The LINNEAN SOCIETY of London BOTANICAL Journal of the Linnean Society

Botanical Journal of the Linnean Society, 2016. With 3 figures

Plastid phylogenetics of Oceania yams (*Dioscorea* spp., Dioscoreaceae) reveals natural interspecific hybridization of the greater yam (*D. alata*)

HANA CHAÏR[1*†], JULIE SARDOS[2†], ANTHEA SUPPLY[1], PIERRE MOURNET[1], ROGER MALAPA[3] and VINCENT LEBOT[4]

[1]*CIRAD, UMR AGAP, F-34398 Montpellier, France*
[2]*Bioversity International, Parc Scientifique Agropolis II, 1990 Boulevard de la Lironde, 34397 Montpellier Cedex 5, France*
[3]*VARTC, PO Box 231, Santo, Vanuatu*
[4]*CIRAD, UMR AGAP, PO Box 946, Port Vila, Vanuatu*

Received 12 February 2015; revised 1 September 2015; accepted for publication 7 December 2015

Phylogenetic relationships of Oceanian staple yams (species of *Dioscorea* section *Enantiophyllum*) were investigated using plastid *trnL-F* and *rpl32-trnL*$^{(UAG)}$ sequences and nine nuclear co-dominant microsatellites. Analysis of herbarium specimens, used as taxonomic references, allowed the comparison with samples collected in the field. It appears that *D. alata*, *D. transversa* and *D. hastifolia* are closely related species. This study does not support a direct ancestry from *D. nummularia* to *D. alata* as previously hypothesized. The dichotomy in *D. nummularia* previously described by farmers in semi-perennial and annual types was reflected by molecular markers, but the genetic structure of *D. nummularia* appears more complex. *Dioscorea nummularia* displayed two haplotypes, each corresponding to a different genetic group. One, including a *D. nummularia* voucher from New Guinea, is closer to *D. tranversa*, *D. alata* and *D. hastifolia* and encompasses only semi-perennial types. The second group is composed of semi-perennial and annual yams. However, some of these annual yams also displayed *D. alata* haplotypes. Nuclear markers revealed that some annual yams shared alleles with *D. alata* and semi-perennial *D. nummularia*, suggesting a hybrid origin, which may explain their intermediate morphotypes and the difficulty met in classifying them.

ADDITIONAL KEYWORDS: *Dioscorea hastifolia* – *Dioscorea nummularia* – *Dioscorea transversa* – *rpl32-trnL*$^{(UAG)}$ – *trnL-F* – Vanuatu.

INTRODUCTION

Yams are members of the genus *Dioscorea* L. (Dioscoreaceae; Dioscoreales). *Dioscorea* is the largest and only dioecious genus in the family, comprising *c.* 640 species (Govaerts, Wilkin & Saunders, 2007) historically assembled into 32–59 sections (Knuth, 1924; Ayensu, 1972). The genus had a pantropical distribution long before the advent of humans, with most of the species being isolated by natural barriers into three continental groups: Asiatic, African and American (Hahn, 1995).The phylogenetic relationships between species of Dioscoreales remain unresolved, although several studies have attempted to clarify them (Caddick *et al.*, 2002; Wilkin *et al.*, 2005; Hsu *et al.*, 2013). There is, however, a paucity of knowledge on the systematic relationships between different species within sections. It is even more complex in areas where yams are considered as indigenous crops connected to local cultures and traditions. In such areas, yam diversity is managed by farmers through the use of wild, spontaneous and cultivated yams (Malapa *et al.*, 2005; Scarcelli *et al.*, 2006; Bousalem *et al.*, 2010; Chaïr *et al.*, 2010), leading to confusion in the systematic identification of specimens.

*Corresponding author. E-mail: hana.chair@cirad.fr
†These authors contributed equally to this work.

Dioscorea section *Enanthiophyllum* Uline is the most economically important section as it contains the main cultivated edible species, notably *D. cayenensis* Lam. and *D. rotundata* Poir. that originated in West Africa, *D. nummularia* Lam., a temperate yam, *D. opposita* Thunb. (probably a synonym of *D. japonica* Thunb.), *D. transversa* R.Br. from Southeast Asia and Oceania; and *D. alata* L. for which the origin remains unknown. Although many studies have attempted to clarify the relationships between African species (Chaïr *et al.*, 2005; Girma *et al.*, 2014), the relationships between Asian and Oceanian species, namely *D. alata*, *D. nummularia* and *D. transversa*, remain unclear.

Dioscorea alata, or greater yam, is believed to have originated from Southeast Asia (Burkill, 1960) and then to have been introduced to the South Pacific islands, where it has a high cultural value. It was dispersed from New Guinea by the first Lapita settlers who spread eastwards from the Bismarck Archipelago > 3000 years ago (Kirch, 2000; Bedford, 2006). It is now the most widely distributed cultivated yam species in the world and is probably also the oldest, with an ancient domestication history (Hahn, 1995; Lebot, 2009). *Dioscorea alata* is a morphologically distinct species, although unknown in the wild, and is not known to hybridize with other *Dioscorea* spp. (Lebot *et al.*, 1998). It was suggested that it could have been domesticated by human selection from wild forms of common origin with *D. hamiltonii* Hook.f., and the synonymous *D. persimilis* Prain & Burk., occurring in an area extending from northern India to Taiwan (Coursey, 1976; http://e-monocot.org). However, recent amplified fragment length polymorphism (AFLP) studies indicated that this species is not the direct ancestor of *D. alata*, and *D. alata* is close to *D. nummularia* and a cultivated form of yam found in Oceanian islands thought to be *D. transversa* (Malapa *et al.*, 2005). Proximity of *D. nummularia* and *D. alata* was confirmed with subsequent *rbcL* and *matK* sequencing (Wilkin *et al.*, 2005).

