

	D. olen	*D. vera*	*D. vieillardii*	*D. minimifolia*	*D. parviflora*	*D. pancheri*
Retroelements						
LTRs						
Ty1/*copia*	**7.14**	**8.20**	**17.52**	**11.43**	**7.51**	**8.51**
Alell	0.70	-	2.97	0.95	0.05	-
Angela	1.95	3.34	2.74	2.85	2.54	2.71
Bianca	0.05	-	-	-	-	-
Ivana	0.86	-	0.96	0.66	-	-
Maximus	1.25	2.00	2.36	2.60	2.98	2.38
TAR	0.30	0.15	0.32	0.28	0.25	0.21
Tork	0.98	0.03	1.74	0.69	0.50	0.59
unclassified	1.06	2.68	6.43	3.40	1.20	2.62
Ty3/*gypsy*	**19.66**	**20.52**	**32.78**	**30.36**	**29.14**	**29.31**
Athila	3.66	1.84	4.03	3.30	4.45	2.94
chromo	9.92	8.32	13.12	12.14	10.16	10.63
Ogre	4.31	8.47	10.27	10.98	10.87	12.25
unclassified	1.77	2.40	5.36	3.94	3.66	3.48
Others						
LTR Caulimovirus	-	-	0.38	0.18	0.11	0.09
Retroelement unclassified	2.92	4.01	1.74	4.41	7.09	1.14
non-LTRs						
LINEs	**0.14**	**-**	**1.22**	**1.15**	**1.28**	**-**
L1	0.03	-	0.11	0.13	0.97	-
unclassified	0.01	-	1.11	0.99	-	-
MITEs						
MITEs	-	0.03	0.07	-	0.03	-
TRIMs						
Cassandra	0.09	-	0.16	-	0.17	-
non-LTR retroelement						
unclassified	0.02	-	-	-	-	-
DNA transposons						
Subclass 1	**3.25**	**0.50**	**5.59**	**3.74**	**2.17**	**1.57**
DNA/CMC	-	-	0.32	-	-	-
DNA/EnSpm	-	-	1.17	-	-	-
DNA/CMC-EnSpm	0.79	0.07	-	1.07	1.12	0.43
DNA/hAT-AC	0.65	0.34	0.46	0.40	0.07	0.33
DNA/hAT-Tag1	0.58	-	0.04	0.17	0.11	0.03
DNA/hAT-Tip100	0.17	0.05	0.28	0.13	0.14	-
DNA/MULE	0.41	0.04	1.77	0.65	0.37	0.20
DNA/PIF	0.66	-	0.25	0.14	0.03	-
DNA transposons unclassified	0.04	-	1.24	1.16	0.33	0.52
Subclass 2						
Helitron	-	0.01	0.26	-	0.22	0.19
Other TEs						
TEs unclassified	1.87	1.53	0.78	2.52	2.26	4.16
Other repeats						
Tandem repeats	**6.72**	**14.94**	**4.19**	**7.32**	**7.89**	**14.75**
Satellite DNAs	6.14	14.37	3.91	7.00	7.63	14.32
rDNAs	0.54	0.57	0.28	0.32	0.26	0.43
telomeric repeats	0.04	-	-	-	-	-
Organellar DNA						
plastids	1.48	1.15	1.54	1.07	0.42	0.27
mitochondria	-	-	0.21	-	-	-
Other repeats						
unclassified	2.70	1.35	3.21	1.71	2.70	1.35

Table 3: Tandem repeats (satDNAs) observed in the species investigated.

Species	Satellite	Length of monomer [bp]	Genomic proportion [%]
D. olen	Sat-ole1	20	0.19
	Sat-ole2	65	0.07
	Sat-ole3	150	0.03
	Sat-ole4	165	0.93
	Sat-ole5	185	0.03
	Sat-ole6	190	3.01
	Sat-ole7	200	0.42
	Sat-ole8	210	1.09
	Sat-ole9	260	0.02
D. vera	Sat-ver1	127	8.58
	Sat-ver2	180	0.84
	Sat-ver3	190	4.96
D. vieillardii	Sat-vie1	30	0.14
	Sat-vie2	50	0.27
	Sat-vie3	60	1.78
	Sat-vie4	65	0.22
	Sat-vie5	120-130	0.26
	Sat-vie6	150	0.11
	Sat-vie7	180	1.11
	Sat-vie8	200	0.02
D. minimifolia	Sat-min1	120	0.22
	Sat-min2 (=Sat-vie7)	180	0.63
	Sat-min3	260	0.69
	Sat-min4	unclassified	5.45
D. parviflora	Sat-par1	47	0.36
	Sat-par2	50	0.09
	Sat-par3	67	0.35
	Sat-par4	115	4.97
	Sat-par5	180	0.73
	Sat-par6	260	1.13
D. pancheri	Sat-pan1	30	0.05
	Sat-pan2	50	0.08
	Sat-pan3	60	9.79
	Sat-pan4	120	0.38
	Sat-pan5 (=Sat-vie7)	180	0.72
	Sat-pan6	260	1.00

Plastid genome

We obtained between 79,119 (50.2 x coverage, *D. pancheri*) and 183,092 (116.2 x coverage, *D. vieillardii*) pairs of reads per individual that mapped to the plastid genome. The GC-content of the plastid genome varied between the endemic New Caledonian species (33 - 37%) and other *Diospyros* species (32%) and is slightly lower than in many other angiosperms (average: ~37%; e.g. *Camellia*, Yang *et al*. 2013; *Ardisia*, Ku *et al*. 2013; *Potenilla*, Ferrarini *et al*. 2013; *Musa*, Martin *et al*. 2013).

The size (~ 157 kb) and composition of the plastid genome of *Diospyros* is similar to that of *Camellia sinensis* (KC143082.1). A fully annotated plastome of *D. vieillardii* is given in Fig. 4. The plastid genome given here is the first fully sequenced plastid genome of Ebenaceae reported in the literature.

The plastid data set of *Diospyros* includes 159,166 characters of which 1178 variable characters are parsimony-uninformative, and 165 (0.1%) variable characters are parsimony-informative. Although these findings are based on only few species, the plastid genomes of recently radiated species like *Diospyros* seem to be not variable enough to use them for inference of phylogenetic relationships among closely related species. To conclude more about the usability of whole plastid genome sequences for inference of relationships, more species and more individuals per species need to be included into this data set. The sole tree resulting from the parsimony analysis is shown in Figure 3A.

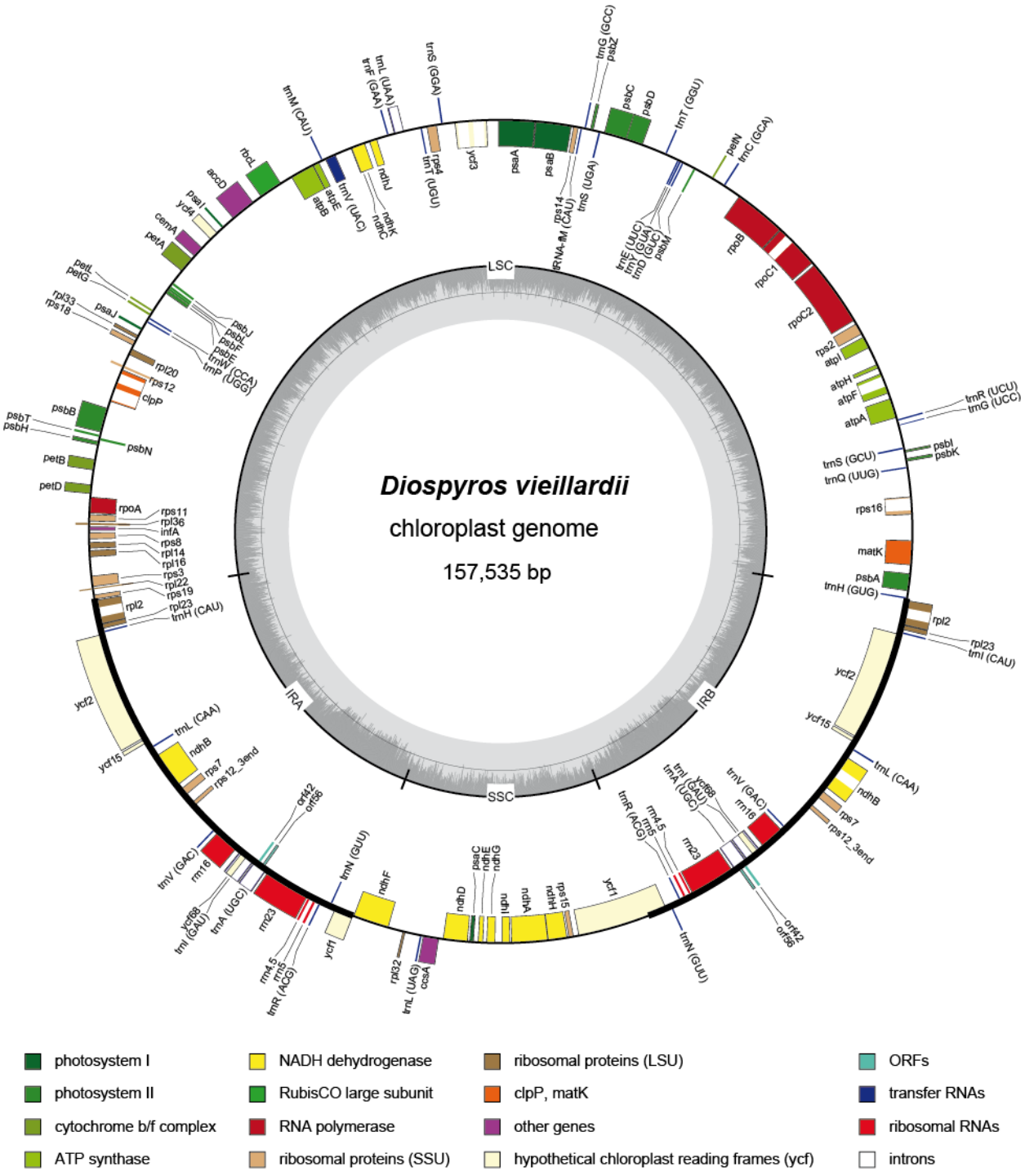
ONGOING WORK

We have prepared a second set of libraries of 24 individuals of 18 endemic New Caledonian *Diospyros* species. These data will be added to the initial data set to have a data set of 21 endemic New Caledonian species, which can be compared to results of our previous work. From these data we will infer the dynamics and role of repetitive DNA in New Caledonian *Diospyros* species, as well as whole plastid genomes for phylogenetic analyses.

ACKNOWLEDGEMENTS

This work was funded by a grant from the Austrian Science Fund (FWF, Project-Number: P 22159-B16) to R. Samuel. Thanks to S Dodsworth for his help with RepeatExplorer. Specimens are deposited in the herbaria of Noumea (NOU), University of Montpellier II (MPU) and the University of Vienna (WU).

Figure 4: Annotated plastome of *Diospyros vieillardii*.



REFERENCES

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. Journal of Molecular Biology 215: 403-410.

Barfuss MHJ, Samuel R, Till W, Stuessy TF. 2005. Phylogenetic relationships in subfamily Tillandsioideae (Bromeliaceae) based on DNA sequence data from seven plastid regions. American Journal of Botany 92: 337-351.

Barrett CF, Davis JI, Leebens-Mack J, Conran JG, Stevenson DW. 2013. Plastid genomes and deep relationships among the commelinid monocot angiosperms. Cladistics 29: 65-87.

Barrett CF, Specht CD, Leebens-Mack J, Stevenson DW, Zomlefer WB, David JI. 2014. Resolving ancient radiations: can complete plastid gene sets elucidate deep relationships among the tropical gingers (Zingiberales)? Annals of Botany 113: 119-133.

Bennett MD, Leitch IJ. 2005. Genome size evolution in plants, In: Gregory TR. (Eds.), The evolution of the genome, Elsevier, San Diego, pp. 90-151.

Casacuberta E, González J. 2013. The impact of transposable elements in environmental adaptation. Molecular Ecology 22: 1503-1517.

Čížková J, Hřibová E, Humplíková L, Christelová P, Suchánková P, Doležel J. 2013. Molecular analysis and genomic organisation of major DNA satellites in banana (*Musa* spp.). PLoS ONE 8: e54808.

Doolittle WF, Sapienza C. 1980. Selfish genes, the phenotype paradigm and genome evolution. Nature 284: 601-603.

Duangjai S, Samuel R, Munzinger J, Forest F, Wallnöfer B, Barfuss MHJ, Fischer G, Chase MW. 2009. A multi-locus plastid phylogenetic analysis of the pantropical genus *Diospyros* (Ebenaceae), with an emphasis on the radiation and biogeographic origins of the New Caledonian endemic species. Molecular Phylogenetics and Evolution 52: 602-620.

Ferrarini M, Moretto M, Ward JA, Šurbanovski N, Stevanović V, Giongo L, Viola R, Cavalieri D, Velasco R, Cetaro A, Sargent DJ. 2013. Evaluation of the PacBio RS platform for sequencing and *de novo* assembly of a chloroplast genome. BMC Genomics 14: 670.

Glenn TC. 2011. Field guide to next-generation DNA sequencing. Molecular Ecology Resources 11: 759-769.

Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. Nucleic Acids Symposium Series 41: 95-98.

Hiesel R, von Haeseler A, Brennicke A. 1994. Plant mitochondrial nucleic acid sequences as a tool for phylogenetic analysis. PNAS 91: 634-638.

Kejnovsky E, Hawkins JS, Feschotte C. 2012. Plant transposable elements: biology and evolution. In: Wendel JF, Greilhuber J, Doležel J, Leitch IJ. (Eds.), Plant genome diversity Volume 1, Springer-Verlag, Wien, pp. 17-34.

Kelly LJ, Leitch AR, Fay MF, Renny-Byfield S, Pellicer J, Macas J, Leitch IJ. 2012. Why size really matters when sequencing plant genomes. Plant Ecology and Diversity 5: 415-425.

Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH. 2005. Use of DNA barcodes to identify flowering plants. PNAS 102: 8396-8374.

Ku C, Hu, J-M, Kuo C-H. 2013. Complete plastid genome sequence of the basal Asterid *Ardisia polysticta* Miq. and comparative analyses of Asterid plastid genomes. PLoS ONE 8: e62548.

Kubis S, Schmidt T, Heslop-Harrison JS. 1998. Repetitive DNA elements as a major component of plant genomes. Annals of Botany 82 (Supplement A): 45-55.

Leitch IJ, Leitch AR. 2013. Genome size diversity and evolution in land plants. In: Leich IJ, Greilhuber J, Doležel J, Wendel JF. (Eds.), Plant genome diversity Volume 2, Springer-Verlag, Wien, pp. 307-322.

Leitch IJ. 2007. Genome size through the ages. Heredity 99: 121-122.

Lohse M, Drechsel O, Bock R. 2007. OrganellarGenomeDRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. Current Genetics 52: 267-274.

Lowry II PP. 1998. Diversity, endemism and extinction in the flora of New Caledonia: a review. In: Peng CF, Lowry II PP. (Eds.), Rare, threatened, and endangered floras of Asia and the Pacific rim, Institute of Botany, Taipei, Taiwan, pp. 181-206.

Macas J, Nemann P, Navrátilová A. 2007. Repetitive DNA in pea (*Pisum sativum* L.) genome: comprehensive characterization using 454 sequencing and comparison to soybean and *Medicago truncalata*. BMC Genomics 8: 427.

Martin G, Baurens F-C, Cardi C, Aury J-A, D´Hont A. 2013. The complete chloroplast genome of banana (*Musa acuminata*, Zingiberales): insight into plastid Monocotyledon evolution. PLoS ONE 8: e67350.

Maurizot P, Vendé-Leclerc M. 2009. New Caledonia geological map, scale 1/500000. Direction de l'Industrie, des Mines et de l'Energie - Service de la Géologie de Nouvelle-Calédonie, Bureau de Recherches Géologiques et Minières.

McCormack JE, Hird SM, Zellmer AJ, Carstens BC, Brumfield RT. 2013. Applications of next-generation sequencing to phylogeography and phylogenetics. Molecular Phylogenetics and Evolution 66: 526-538.

Natali L, Cossu RM, Barghini E, Giordani T, Buti M, Mascagni F, Morgante M, Gill N, Kane NC, Riesenberg L, Cavallini A. 2013. The repetitive component of the sunflower genome as shown by different procedures for assembling next generation sequencing reads. BMC Genomics 14: 686.

Novák P, Neumann P, Pech J, Steinhaisl J, Macas J. 2013. RepeatExplorer: a Galaxy-based web server for genome wide characterization of eukaryotic repetitive elements from next-generation sequence reads. Bioinformatics 29: 792-793.

Oliver KR, Greene WK. 2009. Transposable elements: powerful facilitators of evolution. BioEssays 31: 703-714.

Oliver KR, McComb JA, Greene WK. 2013. Transposable elements: powerful contributors to Angiosperm evolution and diversity. Genome Biology and Evolution 5:1886-1901.

Orgel LE, Crick FH. 1980. Selfish DNA: the ultimate parasite. Nature 284: 604-607.

Parisod C, Alix K, Just J, Petit M, Sarilar V, Mhiri C, Ainouche M, Chalhoub B, Grandbastien M-A. 2009. Impact of transposable elements on the organization and function of allopolyploid genomes. New Phytologist 186: 37-45.

Parisod C, Holdegger R, Brochmann C. 2010. Evolutionary consequences of autopolyploidy. New Phytologist 186: 5-17.

Park J-M, Schneeweiss GM, Weiss-Schneeweiss H. 2007. Diversity and evolution of Ty1-*copia* and Ty3-*gypsy* retroelements in the non-photosynthetic flowering plants *Orobanche* and *Phelipanche* (Orobanchaceae). Gene 387: 75-86.

Paun O, Fay MF, Soltis DE, Chase MW. 2007. Genetic and epigenetic alterations after hybridization and genome doubling. Taxon 56: 649-656.

Pelletier B. 2006. Geology of the New Caledonia region and its implications for the study of the New Caledonian biodiversity. In: Payri C, Richer de Forges B. (Eds.), Compendium of marine species from New Caledonia, Documents Scientifiques et Techniques II4, Institut de Recherche pour le Développement Nouméa, pp. 17-30.

Petrov DA. 2001. Evolution of genome size: new approaches to an old problem. Trends in Genetics 17: 23-28.

Piednoël M, Aberer AJ, Schneeweiss GM, Macas J, Novák P, Gundlach H, Temsch EM, Renner SS. 2012. Next-generation sequencing reveals the impact of repetitive DNA across phylogenetically closely related genomes of Orobanchaceae. Molecular Biology and Evolution 29: 3601-3611.