Dioscorea nummularia, or spiny yam, is native to Melanesia and to Island South-East Asia (ISEA) (http://e-monocot.org). An important centre of diversity is most probably New Guinea, but in the Solomon Islands and Vanuatu spontaneous and wild forms also occur in the forest in addition to several cultivars (Walter & Lebot, 2003). *Dioscorea nummularia* is known in Vanuatu, where this species has been the most documented (Malapa, 2005), as 'wael yam', which means wild yam in Bislama, a local Pidgin English. It is a spontaneous and semi-perennial yam subjected to unusual cultivation practices that are close to paracultivation (Dounias, 2001). Tubers are planted under the canopy and living trees are used as climbing supports for the vines of this semi-perennial plant. Left untouched for 3 to 4 years after plantation, they are then harvested once a year without seasonal constraints. This yam is an important food used in times of food scarcity in Vanuatu (Sardos, 2008; Lebot, 2009). In addition to the common semi-perennial cultivars, some rare annual cultivars, e.g. 'Lapenae', have also been reported (Malapa, 2005).

Additionally, another group of yams belonging to unidentified taxa (Malapa, 2005) and named 'strong yam' by farmers in Vanuatu, is also cultivated in Oceania, often in the same plots as *D. alata*. Strong yams are also annual types and are appreciated for their high dry matter content when compared with *D. alata*. Generally associated with *D. nummularia* [e.g. Kirch (1994) for Futuna or Thaman (1988) in Fiji], some of the strong yam cultivars grown in Vanuatu, but not all, have been recently associated with the Australian species *D. transversa*. Strong yams cultivars named 'Marou' (Malapa *et al.*, 2006) are believed to have been introduced into neighbouring New Caledonia at the beginning of the 20th century by blackbirded workers coming back from Queensland (Bourret, 1973) and to have further spread to Vanuatu.

Dioscorea transversa, or pencil yam, is an Australian species growing in eastern and northern parts of the country. It was commonly harvested, consumed and even stored by Australian Aboriginals (Clarke, 2007). *Dioscorea transversa* is not cultivated in continental Asia and, so far, it has been reported only in Melanesia and Australia. Its edible tubers have high dry matter content and good organoleptic quality, higher than *D. alata* and similar to *D. nummularia* (Lebot, 2009).

Despite the unique square stems with wings at each angle of *D. alata*, confusion over its morphology with *D. nummularia* and *D. transversa* has been reported in the Philippines (Cruz & Ramirez, 1999), Indonesia (Sastrapradja, 1982) and New Caledonia (Bourret, 1973). Consequently, and despite their major importance in local diets, the taxonomy of these three species in section *Enantiophyllum* and their phylogenetic relationships remain unclear: the strong yams cannot be strictly assigned to a particular species, the relationships between *D. nummularia* and *D. alata* are still not resolved and their phylogenetic relationships with *D. transversa* remain unclear.

In the present study, herbarium specimens were used as taxonomic references and two plastid non-coding regions, namely *trnL-F* (Taberlet *et al.*, 1991) and the *rpl32-trnL*$^{(UAG)}$ intergenic spacer (Shaw *et al.*, 2007), widely used for studying intra- and interspecific-level phylogenetic relationships were

combined. Consequently, the phylogenetic relationships of the three yam species commonly planted in Melanesia, namely *D. transversa*, *D. nummularia* and *D. alata*, and the strong yams were investigated. In addition and to explore all possible origins of the strong yams and resolve the difficulty met in classifying them, putative hybridization events between the different taxa were also investigated using a set of nuclear microsatellite markers. Lastly, the findings are discussed to address the impact of the traditional management system on yam phylogenetics in Oceania.

MATERIAL AND METHODS

Description of traditional yams in Oceania

Dioscorea alata, sop-sop yam, has a typical square-winged stem, opposite narrowly heart-shaped leaves and can produce bulbils in addition to its large and long underground tubers. *Diosocrea nummularia*, wael yam, is a robust, high-climbing, spiny vine with large, cordate and elliptical leaves being opposite at the lower portion of the stem and alternate at the upper. *Dioscorea nummularia* is semi-perennial and produces compact and shallow well-developed tubers (Lebot, 2009). *Dioscorea transversa* has lignified stems which develop no wings like *D. nummularia* or discrete ones like those of *D. alata*. Leaves are alternate basally on stems and opposite distally. They are similar in shape to *D. alata* but with thick and shiny laminas similar to *D. nummularia*. Tuber shape varies, the most common ones having thin cylindrical tubers growing deep into the soil (Lebot, 2009). Due to their spiny vines and to the morphological similarities of some cultivars with *D. nummularia* and despite the heterogeneity met in the architecture of their tubers, strong yams are often considered as annual forms of *D. nummularia* (Malapa, 2005), although some of them were confused with *D. transversa* (Malapa *et al.*, 2006). Hereafter, we refer to this taxon as *Dioscorea* sp.

Taxon sampling

To investigate the relationships between the *Dioscorea* species, we used dried plant material obtained from herbarium specimens and from field collections. Detailed information and locations of samples collected are presented in Table 1 and Figure 1A. Three species were included in this study, *D. nummularia* (Dn), *D. alata* (Da) and *D. transversa* (Dtv), with specimens of the strong yams (Dsp). We also included *D. hastifolia* Endl. (Dh), another species of section *Enantiophyllum* which belongs to the Australian genepool. It was widely used by Aboriginal societies (Hallam, 1975; Denham, 2008) and was apparently cultivated in large plots (Grey in Gammage, 2009). Today, it is not consumed and its distribution is restricted to the west coast of Australia.