Renny-Byfield S, Chster M, Kovařík A, Le comber SC, Grandbastien M-A, Deloger M, Nichols RA, Macas J, Novák P, Chase MW, Leitch AR. 2011. Next generation sequencing reveals genome downsizing in allotetraploid *Nicotiana tabacum*, predominantly through the elimination of paternally derived repetitive DNAs. Molecular Biology and Evolution 28: 2843-2854.

Renny-Byfield S, Kovařík A, Chster M, Nichols RA, Macas J, Novák P, Leitch AR. 2012. Independent, rapid and targeted loss of highly repetitive DNA in natural and synthetic allotetraploids of *Nicotiana tabacum*. PLoS ONE 7: e36963.

Renny-Byfield S, Kovařík A, Kelly LJ, Macas J, Novák P, Chase MW, Nichols RA, Pancholi MR, Grandbastien M-A, Leitch AR. 2013. Diploidization and genome size change in allopolyploids is associated with differential dynamics of low- and high-copy sequences. The Plant Journal 74: 829-839.

Russell A, Samuel R, Rupp B, Barfuss MHJ, Šafran M, Besendorfer V, Chase MW. 2010. Phylogenetics and cytology of a pantropical orchid genus *Polystachya* (Polystachyinae, Vandeae, Orchidaceae): Evidence from plastid DNA sequence data. Taxon 59: 389-404.

Sonnhammer ELL, Durbin R. 1996. A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. Gene 167: 1-10.

Steflova P, Tokan V, Vogel I, Lexa M, Macas J, Novák P, Hobza R, Vyskot B., Kejnovsky E. 2013. Contrasting patterns of transposable element and satellite distribution on sex chromosomes (XY_1Y_2) in the dioecious plant *Rumex acetosa*. Genome Biology and Evolution 5: 769-782.

Sveinsson S, Gill N, Kane NC, Cronk Q. 2013. Transposon fingerprinting using low coverage whole genome shotgun sequencing in Cacao (*Theobroma cacao* L.) and related species. BMC Genomics 14: 502.

Swofford DL. 2003. PAUP*. Phylogenetic analysis using parsimony (*and other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Tamura M, Tao R, Yonemori K, Utsunomiya N, Sugiura A. 1998. Ploidy level and genome size of several *Diospyros* species. Journal of the Japanese Horticultural Society 67: 306-312.

Tel-Zur N, Abbo S, Myslabodski D, Mizrahi Y. 1999. Modified CTAB procedure for DNA isolation from epiphytic cacti of genera *Hylocereus* and *Selenicereus* (Cactaceae). Plant Molecular Biology Reporter 17: 249-254.

Turner B, Munzinger J, Duangjai S, Temsch EM, Stockenhuber R, Barfuss MHJ, Chase MW, Samuel R. 2013a. Molecular phylogenetic of New Caledonian *Diospyros* (Ebenaceae) using plastid and nuclear markers. Molecular Phylogenetics and Evolution 69: 740-763.

Turner B, Paun O, Munzinger J, Duangjai S, Chase MW, Samuel R. 2013b. Amplified fragment length polymorphism (AFLP) data suggest rapid radiation of *Diospyros* species (Ebenaceae) endemic to New Caledonia. BMC Evolutionary Biology 13: 269.

Volff J-N. 2006. Turning junk into gold: domestication of transposable elements and the creation of new genes in eukaryotes. BioEssays 28: 913-922.

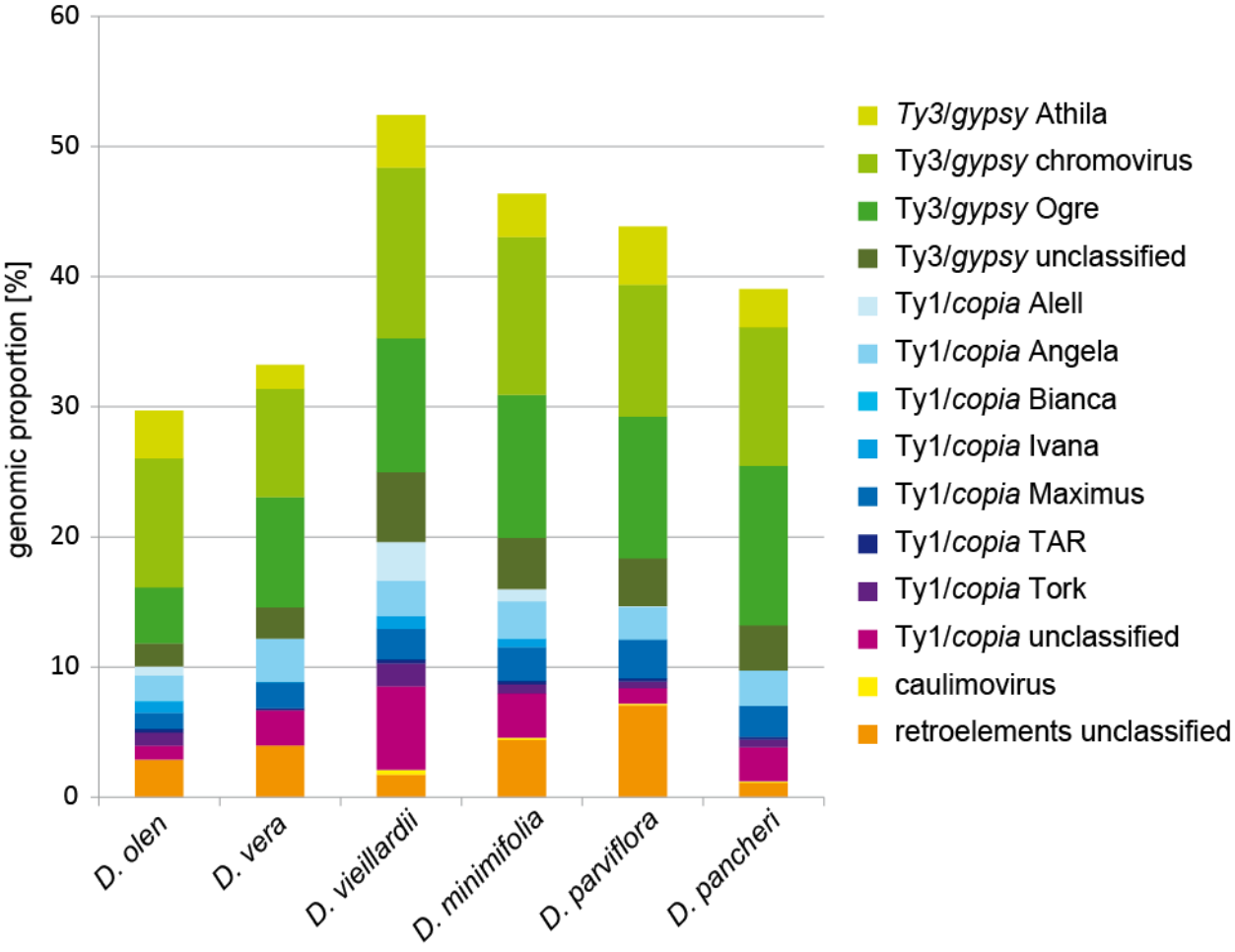
White F. 1992. Twenty-two new and little known species of *Diospyros* (Ebenaceae) from New Caledonia with comments on section *Maba*. Bulletin du Muséum national d'Histoire naturelle 4ème série – section B, Adansonia 2: 179-222.

Wicker T, Sabot F, Hua-Van A, Bennetzen JF, Capy P, Chalhaub B, Flavell A, Leroy P, Morgante M, Panaud O, Paux E, SanMiguel P, Schulmann AH. 2007. A unified classification system for eukaryotic transposable elements. Nature Reviews Genetics 8: 973-982.

Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organellar genomes with DOGMA. Bioinformatics 20: 3252-3255.

Yang J-B, Yang S-X, Li H-T, Yang J, Li De-Zhu. 2013. Comparative chloroplast genomes of *Camellia* species. PLoS ONE 8: e73053.

Supplementary Figure 1: Genomic proportions of LTR retrotransposons.



CONCLUSIONS

The species rich group of *Diospyros* in New Caledonia qualifies as book case example for an explosive adaptive radiation on an oceanic island. The questions aimed in this project (see Aims on p. 19) were generally answered.

The phylogenetic relationships among the investigated New Caledonian *Diospyros* species could be resolved, though not all received high statistical support. *Diospyros vieillardii* is sister to rest of the endemic species. The next branch separating from the remaining endemic group is formed by *D. flavocapra* and *D. umbrosa*, this being followed by a clade composed of *D. cherrieri* and *D. veillonii*. Species from calcareous substrates (*D. calciphila*, *D. inexplorata* and *D.* sp. Pic N'ga) formed a group in the sequencing analyses (both DNA and RAD). In both AFLP and RAD analyses *D. minimifolia* formed a clade together with the species of calcareous substrates. In neither of the analyses the individuals of *D. minimifolia* and *D. parviflora* formed unique clades, they also show great variability in leaf morphology thus indicating that they are not true species but could be of polytopic, perhaps hybrid origin. Apart from these two species most of the other species seem to be good species forming unique clades in either AFLP or in RAD analysis. In AFLP *D. vieillardii* was split into two groups, but they formed a well-supported clade in DNA sequencing and RAD analyses. In the RAD data *D. erudita*, *D. perplexa*, *D. pustulata*, and *D. revolutissima* failed to form unique clusters. No clear correlation was observed between phylogenetic grouping of *Diospyros* species and ecological conditions or geography. However in the case of *D. minimifolia* and *D. parviflora* (where populations of the same species do not cluster together) we do observe some geographical pattern of grouping.

According to the dating analyses, ancestors of the present *Diospyros* species reached New Caledonia around 9 mya via a long distance dispersal, most probably from islands in the Pacific Ocean (Indo-Malayan – Hawaiian archipelagos). Lineages forming a group of closely related species, among which relationships could not be clarified unambiguously, started to diversify around 4 mya. *Diospyros* being a woody plant with a generation time of several years we can conclude that most likely not more than a half million generations have existed since that time.

The genomes of the New Caledonian *Diospyros* species seem to be plastic/porous meaning that only a few genes are responsible for the species identity and that these genes are flexible enough to allow fast adaptation to new ecological conditions. Genes involved in binding and transporting compounds to/through the cell membrane were found to show species specific variants. Considering that these investigations were conducted with pairs of sister species occurring on different substrates (which are different in their heavy-metal content, as well as

nutrient availability) we can conclude, that adaptive radiation has played a role in shaping this group of New Caledonian *Diospyros* species.

Due to the low resolution of the phylogenetic trees based on nuclear and plastid markers it is not possible to predict anything about hybrid speciation in this group. However, analyses of the genetic structure of the New Caledonian *Diospyros* species using the AFLP and RAD data showed several admixed individuals which could be of hybridogenic origin.

Diospyros species from this clade of closely related New Caledonian species have larger genome than species from other clades or other regions of the world. Genome size differences observed are not due to polyploidy; in contrast to many other plant groups, polyploidy seems to be rare in the genus *Diospyros*. The endemic species, generally having higher genome sizes, have also more repeated elements than the other species with smaller genomes. We did not find group/species specific repeated elements. In *Diospyros* the most frequently observed repeated elements are LTRs, especially the Ty3/*gypsy* elements are the most abundant. This correlates well with observations in other plant groups. A second prominent group of repetitive DNA is formed by satellite repeats.

The size and genetic composition of the plastid genome of *Diospyros* is similar to the plastid genome of *Camellia sinensis*. However, the plastid genomes of endemic *Diospyros* (34%) and *Camellia* (37%) differ in GC content.

APPENDIX

Abstracts of conference contributions (oral presentations and posters)

Plant Genome Evolution

Amsterdam The Netherlands, 8-10 September 2013.

Evolution of New Caledonian *Diospyros* species (Ebenaceae)

Barbara Turner[1], Ovidiu Paun[1], Jérôme Munzinger[2], Sutee Duangjai[3], Mark W. Chase[4], Rosabelle Samuel[1]

[1]Department of Systematic and Evolutionary Botany, Faculty of Life Sciences, University Vienna, Rennweg 14, 1030 Wien, Austria; [2]IRD, UMR AMAP, TA A51/PS2, 34398 Montpellier Cedex 5, France; [3]Department of Forest Biology, Faculty of Forestry, Kasetsart University, Bangkok, Thailand; [4]Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3DS, UK.

Poster presentation: P045

In New Caledonia there are 31 species of *Diospyros* found and all but one (*D. fasciculosa*) are endemic. Molecular phylogenetic analyses of plastid and nuclear markers show that the New Caledonian *Diospyros* species are occurring in four different clades of which three contain only one to five species. The fourth group comprises 24 closely related species for which relationships remain mostly unresolved. Although species of this endemic group are morphologically distinct and largely occupy different niches, they exhibit little or no sequence divergence. The broadly distributed *D. vieillardii*, which is adapted to ultramafic soil but has the ability to grow on other soil types as well, has been shown to be sister to the rest of the group. We used Amplified Fragment Length Polymorphism (AFLP), a genome-wide molecular marker, to investigate species boundaries and their relationships. Distance-based and Bayesian analyses of AFLP data resulted in comparable results and suggest a process of rapid radiation. The analysed individuals are circumscribed into two subgroups, but they often do not follow morphological species boundaries. Bayesian analysis using STRUCTURE suggests a degree of admixture between the two gene pools for most species. Further, restriction-site associated DNA sequencing (RAD-seq), a next generation sequencing based technique that samples at reduced complexity across the investigated genomes, also supports rapid radiation and frequent interspecific hybridization. In the light of our AFLP and RAD-seq results the evolution of this group started after a long distance dispersal of an ancestor similar to present-day *D. vieillardii.* This gave rise to a couple of lineages, which later rapidly radiated across the available habitats of the island, but retained the propensity to frequently hybridize. Our results show the importance of rapid radiation across heterogenic habitats for successful colonization of islands.

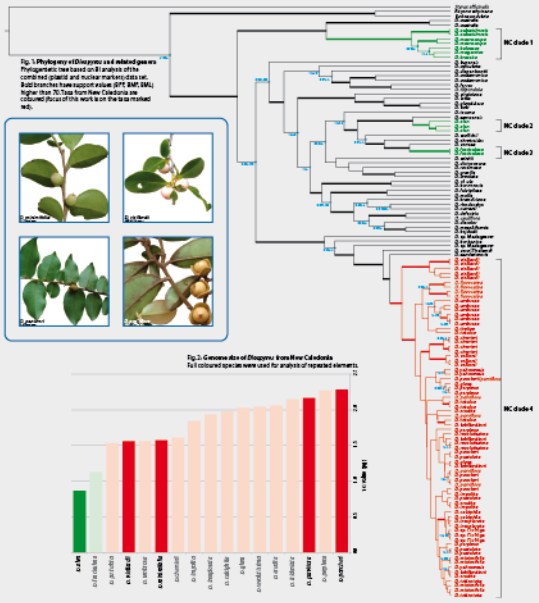
Genome evolution after rapid adaptive radiation in New Caledonian *Diospyros* species (Ebenaceae)

Barbara Turner[1], Ovidiu Paun[1], Jérôme Munzinger[2], Sutee Duangjai[3], Mark. W. Chase[4] & Rosabelle Samuel[1]

[1] Dept. of Systematic and Evolutionary Botany, Faculty of Life Sciences, University of Vienna, Rennweg 14, 1030 Vienna, Austria
[2] IRD, UMR AMAP, TA A51/PS2, 34398 Montpellier Cedex 5, France
[3] Dept. of Forest Biology, Faculty of Forestry, Kasetsart University, Bangkok, Thailand
[4] Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3DS, UK
Contact e-mails: barbara.turner@univie.ac.at, ovidiu.paun@univie.ac.at, mary.rosabella.samuel@univie.ac.at

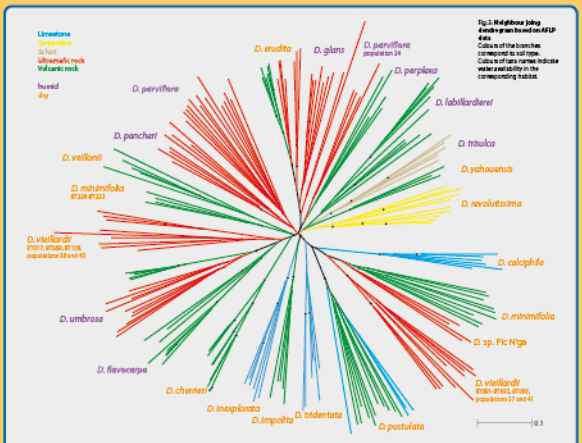
Diospyros L. is the largest genus of the family Ebenaceae, comprising over 500 species and it is widely distributed in the tropics and subtropics. In New Caledonia there are 31 species of *Diospyros*, and all but one of them, *D. fasciculosa*, are endemic (White, 1993).

Molecular phylogenetic analyses of plastid and nuclear markers show that the New Caledonian *Diospyros* species group in four different clades of which three contain only one to five species. The fourth group comprises 24 closely related species for which relationships remain mostly unresolved (Fig. 1). Although species of this endemic group are morphologically distinct and largely occupy different niches, they exhibit little or no sequence divergence. The broadly distributed *D. vieillardii*, which is adapted to ultramafic soil, has been shown to be sister to the rest of the group.



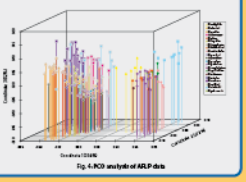
Genome sizes vary (up to 2.6 fold) between New Caledonian *Diospyros* species (Fig. 2). Species from the species-rich (forth) clade generally have larger genomes than those from other clades. Chromosome counts reveal the investigated species to be diploid.

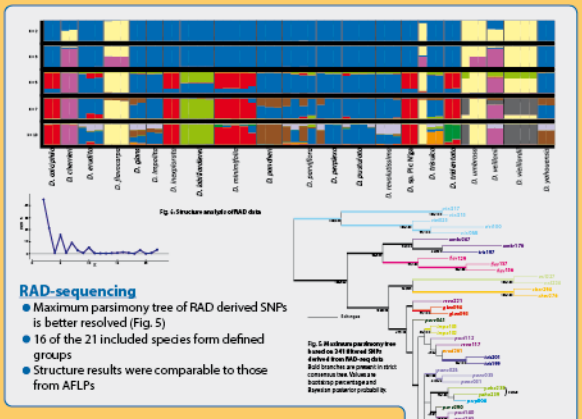
Dating analysis show the ancestor of clade four having arrived in New Caledonia via long distance dispersal around nine million years ago (mya). *Diospyros vieillardii* which is sister to the rest of the clade diverged around 7 mya (Turner *et al.*, 2013). Taking into account that these are woody plants with a relative long generation time (compared to herbaceous plants) we can consider them to be young by means of generations since they colonized New Caledonia.