Of the 28 specimens from field collections, provided by the Vanuatu Agricultural and Technical Center (VARTC), six belong to *D. alata*, seven belong to the local strong yam group and 15 specimens are identified as *D. nummularia* and classified as wael yam, including one specimen, DnLapenae, which is cultivated by farmers as an annual crop. All *D. alata* specimens were collected in Vanuatu: three originated from Vanuatu (Da1003, Da357 and DaMVu) and the three others from India (Da313, Da402 and DaFIn). Seven herbarium specimens were obtained from the National Herbarium of Australia (CANB): one *D. nummularia* specimen collected in Papua New Guinea (PNG) (Dn27777), four *D. transversa* specimens collected in East Australia (Dtv416 and Dtv12068) and in the Torres Strait Islands (Dtv6602 and Dtv3434) and two *D. hastifolia* specimens collected in Western Australia (Dh4322 and Dh15249). The leaves of these specimens were collected between 1935 and 2004. A specimen of *D. bulbifera* L. (Section *Opsophyton* Uline), provided by VARTC, was included in this study to be used as outgroup.

DNA extraction, amplification and sequencing

Total genomic DNA from field samples was extracted from specimens as described by Risterucci *et al.* (2000) and then purified using a column purification kit as described by the manufacturer (Macherey-Nagel; cat. no. 740 571 100). DNA extractions from herbarium samples were conducted using the Qiagen 96 Plant kit for lyophilized tissues (Qiagen). For each set of DNA extraction, at least two negative extraction controls were performed.

Two plastid DNA regions, *trnL-F* (Taberlet *et al.*, 1991) and *rpl32-trnL*$^{(UAG)}$ (Shaw *et al.*, 2007), which have proven their usefulness in investigating low taxonomic levels where species may be very closely related (Kelchner, 2000; Hodkinson *et al.*, 2002; Viruel, Catalàn & Segarra-Moragues, 2012), were amplified. However, the published primers failed to amplify DNA from herbarium specimens. Indeed, short fragments are abundant in herbarium DNA and can decrease the success of PCR amplification (Särkinen *et al.*, 2012). New primer pairs were thus designed to amplify both regions by small overlapping sequences of maximum 200–400 bp. To identify strictly homologous regions enabling primer design, sequences from *Dioscorea* spp. were downloaded from GenBank and aligned. For the *trnL-F* region, sequences of 12 species were used [*D. abyssinica* (D89715.1), *D. alata* (DQ841331.1), *D. bulbifera*

Table 1. Sample collection and herbarium voucher specimens, common and Bislama names, growth cycle, geographical origin, provider and GenBank accession of the *Dioscorea* spp. used in this study

Sample ID	Affiliated species	Common name	Bislama* name	Growth cycle	Geographical origin	Source (provider)	*trnL-F* accession no.	*rpl32-trnL*$^{(UAG)}$ accession no.	Haplotype
Da1003	*D. alata*	Greater yam	Sopsop yam	Annual	Vanuatu	VARTC	KM888689	–	(Ha1)
Da357	*D. alata*	Greater yam	Sopsop yam	Annual	Vanuatu	VARTC	KM888687	KM888719	Ha1
Da313	*D. alata*	Greater yam	Sopsop yam	Annual	India	VARTC	KM888686	KM888721	Ha2
Da402	*D. alata*	Greater yam	Sopsop yam	Annual	India	VARTC	KM888688	KM888720	Ha2
DaFIn	*D. alata*	Greater yam	Sopsop yam	Annual	India	VARTC	KM888690	KM888722	Ha2
DaMVu	*D. alata*	Greater yam	Sopsop yam	Annual	Vanuatu	VARTC	KM888685	KM888723	Ha2
Dh4322	*D. hastifolia*	Warran yam	NA	Annual	West Australia	CANB	KM888708	KM888726	Hh1
Dh15249	*D. hastifolia*	Warran yam	NA	Annual	West Australia	CANB	KM888707	KM888725	Hh2
Dn1	*D. nummularia*	Spiny yam	Wael yam	Semi-perennial	Vanuatu	VARTC	KM888696	KM888727	Hn2
Dn2	*D. nummularia*	Spiny yam	Wael yam	Semi-perennial	Vanuatu	VARTC	KM888697	KM888728	Hn2
Dn3	*D. nummularia*	Spiny yam	Wael yam	Semi-perennial	Vanuatu	VARTC	KM888698	KM888729	Hn1
Dn4	*D. nummularia*	Spiny yam	Wael yam	Semi-perennial	Vanuatu	VARTC	KM888699	KM888753	Hn1
Dn5	*D. nummularia*	Spiny yam	Wael yam	Semi-perennial	Vanuatu	VARTC	KM888700	KM888730	Hn2
Dn6	*D. nummularia*	Spiny yam	Wael yam	Semi-perennial	Vanuatu	VARTC	KM888701	KM888731	Hn2
Dn7	*D. nummularia*	Spiny yam	Wael yam	Semi-perennial	Vanuatu	VARTC	KM888702	KM888732	Hn2
Dn8	*D. nummularia*	Spiny yam	Wael yam	Semi-perennial	Vanuatu	VARTC	KM888703	KM888733	Hn1
Dn9	*D. nummularia*	Spiny yam	Wael yam	Semi-perennial	Vanuatu	VARTC	KM888704	KM888734	Hn1
Dn10	*D. nummularia*	Spiny yam	Wael yam	Semi-perennial	Vanuatu	VARTC	KM888705	KM888735	Hn1
Dn11	*D. nummularia*	Spiny yam	Wael yam	Semi-perennial	Vanuatu	VARTC	KM888706	KM888736	Hn2
Dn206	*D. nummularia*	Spiny yam	Wael yam	Semi-perennial	Vanuatu	VARTC	KM888692	KM888737	Hn2
Dn333	*D. nummularia*	Spiny yam	Wael yam	Semi-perennial	Vanuatu	VARTC	KM888693	KM888739	Hn1
Dn335	*D. nummularia*	Spiny yam	Wael yam	Semi-perennial	Vanuatu	VARTC	KM888694	KM888740	Hn2
DnLapenae	*D. nummularia*	Spiny yam	Wael yam	Annual	Vanuatu	VARTC	KM888691	KM888741	Hn1
Dn27777	*D. nummularia*	Spiny yam	Wael yam	Semi-perennial	Papua New-Guinea	CANB	-	KM888752	(Hn2)
Dtv416	*D. transversa*	Pencil yam	NA	Annual	East Australia	CANB	KM888716	KM888750	Htv1
Dtv6602	*D. transversa*	Pencil yam	NA	Annual	Torres Strait Islands	CANB	KM888715	KM888748	Htv2
Dtv12068	*D. transversa*	Pencil yam	NA	Annual	East Australia	CANB	KM888717	KM888749	Htv1
Dtv3434	*D. transversa*	Pencil yam	NA	Annual	Torres Strait Islands	CANB	-	KM888751	(Htv2)
Dsp336	*Dioscorea* sp.	NA	Strong yam	Annual	Vanuatu	VARTC	KM888714	KM888747	Hn1
Dsp331	*Dioscorea* sp.	NA	Strong yam	Annual	Vanuatu	VARTC	KM888695	KM888738	Ha1
Dsp1033	*Dioscorea* sp.	NA	Strong yam	Annual	Vanuatu	VARTC	KM888711	KM888742	Hn1
Dsp1652	*Dioscorea* sp.	NA	Strong yam	Annual	Vanuatu	VARTC	KM888712	KM888743	Hn1
DspKwala	*Dioscorea* sp.	NA	Strong yam	Annual	Vanuatu	VARTC	KM888709	KM888744	Hn1
DspMarou	*Dioscorea* sp.	NA	Strong yam	Annual	Vanuatu	VARTC	KM888710	KM888745	Hn1
DspRul	*Dioscorea* sp.	NA	Strong yam	Annual	Vanuatu	VARTC	KM888713	KM888746	Ha1
Db	*D. bulbifera*	NA	NA	Perennial	Vanuatu	VARTC	KM888718	KM888724	–