AFLP

- most of the morphologically-defined taxa form clear groups (Fig. 3)
- backbone of the phylogenetic tree could not be resolved
- Bayesian Structure and Dice-distance based PCO analysis (Fig. 4) grouped the taxa into two groups, each with three subgroups.
- grouping in neither of the analyses corresponds to ecological or geographical attributes.



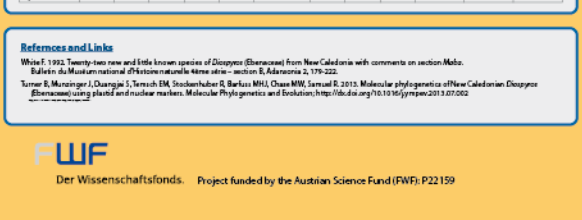


RAD-sequencing

- Maximum parsimony tree of RAD derived SNPs is better resolved (Fig. 5)
- 16 of the 21 included species form defined groups
- Structure results were comparable to those from AFLPs

Repeated Elements

According to the differences in genome size we expected to find different amounts of repeated elements in the genomes of the investigated species. We sequenced 14 individuals of *Diospyros* with different genome sizes (Fig. 2) at a low coverage. Species with lagers genomes (e.g. *D. parviflora* and *D. panchei*) generally more repeated elements, but no specific types have been found. Like common in many plant groups, the most abundant group of repeated elements found were LTR Ty3-gypsy retroelements.



Conclusions

The results obtained so far show *Diospyros* in New Caledonia to be a classical example of recent rapid radiation. Because of their relatively young age, sequence divergence between the species is low. Accelerated rates of evolution at few genes as a result of positive selection could have resulted in the morphological and ecological diversification apparent today in this group. Barriers to gene flow between these species may still be porous, with only few genes responsible for ecological and morphological adaptations evolving on distinct trajectories under selection, whereas the rest of the genomes seem open to gene flow. Finding these few genes is difficult because the information is "diluted" relative to the rest of their genomes.

References and Links

White F. 1993. Twenty-two new and little known species of *Diospyros* (Ebenaceae) from New Caledonia with comments on section *Maba*. Bulletin du Museum national d'Histoire naturelle 4ème série – section B, Adansonia 2, 179-222.

Turner B, Munzinger J, Duangjai S, Tansch EM, Stockenhuber R, Barfuss MHJ, Chase MW, Samuel R. 2013. Molecular phylogenetics of New Caledonian *Diospyros* (Ebenaceae) using plastid and nuclear markers. Molecular Phylogenetics and Evolution; http://dx.doi.org/10.1016/j.ympev.2013.07.002

FWF Der Wissenschaftsfonds. Project funded by the Austrian Science Fund (FWF): P22159

Poster presented at "Plant Genome Evolution 2013"

15. Treffen der Österreichischen Botanikerinnen und Botaniker

Austrian botanical meeting, Innsbruck Austria, 27-29 September 2012

Diversification of endemic New Caledonian *Diospyros* (Ebenaceae)

Barbara Turner[1], Jérôme Munzinger[2], Sutee Duangjai[3], Michael H. J. Barfuss[1], Bruno Wallnöfer[4], Ovidiu Paun[1], Mark. W. Chase[5] & Rosabelle Samuel[1]

[1] Dept. of Systematic and Evolutionary Botany, Faculty of Life Sciences, University of Vienna, Rennweg 14, 1030 Vienna, Austria, [2] IRD, UMR AMAP, TA A51/PS2, 34398 Montpellier Cedex 5, France, [3] Dept. of Forest Biology, Faculty of Forestry, Kasetsart University, Bangkok, Thailand, [4] Natural History Museum, Burgring 7, 1010 Vienna, Austria, [5] Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3DS, UK.

Oral presentation

Diospyros is one of the largest genera of angiosperms, comprising approximately 500 species of which 31 are distributed in New Caledonia. Of these species all except *D. fasciculosa* are endemic to New Caledonia. Molecular studies based on plastid markers highlighted the presence of four lineages and two types of endemism in the genus. The first being paleoendemism with an Australian origin e.g., *D. macrocarpa*, *D. brassica*, *D. balansae*, and the second neoendemism with a recent Indomalesian origin that includes species like *D. vieillardii*, *D. umbrosa*, *D. parviflora* etc. Phylogenetic analysis based on low-copy nuclear genes such as chloroplast-expressed glutamine synthetase (*ncpGS*) and phytochrome A (*PhyA*) shows a similar pattern as that of plastid regions. Sequence divergence among neoendemics of clade II is low in results of both plastid and low-copy nuclear markers. The position of *D. vieillardii* as sister to the rest of the neoendemic species of the clade II is confirmed, which is associated with the ability of this species to grow in ultramafic soil, a special substrate that emerged early during radiation of the New Caledonian *Diospyros*. AFLP analysis is used to evaluate species boundaries of the neoendemics, which again supports the isolated position of *D. vieillardii*. We also want to compare the AFLP results with those of next generation sequencing technique RAD (restriction site associated DNA). Further analysis of data is in progress.

ICPHB 2012

International Conference on Polyploidy, Hybridization and Biodiversity Průhnice Czech Republic, 7-10 May 2012.

Speciation of New Caledonian *Diospyros* (Ebenaceae)

Barbara Turner[1], Jérôme Munzinger[2+], Sutee Duangjai[3], Eva Temsch[1], Bruno Wallnöfer[4], Mark. W. Chase[5] & Rosabelle Samuel[1]

[1] Dept. of Systematic and Evolutionary Botany, Faculty of Life Sciences, University of Vienna, Rennweg 14, 1030 Vienna, Austria; [2] IRD, UMR AMAP, Laboratoire de Botanique, 98848 Nouméa Cedex, New Caledonia; [+] current address: IRD, UMR AMAP, TA A51/PS2, 34398 Montpellier Cedex 5, France; [3] Dept. of Forest Biology, Faculty of Forestry, Kasetsart University, Bangkok, Thailand; [4] Natural History Museum, Burgring 7, 1010 Vienna, Austria; [5] Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3DS, UK.

Poster presentation: P1-31

Diospyros is one of the largest genera of angiosperms, comprising approximately 500 species of which 31 are distributed in New Caledonia. Of these species, all except *D. fasciculosa* are endemic to New Caledonia. Recent molecular studies on family Ebenaceae where a subset of New Caledonian *Diospyros* was included highlighted the presence of four lineages and two types of endemism: (1) paleoendemics, which suggested ancient origin (included in a clade dated to the upper Eocene based on Australian fossils), e.g., *D. macrocarpa*, *D. brassica*, *D. balansae*, and (2) neoendemics, recent Indo-Malesian elements that include species like *D. fasciculosa* and *D. olen*, and a highly diverse clade comprising the remaining species such as *D. vieillardi*, *D. umbrosa*, *D. parviflora* etc. Species boundaries among most of these neoendemics seem to be unclear and are not well accepted by all authors. Clear delimitations of taxa are needed for conservation purposes.

The ongoing project on New Caledonian *Diospyros* uses AFLP analysis to determine species boundaries (i.e., taxonomic units) of the neoendemics. Molecular phylogenetics using rapidly evolving plastid and low-copy nuclear genes such as chloroplast expressed Glutamine Synthetase (*ncpGS*) and alcohol dehydrogenase (*Adh*) are used to detect hybridization and introgression that could have given rise to speciation as well as reproductive isolation that has evolved as a consequence of divergent selection on traits in different environments and thus ecological speciation. Investigations on the variation in genome size of the New Caledonian *Diospyros* are in progress.

Speciation of New Caledonian *Diospyros* (Ebenaceae)

Barbara Turner[1], Jérôme Munzinger[2+], Sutee Duangjai[3], Eva Temsch[1], Bruno Wallnöfer[4], Mark. W. Chase[5] & Rosabelle Samuel[1]

[1] Dept. of Systematic and Evolutionary Botany, Faculty of Life Sciences, University of Vienna, Rennweg 14, 1030 Vienna, Austria
[2] IRD, UMR AMAP, Laboratoire de Botanique, 98848 Nouméa Cedex, New Caledonia
[+] current address: IRD, UMR AMAP, TA A51/PS2, 34398 Montpellier Cedex 5, France
[3] Dept. of Forest Biology, Faculty of Forestry, Kasetsart University, Bangkok, Thailand
[4] Natural History Museum, Burgring 7, 1010 Vienna, Austria
[5] Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3DS, UK

universität wien

Introduction

Diospyros L. is the largest genus of the family Ebenaceae, having over 500 species and is widely distributed in the tropics and subtropics. In New Caledonia there are 31 species of *Diospyros*, and all but one of them, *D. fasciculosa*, are endemic (White, 1993). The origins of the endemic New Caledonia *Diospyros* are unclear, but two possible scenarios have been proposed. The first is that they may have recently dispersed from the Indomalaysian region, an area with high species diversity. The second, supported by fossil evidence, is that these endemics are the result of diversification of ancient Gondwanan lineages. Although there are only a few species distributed in Australia, fossil evidence supports the hypothesis that *Diospyros* has been a component of the Australian flora since at least the Eocene (Christophel and Basinger, 1982).

Aims

- Clarification of geographical origin and relationships among the New Caledonian *Diospyros* species.
- Reconstruction of the evolution of the neoendemic *Diospyros* species of clade IV.
- Investigation of genome size variation among the neoendemic *Diospyros* species of clade IV.

Methods

To answer the questions above we have sequenced 4 chloroplast markers (*rbcL*, *atpB*, *trnK-matK* and *trnS-trnG* spacer) and 2 low copy nuclear genes (*ncpGS* and *PhyA*). In addition to that we measured genome size using flow-cytometry.


Fruit of *D. macrocarpa*


Male flower of *D. olen*


D. vieillardii with fruits


D. labillenieri with fruits


D. revolutissima with fruits


D. pustulata with fruits

Results

Chloroplast markers

Results of our phylogenetic analyses support multiple origins for the New Caledonian species of *Diospyros*; they are found in four lineages (Fig. 1).
I) The first clade is sister to the Australian species, *D. australis* and *D. pentamera*. Evidence from fossil record indicates that *D. australis* or species morphologically similar to it have been on the Australian landmass since at least the Eocene. This allows us to infer that the New Caledonian species in this clade are most likely paleoendemics.
II + III) The second and third lineages consist of endemic *D. olen* and non-endemic *D. fasciculosa*. They are embedded within a clade consisting of species from India, Sri Lanka and Southeast Asia. *D. olen* is a neoendemic that originated from relatively recently introduced Indomalaysian elements.
IV) The fourth lineage comprises the rest of the New Caledonian *Diospyros* species. They cluster together with Hawaiian species, *D. sandwicensis* , and the widely distributed species, *D. vera*. These species are considered neoendemics. Within this lineage, *D. vieillardii* is sister to the rest, which is associated with ultramafic soil, indicating that the ability to grow on such special substrates emerged early during the radiation of the New Caledonian Diospyros species.

Fig. 1: Phylogeny of *Diospyros* and related genera. Phylogenetic tree based on ML analysis of the combined chloroplast data set. Bootstrap percentages (BP) higher than 70 are given. Geographic distributions are indicated by color. Taxa in red are species from New Caledonia.

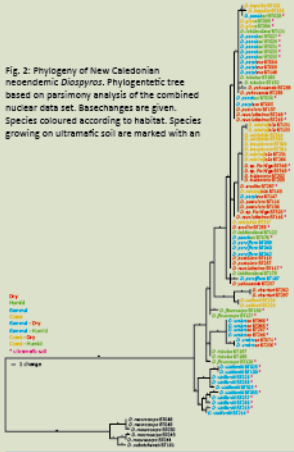


Fig. 2: Phylogeny of New Caledonian neoendemic *Diospyros*. Phylogenetic tree based on parsimony analysis of the combined nuclear data set. Basechanges are given. Species coloured according to habitat. Species growing on ultramafic soil are marked with an

Nuclear markers

To clarify the relationships within the neoendemic clade (IV) we analyzed two nuclear low-copy markers. As outgroup we used the paleoendemic clade (I) (Fig. 2). Although these markers are considered as variable markers, the resolution within the neoendemic clade is very low. The general topology of the phylogenetic tree reflects that of chloroplast markers.


Map of New Caledonia
Dots are indicating sampling locations

Genome size

The genome size of the neoendemic species is larger than it is in other *Diospyros* species. As the neoendemic *Diospyros* species are diploid the increased genome size could be due to transposon proliferation during the colonisation of the ultramafic soils. In addition to general genome size measurements we also compared different preservation techniques and their impact on flow cytometry results (Fig. 3).

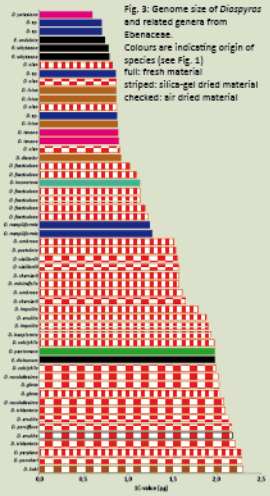


Fig. 3: Genome size of *Diospyros* and related genera from Ebenaceae.
Colours are indicating origin of species (see Fig. 1)
full: fresh material
striped: silica-gel dried material
checked: air dried material

Further research

As the nuclear low-copy markers did not give the desired resolution (variation between individuals is in some cases greater than variation between species) we are currently trying to resolve the relationships among the New Caledonian neoendemic *Diospyros* species using classical AFLP and next generation sequencing technique RAD (restriction site associated DNA). Chromosome counts will be continued for the neo-endemics.

FWF
Der Wissenschaftsfonds.
The research was funded by the Austrian Science Fund (FWF): P22159
Contact: barbara.turner@univie.ac.at

Poster presented at “ICPHB 2012”

IBC 2011

XVIII International botanical congress, Melbourne Australia, 23-30 July 2011.

Pattern and mode of speciation of New Caledonian *Diospyros* (Ebenaceae)

Samuel, R[1], Turner, B[1], Duangjai, S[2], Munzinger, J[3], Wallnoefer, B[4], Chase, M[5]

[1]Dept of Systematic and Evolutionary Botany, University of Vienna, Austria; [2]Dept of Forest Biology, Kasetsart University, Bangkok, Thailand; [3]Laboratoire de Botanique, Centre IRD de Nouméa, New Caledonia; 4Natural History Museum, Vienna, Austria; [5]Jodrell Laboratory, Royal Botanic Gardens, Kew, UK

Oral presentation: Symposium 168 – Theme 06

Diospyros is one of the largest genera of angiosperms, comprising approximately 500 species of which 31 are distributed in New Caledonia. Of these species all except *D. fasciculosa* are endemic to New Caledonia. Recent molecular studies on family Ebenaceae where a subset of New Caledonian *Diospyros* was included highlighted the presence of four lineages and two types of endemism. The first being paleoendemics, which suggested ancient Gondwana (Australian) origin, e.g., *D. macrocarpa, D. brassica, D. balansae*, and the second neoendemics, recent Indo–Malesian elements that include species like *D. vieillardi, D. umbrosa, D. parviflora* etc. Species boundaries among most of the neoendemics seem to be unclear and are not well accepted by all authors. The ongoing project on New Caledonian *Diospyros* uses AFLP analysis to determine species boundaries (i.e. taxonomic units) of the neoendemics. Molecular phylogenetics using rapidly evolving plastid and low-copy nuclear sequences will detect hybridization and introgression that could have given rise to speciation as well as reproductive isolation that has evolved as a consequence of divergent selection on traits in different environments and thus ecological speciation. Speciation is often accompanied by chromosomal rearrangement both numerical and structural, which will be investigated for the New Caledonian *Diospyros*.

IBC 2011

XVIII International botanical congress, Melbourne Australia, 23-30 July 2011.

Origin and evolution of New Caledonian *Diospyros* (Ebenaceae): a phylogenetic approach

Rosabelle Samuel[2], Barbara Turner[2], Sutee Duangjai[1], Jerome Munzinger[3], Bruno Wallnöfer[4], Michael H. J. Barfuss[2], Mark. W. Chase[5],

[1]Dept. of Forest Biology, Faculty of Forestry, Kasetsart University, Bangkok, Thailand; [2]Dept. of Systematic and Evolutionary Botany, Faculty of Life Sciences, University of Vienna, Rennweg 14, 1030 Vienna, Austria; [3]Laboratoire de Botanique, Centre IRD de Nouméa, BP A5, Noumea Cedex, New Caledonia; [4]Natural History Museum, Burgring 7, 1010 Vienna, Austria; [5]Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3DS, UK.

Poster presentation: Theme 06

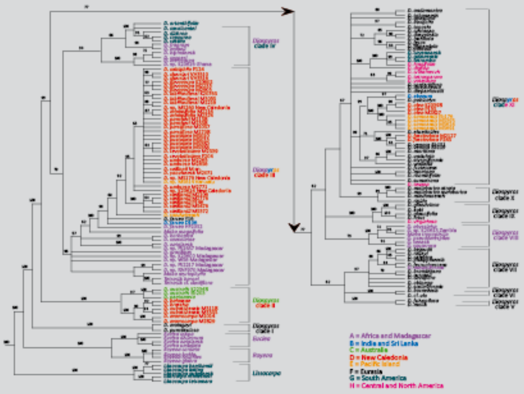
Diospyros is one of the largest genera comprising approximately 500 species of which 31 are distributed in New Caledonia. The recent molecular studies on family Ebenaceae where a subset of New Caledonian *Diospyros* was included highlighted the presence of four lineages and two types of endemism. The first being paleoendemics, which suggested ancient Gondwanan (Australian) origin, e.g., *D. macrocarpa*, *D. brassica* and *D. balansae*; and the second being neoendemics, elements coming relatively recently from the Indo-Malaysian region that include species like *D. vieilardi*, *D. umbrosa* and *D. parviflora*. This neoendemic species of New Caledonia group together with Hawaiian *D. sandwicensis* and widely distributed *D. ferrea*. The level of DNA sequence divergence among the neoendemic species is relatively low and does not appear to be correlated with the level of phenotypic diversity. The steep environmental gradients and unusual soil types in New Caledonia appear to have facilitated speciation in this group of neoendemics. Species boundaries among them are unclear and under discussion by different authors. In terms of conservation priorities for New Caledonian species, our results support the existence of four genetically distinct groups on this island. Each lineage of New Caledonian *Diospyros* should be treated as a separate conservation unit.