*Pidgin language of Vanuatu; (), specimens which amplified for either *trnL-F* or *rpl32-trnL*$^{(UAG)}$.

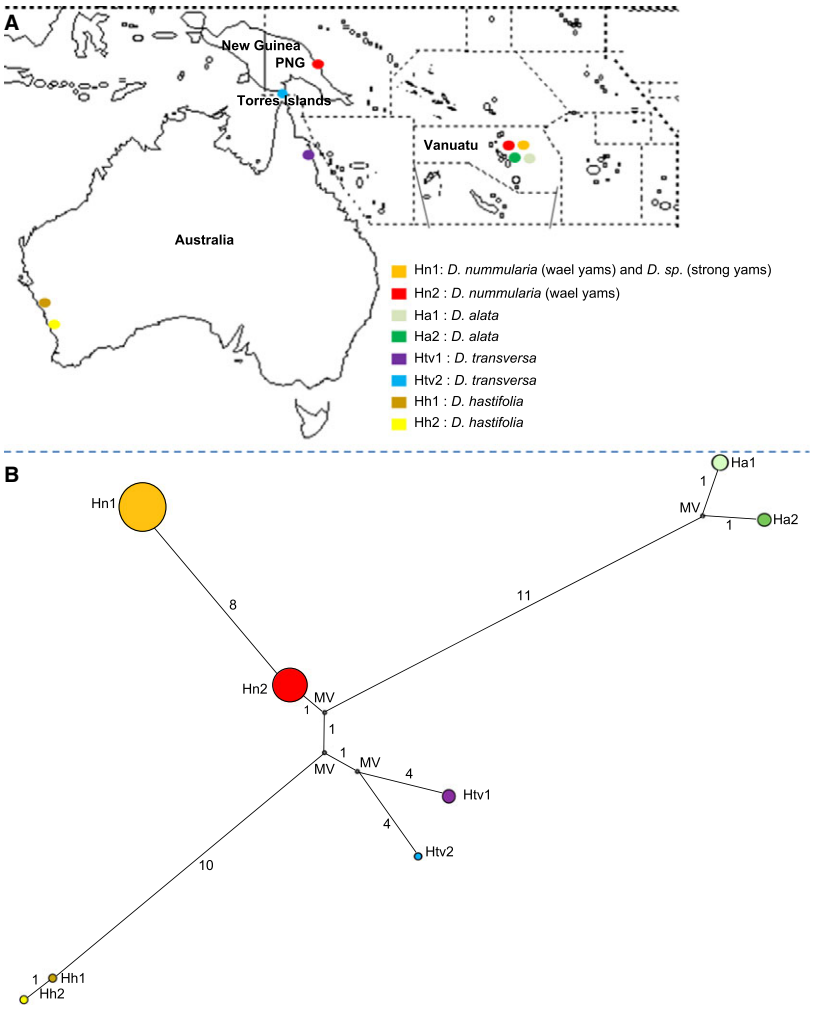

Figure 1. A, sampling localities and plastid haplotypes of the 32 specimens of yams analysed. B, MJ network obtained from the analyses of 32 *trnL-F* and *rpl32-trnL*$^{(UAG)}$ concatenated sequences, showing the relationships between *D. nummularia*, *Dioscorea* sp., *D. transversa*, *D. alata* and *D. hastifolia* haplotypes. MV, median vectors.