Origin and evolution of New Caledonian *Diospyros* (Ebenaceae): a phylogenetic approach

Sutee Duangjai[1], Barbara Turner[2], Jérôme Munzinger[3], Bruno Wallnöfer[4], Michael H. J. Barfuss[2], Mark. W. Chase[5], Rosabelle Samuel[2]

[1]Dept. of Forest Biology, Faculty of Forestry, Kasetsart University, Bangkok, Thailand
[2]Dept. of Systematic and Evolutionary Botany, Faculty of Life Sciences, University of Vienna, Rennweg 14, 1030 Vienna, Austria
[3]Laboratoire de Botanique, Centre IRD de Nouméa, BP A5, Noumea Cedex, New Caledonia
[4]Natural History Museum, Burgring 7, 1010 Vienna, Austria
[5]Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3DS, UK

universität wien



Introduction and Aims

Diospyros L. is the largest genus of the family Ebenaceae, having over 500 species and is widely distributed in the tropics and subtropics. In New Caledonia there are 31 species of *Diospyros*, and all but one of them, D. fasciculosa, are endemic (White, 1993). The origins of the endemic New Caledonia *Diospyros* are unclear, but two possible scenarios have been proposed. The first is that they may have recently dispersed from the Indomalaysian region, an area with high species diversity. The second, supported by fossil evidence, is that these endemics are the result of diversification of ancient Gondwanan lineages. Although there are only a few species distributed in Australia, fossil evidence supports the hypothesis that *Diospyros* has been a component of the Australian flora since at least the Eocene (Christophel and Basinger, 1982).

Here we aim (1) to clarify phylogenetic relationships within the pantropical genus *Diospyros* (Ebenaceae sensu lato), (2) estimate relationships among the New Caledonian *Diospyros* species, and (3) determine geographical of New Caledonian endemics.

To assess potential intra-specific variation versus inter-specific variation, multiple accessions of some New Caledonian taxa were included (*D. cherrieri, D. fasciculosa, D. flavocarpa, D. labillardierei, D. macrocarpa, D. minimifolia, D. olen, D. oubatchensis, D. pancheri, D. parviflora, D. pustulata*, D. *revolutissima, D. umbrosa* and *D. vieillardii*).

Outgroup sampling included members of the other genera of Ebenaceae, i.e. *Euclea, Lissocarpa* and *Royena*.

The phylogenetic analysis was based on multiple regions DNA sequences of cp genome (*rbcL, atpB, matK, ndhF, trnK* intron, *trnL* intron, *trnL-trnF* spacer and *trnS-trnG* spacer).

Fig. 1. Phylogeny of *Diospyros* and related genera. Strict consensus of 15000 equally most-parsimonious trees resulting from MP analysis of the combined data set. Bootstrap percentages (BP) higher than 50 are given above branches. Geographic distributions are indicated by color. Taxa in red are species from New Caledonia.

Results and conclusions

Monophyly of each genera of Ebenaceae is strongly supported (BP 100).

Dispersal/vicariance (DIVA) analysis of the combined data showed eight distribution areas according to their biogeographic background: Africa and Madagascar (A), India and Sri Lanka (B), Australia (C), New Caledonia (D), Hawaii and Pacific Islands (E), Eurasia (F), South America (G), and Central and North America (H).

Results of our phylogenetic analyses support multiple origins for the New Caledonian species of *Diospyros*; they are found in four lineages in three of the major clades (II, III, and XI) (Figs. 1 – 3).

The first clade consisting of *D. balansae, D. brassica, D. macrocarpa*, and *D. oubatchensis*, is sister to the Australian species, *D. australis* and *D. pentamera*, in clade II. Evidence from the fossil record indicates that *D. australis* or species morphologically similar to it have been on the Australian landmass since at least the Eocene. This allows us to infer that the New Caledonian species in this clade are most likely paleoendemics.

The second New Caledonian lineage comprises species belonging to section *Maba* (White, 1992). All fifteen species of this group included in our study cluster together in clade III with Hawaiian species, D. sandwicensis , and the widely distributed species, *D. ferrea*. A recent radiation of species took place following its introduction into New Caledonia (Fig. 4). These species are considered neoendemics. The steep environmental gradients and unusual soil types in New Caledonia appear to have facilitated speciation in this group (White, 1992).

Within this lineage, *D. vieillardii* is sister to the rest, which is of particular note as it is a plant associated with ultramafic soil (White, 1993), indicating that the ability to grow on such special substrates emerged early during the radiation of the New Caledonian *Diospyros* species in section Maba.

The third and fourth lineages consist of endemic *D. olen* and non-endemic *D. fasciculosa*. The New Caledonian species *D. fasciculosa* and *D. olen*, are embedded within a clade consisting of species from India, Sri Lanka and Southeast Asia, i.e., *D. ebenum, D. ehretioides, D. maritima, D. pubicalyx, D. styraciformis, D. venosa* and *D. wallichii*, and D. *frutescens* as the sister to the rest of the clade. This clade is further embedded within Southeast Asian species. Therefore *D. olen* is a neoendemic that originated from relatively recently introduced Indomalaysian elements.

In terms of conservation priorities for New Caledonian species, our results support the existence of four genetically distinct groups of *Diospyros* on these islands, two of which are closely related. Because biodiversity should be measured not only based on number of species but also using accumulated evolutionary history (Williams et al., 1991; Mooers, 2007; Forest et al., 2007), each lineage of New Caledonian *Diospyros* should be treated as a separate conservation unit.

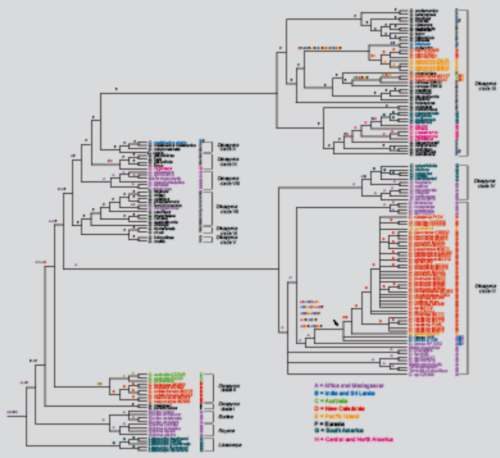


Fig. 2. Biogeographical optimization based on dispersal-vicariance analyses performed with the program DIVA (Ronquist, 1996) using one of the 15000 equally-most parsimonious trees resulting from MP phylogenetic analysis of Ebenaceae. Taxa in red are species from New Caledonia. The proposed biogeographical scenarios of New Caledonian species are shown to the right of the tree

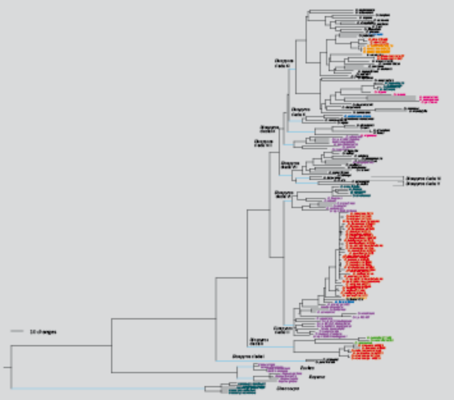
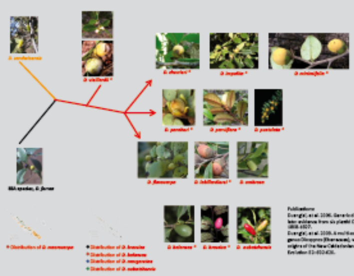


Fig. 3. One of the 15000 equally parsimonious trees resulting from the MP analysis of the combined data set represented as a phylogram with branch lengths proportional to the number of character substitutions. Some of the New Caledonian taxa are shown in the right of the tree. (* from http://www.endemia.nc/)



Der Wissenschaftsfonds.

Poster presented at “IBC 2011”

CURRICULUM VITAE

Personal Details

Name	Turner
First name	Barbara Christa
Academic degree	Mag. rer. nat.

Education

September 2010 – February 2014	PhD student in Biology, University Vienna
07.08.2008	Diploma-exam, with distinction
October 2003 – July 2008	Studies of biology/botany, University Vienna
12.06.2002	High-school graduation (Matura), with distinction
September 1997 – June 2002	Federal College of Horticulture (HBLVA für Gartenbau, Schönbrunn, Grünbergstraße 24), Vienna
September 1993 – June 1997	Secondary school (Bundesrealgymnasium Rainergasse 39), Vienna
September 1989 – June 1993	Primary school (Übungsvolkschule Ettenreichgasse 45b), Vienna

Work experience at University

September 2010 – April 2014	PhD student Evolution and biodiversity of New Caledonian *Diospyros*, project leader: ao. Univ.-Prof. Dr. Mary Rosabelle Samuel (FWF no. P22159)
March – May 2009	Scientific employee Molecular Phylogeny and evolution of genus *Polystachya*, project leader: ao. Univ.-Prof. Dr. Mary Rosabelle Samuel (FWF no. P1908)
November – December 2008	Scientific employee Evolutionary radiation in *Hypochaeris*, project leader: Univ.-Prof. Dr. Tod F. Stuessy (FWF no. P18446)
April 2007 – July 2008	Diploma student Molecular Phylogeny and evolution of genus *Polystachya*, project leader: ao. Univ.-Prof. Dr. Mary Rosabelle Samuel (FWF no. P1908)

Work experience (non-academic)