(EF619352.1), *D. cayenensis* (D89708.1), *D. rotundata* (D89695.1), *D. esculenta* (Lour.) Burkill (DQ841298.1), *D. esculenta* var. *spinosa* (GQ265290.1), *D. glabra* Roxb. (DQ841321.1), *D. hispida* Dennst. (DQ841323.1), *D. pentaphylla* L. (GQ265289.1), *D. persimilis* Prain & Burkill (DQ841328.1) and *D. praehensilis* (D89698.1)] and sequences of six species were used for *rpl32-trnL*$^{(UAG)}$ [*D. abyssinica* (JF705568.1), *D. bulbifera* (JF705571.1), *D. dumetorum* (Kunth) Pax (JF705570.1), *D. praehensilis* (JF705573.1), *D. rotundata* (JF705572.1) and *D. elephantipes* (L'Hér.) Engl. (EF380353.1)]. These sequences were aligned with BioEdit version 7.2.0 (Hall, 1999). Using the software Primer 3 plus (Untergasser *et al.*, 2007), primer pairs positioned in the conserved regions were then designed. All primer sequences used for PCR amplification and sequencing of the resulting fragments are presented in Table 2.

The PCR protocol was conducted using between 20 and 300 ng of DNA, 0.625 U Hotspot Taq polymerase (Promega), 5 μL 5× buffer (Promega), 50 mM $MgCl_2$, 0.4 μM each primer and 5 mM dNTPs. PCRs were performed in a PTC-100 thermocycler (MJ Research) with the following programme: 5 min at 94 °C, then ten cycles of 45 s at 94 °C, and touch-down of 1 min at T_a starting at 55 °C with −0.5 °C at each cycle, 2 min at 72 °C, then 26 cycles of 45 s at 94 °C, 1 min at 50 °C, 2 min at 72 °C and finally 5 min at 72 °C. Sequences from herbarium specimens were amplified with the newly designed primers and using the PCR mix provided with the GoTaq Long (Promega) with 2 μL 1:10 DNA and by adding 0.25 μL 100× bvine serum albumin. All PCR products were

Table 2. Characteristics of two probes of four clustering primer pairs developed for amplification and sequencing of the two plastid sequences: *trnL-F* and *rpl32-trnL*$^{(UAG)}$

PROBEDB_ACC	Locus		Primer sequence (5′–3′)	Allele size (bp)
Pr032251208	PDaCIRtrnL-trnF_frag1	F:	GGGATATGGCGAAATTGGTA	228
		R:	TTGAGTCTCTGCACCTATCCTTT	
	PDaCIRtrnL-trnF_frag2	F:	AAAGGATAGGTGCAGAGACTCAA	322
		R:	TTCTCGTCCGATTAATTCGTTT	
	PDaCIRtrnL-trnF_frag3	F:	TCAACCGAAGTTGAAGGAAGA	196
		R:	GGACTTGAACCCTCACGATT	
	PDaCIRtrnL-trnF_frag4	F:	AATCGTGAGGGTTCAAGTCC	263
		R:	GCGTGTCAGGAACCAGATTT	
Pr032251207	PDaCIRtrnL-rpl32_frag1	F:	GCTTCCTAAGAGCAGCGTGT	206
		R:	TGTTAAAACTGAACCCTATGACGA	
	PDaCIRtrnL-rpl32_frag2	F:	TGCTCAATCAATGATCTATCGTC	375
		R:	CAAAACCTAATTGATTTGAGAAATATG	
	PDaCIRtrnL-rpl32_frag3	F:	TATTTCTCAAATCAATTAGGTTTTGG	265
		R:	GTATGYGGAATACCAATTCTTTTGTC	
	PDaCIRtrnL-rpl32_frag4	F:	AGACAAAAGAATTGGTATTCCACA	293
		R:	GCGGTTCCAAAAAAACGTACTTC	

visualized by electrophoresis on 1.5% agarose gels using DNA ladder Exactladder DNA Premix 2 log (Ozyme). PCR products were purified using a PCR QIAquick kit (Promega). Direct sequencing was conducted on both strands on an ABI 3500 automated DNA sequencer (Applied Biosystems). All DNA sequences obtained in this study have been deposited in GenBank (accession numbers KM888685–KM888753).

Editing and sequence alignment

DNA sequences were edited and aligned manually against the *trnL-F* and *rpl32-trnL*$^{(UAG)}$ sequences of *D. elephantipes* (EF380353.1|:46494–47415 and EF38 0353.1|:122931–124105, respectively) using GenalysWin version 3.4.8 (CNG) (Takahashi *et al.*, 2003). Sequence statistics were analysed using MEGA version 4.0.2 (Tamura *et al.*, 2013).

Phylogenetic analyses

The best-fit partitioning scheme for our dataset was investigated using PartitionFinder (Lanfear *et al.*, 2012) using the Bayesian information criterion (BIC) as the information-theoric measure. GEVALT software (Davidovich, Kimmel & Shamir, 2007) as implemented in Haplophyle (http://haplophyle.cirad.fr) was used for haplotype definition analysing the concatenated *trnL-F* and *rpl32-trnL*$^{(UAG)}$ sequences. Then, a median-joining network analysis (MJ network) (Bandelt, Forster & Rohl, 1999) was performed with Haplophyle software. Considering that microstructural mutations and their underlying biological patterns are important features to be considered for phylogenetic analysis (e.g. Benson, 1997; Kelchner, 2000), they were coded as mutations for this analysis.

For phylogenetic analyses, maximum-likelihood (ML) and Bayesian Markov chain Monte Carlo (MCMC) methods were used. ML analysis was conducted using PhyML 3.0 (Guindon & Gascuel, 2003; Guindon *et al.*, 2010) and the Bayesian analysis (BA) was performed with MrBayes 3.1.2 (Huelsenbeck & Ronquist, 2001). Each method was applied to three data sets: the *trnL-F* sequences, the *rpl32-trnL*$^{(UAG)}$ sequences and the *trnL-F* and *rpl32-trnL*$^{(UAG)}$ concatenated sequences. Structural mutations were not considered in these analyses (Borsch & Quandt, 2009). For ML analysis, the best model of nucleotide evolution was estimated using the Akaike information criterion (AIC) implemented in JModeltest 2.1.4 (Posada, 2008). The HKY+G (Haegawa Kishimo and Yano) model with gamma-distributed rate variation across sites (*GtrnL-trnF* = 0.51; *Grpl32-trnL*$^{(UAG)}$ = 0.59 and *G* concatenated = 0.57) was selected for the three datasets as the best model among the 24 compared. The level of support for branches was tested using bootstrap support (BS) analysis with 500 replicates (Felsenstein, 1985).

For BA, the general time reversible model (GTR) with six types of substitution (6GTR) and a gamma-distributed rate variation across sites as identified by PartitionFinder as the best model for the concatenated sequences was chosen for sequence evolution. Each dataset was analysed by launching

simultaneously two runs of two MCMCs for 4000 000 generations each. Trees were sampled every 1000 generations. The first 500 000 trees were not considered and the remaining trees were used to construct consensus trees with Bayesian posterior probabilities (PP) compatible with the single tree. The consensus trees for each of the three datasets were viewed using FigTree version 1.4.0 (http://tree.bio.ed.ac.uk/software/figtree/).

Nuclear microsatellite analysis

Nine nuclear microsatellite markers (Andris *et al.*, 2010) were selected and used further for analysis: mDaCIR3, mDaCIR11, mDaCIR17, mDaCIR18_1, mDaCIR18_2, mDaCIR20, mDaCIR59, mDaCIR62_1 and mDaCIR62_2. PCR amplifications were performed on PTC-100 thermocyclers (MJ Research) and genotyping was carried out on an IR2-DNA analyser (LiCor 4200 Sequencer) as described by Chaïr *et al.* (2010). AFLP Quantar Pro 1.0 Software was used for automated data collection and to determine allele sizes.

Dioscorea alata and *D. nummularia* are polyploids (Malapa, 2005; Arnau *et al.*, 2009) and each allele was scored as 1 (present) or 0 (absent). A distance matrix was calculated using the Dice dissimilarity index between pairs of specimens (Dice, 1945). A minimum 80% proportion of valid data was required for each unit pair. The genetic relationships between specimens were assessed by constructing a weighted neighbour-joining (NJ) tree (Saitou & Nei, 1987) and by using principal coordinates analysis (PCoA) implemented in Darwin version 5 software (Perrier & Jacquemoud-Collet, 2006). To assess the degree of statistical support for the different branches in the NJ tree, we performed 500 replicates of bootstrap analysis on the data set.

RESULTS

Amplification and sequencing

For the *trnL-F* region, the newly designed primers succeeded in amplifying five of the herbarium specimens, but failed for Dn27777 and Dtv3434. Sequences were thus obtained for 33 samples. For *rpl32-trnL*$^{(UAG)}$, 35 sequences were obtained, but the one amplified from specimen Da1003 appeared truncated and was discarded from the subsequent analysis. Concatenation of the two sequences requires the same set of samples for all the loci and thus our analysis was limited to 32 samples for which we obtained complete sequences for both regions.

Base composition and alignment

Sequence characteristics for each of the two markers and the data set are shown in Table 3. The total aligned sequence length for *trnL-F* consisted of 751 bp, which included 20 structural mutations [a 4-bp inverted sequence and 19 insertions or deletions (indels)]. Twelve of the indels discriminated the targeted specimens from the outgroup. The total aligned sequence length for *rpl32-trnL*$^{(UAG)}$ consisted of 871 bp, including 24 structural mutations [a 25-bp inverted sequence and 23 indels]. Nineteen of these indels separated the targeted specimens from the outgroup. The total aligned sequence length of concatenated sequences consisted of 1621 bp. Once indels and inverted sequences were deleted, the *trnL-F* and *rpl32-trnL*$^{(UAG)}$ and concatenated sequence matrices consisted of 688, 738 and 1426 bp, respectively. The *trnL-F* region had 17 (2.47%) variable sites and six (0.87%) potentially parsimony-informative characters. The *rpl32-trnL*$^{(UAG)}$ region had 29 (3.93%) variable sites and 15 (2.03%) potentially parsimony-informative positions. The concatenated sequences had 46 (3.22%) variable sites and 21 (1.47%) potentially parsimony-informative characters.

PartitionFinder was then used to determine which partitioning scheme best fit the dataset for further analysis. The lowest BIC score was assigned to the concatenated sequence when compared with the analysis of the two individual sequences.

Table 3. DNA site variation for each marker and for the data set of *Dioscorea* spp. used in the phylogenetic analyses

Marker	*trnL-F*	*rpl32-trnL*$^{(UAG)}$	Concatenated sequences
Number of specimens	33	34	32
Total aligned length (bp)	751	871	1621
Aligned length analysed (bp) (without mutations)	688	738	1426
Conserved characters (bp)	671	709	1380
Variable characters (bp)	17 (2.47%)	29 (3.93%)	46 (3.22%)
Potentially informative characters (bp)	6 (0.87%)	15 (2.03%)	21 (1.47%)

Haplotype identification and phylogenetic network analysis

In our sample, eight haplotypes were identified using GEVALT software, corresponding to two haplotypes per species, and one more haplotype corresponding to the outgroup (Fig. 1B). The specimens from the strong yam group, classified in this study as *Dioscorea* sp., share haplotypes with either *D. alata* (Ha1) for two specimens or *D. nummularia* (Hn1) for the five other specimens, but none with *D. transversa* or *D. hastifolia* (Table 1). We noted for *D. nummularia* that Hn1 is composed of both *Dioscorea* sp. and *D. nummularia* specimens, including the annual *D. nummularia* 'Lapenae', whereas Hn2 is exclusively composed of *D. nummularia*. The two *D. transversa* haplotypes correspond to the specimens collected in East Australia (Ht1) and to the accession collected in the Torres Strait Islands (Ht2). Haplotype Ha2 encompasses three *D. alata* specimens from India out of four whereas Ha1 was composed of the remaining *D. alata* specimen from Vanuatu and the two *Dioscorea* sp. specimens.