September 2009 – August 2010 August 2002 – September 2003 July – August 2001	Lab technician, HBLA u. BA für Wein und Obstbau, Klosterneuburg
April – August 2000 July 1999	Internships for horticultural education

Teaching experience

Since 2010	Methods in evolution and systematics of plants (course leader: ao. Univ.-Prof. Dr. Karin Vetschera)
2009	Practical course on genetics and molecular biology (course leader: ao. Univ.-Prof. Dr. Josef Loidl)
Since 2008	Different courses on molecular phylogenetics (course leader: ao. Univ.-Prof. Dr. Mary Rosabelle Samuel)

Conferences

September 2012	15. Treffen der Österreichischen Botanikerinnen und Botaniker, Innsbruck, Austria
May 2012	International Conference on Polyploidy, Hybridization and Biodiversity, Pruhonice, Czech Republic
November 2007	Orchid evolutionary biology and conservation: from Linnaeus to the 21st century, RBG Kew, London, UK

Scientific stays abroad

June 2013	One week visit to the lab of Prof. Andrew Leitch (Queen Mary University of London) to learn data analysis of repeated elements (financed from FWF project P22159)
February – April 2011	Field trip to New Caledonia (financed from FWF project P22159)
January – February 2009	Scientific stay in Jodrell laboratory, Royal Botanic Gardens Kew, London, UK (financed from Synthesys grant)

Scientific memberships

Since 2010	International Association for Plant Taxonomy, IAPT

Language skills

German	First language
English	Written and spoken (CEFR C1/C2)
French	Basic knowledge (CEFR A2)

Publications

Turner B, Paun O, Munzinger J, Duangjai S, Chase MW, Samuel R. 2013b. Analyses of amplifies fragment length polymorphisms (AFLP) indicate rapid radiation of *Diospyros* species (Ebenaceae) endemic to New Caledonia. BMC Evolutionary Biology 13: 269.

Turner B, Munzinger J, Duangjai S, Temsch EM, Stockenhuber R, Barfuss MHJ, Chase MW, Samuel R. 2013a. Molecular phylogenetics of New Caledonian *Diospyros* (Ebenaceae) using plastid and nuclear markers. Molecular Phylogenetics and Evolution 69: 740-763.

Rebernig CA, Weiss-Schneeweiss H, Blöch C, **Turner B**, Stuessy TF, Obermayer R, Villasenõr JL, Schneeweiss GM. 2012. The evolutionary history of the white-rayed species of *Melampodium* (Asteraceae) involved in multiple cycles of hybridization and polyploydization. American Journal of Botany 99: 1043-1057.

Weiss-Schneeweiss H, Blöch C, **Turner B**, Villasenõr JL, Stuessy TF, Schneeweiss GM. 2012. The promiscuous and the chaste: frequent allopolyploid speciation and its genomic consequences in American daisies (*Melampodium* sect. *Melampodium*; Asteraceae). Evolution 66: 211-228.

Rupp B, Samuel R, Russell A, Temsch EM, Chase MW, Leitch I. 2010. Genome size in *Polystachya* (Orchidaceae) and its relationships to epidermal characters. Botanical Journal of the Linnean Society 163: 223-233.

Russell A, Samuel R, Klejna V, Barfuss MHJ, **Rupp B**, Chase MW. 2010. Reticulate evolution in diploid and tetraploid species of *Polystachya* (Orchidaceae) as shown by plastid DNA sequences and low-copy nuclear genes. Annals of Botany 106: 37-56.

Russell A, Samuel R, **Rupp B**, Barfuss MHJ, Safran M, Besendorfer V, Chase MW. 2010. Phylogenetics and cytology of a pantropical orchid genus *Polystachya* (Polystachyinae, Vandeae, Orchidaceae): Evidence from Plastid DNA sequence data. Taxon 59: 389-404.

Rupp B. 2008. Genome size and molecular phylogeny of selected *Polystachya* species (Orchidaceae). Diplomarbeit, Universität Wien.

Conference contributions

Turner B, Paun O, Munzinger J, Duangjai S, Chase MW, Samuel R. 2013. Evolution of New Caledonian *Diospyros* species (Ebenaceae). Plant Genome Evolution, Amsterdam, The Netherlands, September 8-10, 2013. Poster

Turner B, Munzinger J, Duangjai, S, Barfuss MHJ, Wallnöfer B, Paun O, Chase MW, Samuel R. 2012. Diversification of endemic New Caledonian *Diospyros* (Ebenaceae). In: Berichte des naturwissenschaftlich-medizinischen Vereins in Innsbruck, Supplementum 20, 15. Treffen der Österreichischen Botanikerinnen und Botaniker, Innsbruck, Austria, September 27-29, 2012. Talk

Turner B, Munzinger J, Duangjai S, Temsch E, Wallnöfer B, Chase MW, Samuel R. 2012. Speciation of New Caledonian *Diospyros* (Ebenaceae). International Conference on Polyploidy, Hybridization and Biodiversity, Pruhonice, Czech Republic, May 7-10, 2012. Poster

Samuel R, **Turner B**, Duangjai S, Munzinger J, Wallnöfer B, Chase MW. 2011. Pattern and mode of speciation of New Caledonian *Diospyros* (Ebenaceae). IBC2011 XVIII international botanical congress, Melbourne, Australia, 23-30 July 2011. Talk

Samuel R, **Turner B**, Duangjai S, Munzinger J, Wallnöfer B, Chase MW, Barfuss MHJ. 2011. Origin and evolution of New Caledonian *Diospyros* (Ebenaceae): a phylogenetic approach. IBC2011 XVIII international botanical congress, Melbourne, Australia, 23-30 July 2011. Poster

Russell A, Samuel R, Barfuss MHJ, **Turner B**, Chase MW. 2011. Evolutionary inference from multiple incongruent DNA data matrices: reticulate evolution of *Polystachya* (Orchidaceae). IBC2011 XVIII international botanical congress, Melbourne, Australia, 23-30 July 2011. Poster

Rebernig CA, Weiss-Schneeweiss H, Blöch C, Schneeweiss GM, **Rupp B**, Stuessy TF. 2010. Recurrent polyploidization and hybridization in the evolution of *Melampodium* series *Leucantha* (Astreraceae). In: Verhandlungen des naturforschenden Vereins in Brünn, Specialausgabe, Structural and Functional Diversity of the Eukaryotic Genome, Augustiner-Abtei Alt Brünn, October 14-16, 2010. Poster

Russell A, Samuel R, Klejna V, Barfuss M, **Rupp B**, Chase M. 2010. Analysis of multiple nuclear and plastid loci reveals reticulate evolution in diploid and tetraploid species of genus *Polystachya* (Orchidaceae). In: International Conference: New Frontiers in Plant Systematics and Evolution (NFPSE 2010). Beijing, China, 7.-9.7.2010. Poster

Russell A, Samuel R, Barfuss MJH, **Rupp B**, Klejna V, Chase MW. 2009. Low copy nuclear genes reveal hybrid speciation in *Polystachya* (Orchidaceae). In: International Conference on Polyploidy, Hybridization and Biodiversity, Saint Malo, France. 17.-20.5.2009. Poster

Rebernig CA, Weiss-Schneeweiss H, Blöch C, Schneeweiss GM, **Rupp B**, Stuessy TF. 2009. Unravelling multiple cycles of hybridization and polyploidization in the evolutionary history of *Melampodium* series *Leucantha* (Asteraceae). International Conference on Polyploidy, Hybridization and Biodiversity, Saint Malo, France. 17.-20.5.2009. Poster

Weiss-Schneeweiss H, Blöch C, **Rupp B**, Stuessy TF. 2009. Origin and genome evolution of polyploidy species of the genus *Melampodium* (Asteraceae). International Conference on Polyploidy, Hybridization and Biodiversity, Saint Malo, France. 17.-20.5.2009. Poster

Russell A, Samuel R, **Rupp B**, Barfuss MJH, Safran M, Klejna V, Chase MW. 2009. Phylogeny, cytology and biogeography of the pantropical orchid genus *Polystachya*. In: Systematics. National Herbarium of the Netherlands and National Museum of Natural History Naturalis, Leiden, The Netherlands. 10.-14.8.2009. Talk

Russell A, Samuel MR, Rupp B, Barfuss MHJ, Chase MW. 2008. Phylogenetics, African biodiversity and intercontinental dispersal in *Polystachya* (Orchidaceae). In: X Young Systematists Forum. Flett Theatre, Natural History Museum, London, UK. 2. December 2008. Poster

Russell A, Samuel MR, Barfuss MHJ, **Rupp B**, Chase MW. 2008. Molecular systematics of *Polystachya* (Orchidaceae). In: Fourth International Conference: The Comparative Biology of the Monocotyledons. Copenhagen, Denmark, 11-15 August 2008. Abstracts. Talk

Russell A, **Rupp B**, Safran M, Samuel R, Barfuss MHJ, Klejna V, Schneeweiss H, Besendorfer V, Reich D, Chase MW. 2008. Molecular Systematics of *Polystachya* (Orchidaceae). In: Botany without Borders. University of British Columbia, Vancouver, BC, Canada, July 26-30 2008. Poster

Russell A, **Rupp B**, Safran M, Samuel R, Barfuss MHJ, Weiss-Schneeweiss H, Besendorfer V, Reich D, Chase MW. 2007. Molecular systematics of *Polystachya*. Orchid evolutionary biology and conservation: from Linnaeus to the 21st century, 31st October - 2nd November 2007. Royal Botanic Gardens Kew. Poster