The MJ network was analysed for the eight haplotypes obtained. Haplotypes from a given species appear linked in the MJ network, leading to the identification of four genetic groups corresponding to the four species studied at the exception of the affiliation of the *Dioscorea* sp. specimens to *D. nummularia* and *D. alata*. Both *D. alata* haplotypes and both *D. transversa* haplotypes appear to be derived by one and four mutations, respectively, from common ancestors, whereas *D. hastifolia* and *D. nummularia* share a common pattern with their haplotypes deriving by one and by eight mutations, respectively, from each other. According to the network topology, a common ancestor is shared by *D. alata*, *D. nummularia* and a wider genetic group composed of *D. hastifolia* and *D. transversa*, the two Australian species displaying a common ancestor between the one in common with *D. alata* and *D. nummularia* and their divergence. Considering haplotype Hn2, *D. nummularia* is the closest to this ancestor (one mutation) followed by *D. transversa* (six mutations), whereas *D. alata* and *D. hastifolia* are more distant (12 and 11 mutations, respectively).

Phylogenetic analysis using plastid sequences

BA and ML analyses were performed for the two plastid regions separately and for the combined dataset. The topologies of the trees obtained with BA and ML analyses were congruent and confirmed haplotyping. In addition, herbarium specimens that failed to produce one of the sequences, the specimens of *D. nummularia* collected in Papua New Guinea (Dn27777) and of *D. transversa* collected in the Torres Strait Islands (Dtv3434) for which no sequence was obtained for the *trnL-F* region, exhibited Hn2 and Htv2 haplotypes, respectively, in both trees obtained from the *rpl32-trnL*$^{(UAG)}$ sequence. Likewise, the accession of *D. alata* from Vanuatu Da1003 exhibited an Ha1 haplotype in the tree constructed from the *trnL-F* sequence. The posterior probability values obtained from the BA were weaker than those obtained from the ML analysis. We thus describe here the ML tree obtained from the concatenated sequence matrix.

The ML tree obtained after analysis of the concatenated *trnL-F* and *rpl32-trnL*$^{(UAG)}$ sequences is presented in Figure 2. It is composed of eight distinct branches that correspond to the haplotypes previously identified, in addition to the outgroup. *Dioscorea hastifolia* (BS = 100) and *D. alata* (BS = 97) are strongly supported, whereas *D. transversa* (BS = 53) is weakly supported. One of the haplotypes in *D. nummularia*, Hn1, is moderately supported (BS = 82), whereas divergence among Hn1 and Hn2 appears unresolved (BS = 56). Equally and according to the weak BS values for these branches (BS = 36 and BS = 46) divergence between *D. hastifolia*, *D. alata* and *D. transversa* is unresolved in the tree. Nevertheless and despite the overall low BS of the branches, the topology of the tree supports the results obtained with the network analysis. It suggests that Hn1 clusters apart from *D. alata*, *D. hastifolia* and *D. transversa*, which seem somehow to share a common ancestor with Hn2.

Phylogenetic analysis based on SSR data

Simple sequence repeat (SSR) analysis was conducted to identify specimens with hybrid status. SSR markers failed to amplify from herbarium specimens. Consequently, *D. transversa* and *D. hastifolia* could not be included in the analysis. In addition, three samples obtained from fresh leaves (DaFIn, Dsp1033 and DspKwala) had too many missing data to be kept for further analyses. Overall, 25 samples were thus considered for this part of the study. Eighty-two alleles were identified using nine microsatellite primer pairs (Table 4). Among these alleles, 49 were found to be species-specific, of 22 in *D. alata*, 15 in *D. nummularia* and 12 alleles in *Dioscorea* sp. specific. Among the 46 alleles obtained for the five *Dioscorea* sp. specimens analysed, eight alleles were shared with *D. alata* and 17 with *D. nummularia*. This number of shared alleles between *Dioscorea* sp. and the two other species is higher than the nine alleles found to be common to all specimens analysed. It is also higher than the three alleles common to *D. alata* and *D. nummularia*.

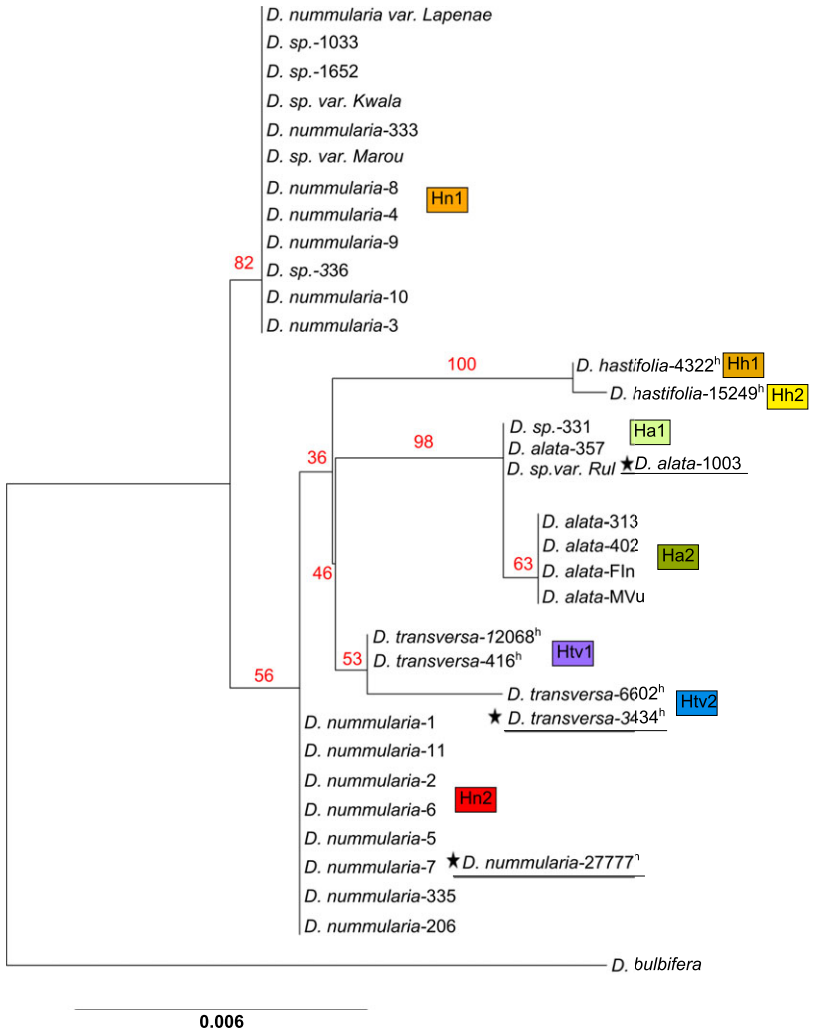

Figure 2. ML-based phylogenetic tree obtained from the analyses of 32 *trnL-F* and *rpl32-trnL*$^{(UAG)}$ concatenated sequences. ^{h}Herbarium specimens. Colour boxes: haplotypes. Specimens underlined and preceded with an asterisk correspond to sequences obtained with one of the two intergenic spacers with position on the ML tree with *trnL-F* or *rpl32-trnL*$^{(UAG)}$ sequence.

Pairwise dissimilarities were computed for the 25 specimens using the Dice distance and were then depicted by an NJ tree (Fig. 3A). Three main clusters were identified. Cluster A composed of specimens with haplotype Ha. It was subdivided into two subclusters: the first was composed of *D. alata* specimens and the second was composed of the two strong yam specimens with haplotype Ha1. The second cluster, cluster B, assembled exclusively wael yams with haplotype Hn1 except specimen Dn5 which presented haplotype Hn2. It was also subdivided into two subclusters with 100% bootstrap support. The third cluster, C, was also subdivided into two subclusters, the first one with the remaining wael yams with haplotype Hn2 and the second one with remaining strong-yams with haplotype Hn1. Bootstrap support values between the different clusters and subclusters were highly significant (93–100%).

To investigate the relationships between specimens further, a PCoA was carried out (Fig. 3B). The first two eigenvalues obtained explained 63.61% of the total variance. The differentiation between *D. alata* and *D. nummularia* specimens appeared clearly on each side of the first axis. The second axis further differentiated the two haplotypes of *D. nummularia*. We noted that four *D. nummularia* samples composed of individuals exhibiting either Hn1 or Hn2 haplotypes are positioned between two clusters exclusively composed of Hn1 and Hn2 haplotypes, respectively. The PCoA also showed that the strong yam specimens do not constitute a single cluster and tend to have an intermediate position between

Table 4. Number of alleles obtained using the nine microsatellite markers and their distribution among the studied species

Species	Farmers' classification	*N*	Total alleles	Total alleles per species	Specific alleles per species	Shared alleles *Dioscorea* sp./*D. alata*	Shared alleles *Dioscorea* sp./*D. nummularia*	Shared alleles three species	Shared alleles *D. nummularia*/*D. alata*
D. alata	Sopsop Yam	5	82	38	22	8	17	9	3
Dioscorea sp.	Strong Yam	5		51	12				
D. nummularia	Wael Yam	15		46	15				

D. alata and the Hn2 haplotype of *D. nummularia*, even though three samples, namely Dsp336, DspMarou and Dsp1652, appear closer to Hn2 than to *D. alata*.

DISCUSSION

Phylogenetic relationships between *D. alata*, *D. nummularia*, *D. transversa* and *Dioscorea* sp

There is no comprehensive study of the phylogenetics of yams native to Oceania. Although a worldwide phylogenetic study of *Dioscorea* has been carried out (Wilkin *et al.*, 2005), section *Enantiophyllum* and especially the yams cultivated in Oceania, were poorly and sparsely investigated (Caddick *et al.*, 2002; Wilkin *et al.*, 2005; Hsu *et al.*, 2013). In an attempt to understand the relationships between these different species, our study using plastid sequence analyses indicates that *D. alata*, *D. transversa*, *D. hastifolia* and *D. nummularia* are closely related species. Bootstrap supports obtained on the phylogenetic trees were weak. This is mainly due to the number of markers and to the sample used. Our study targeted four closely related species in order to understand their relationships and shed light on strong yams; such a pattern may explain the low level of variation and consequently the overall low BS obtained. However, the tree topologies are informative. The ML tree suggests the divergence of Hn1 and Hn2 and the emergence of *D. alata*, *D. transversa* and *D. hastifolia* from a close common ancestor with Hn2. Such a finding needs to be checked with more variable markers and by including more related species. In addition, the species of section *Enanthiophyllum* from Oceania should be included in a global phylogenetic analysis of this section.

The level of separation between *D. nummularia* haplotypes and *D. alata* clearly shows that *D. nummularia* is not the direct ancestor of *D. alata*, despite the results obtained in recent studies (Malapa *et al.*, 2005). Although *D. alata* is the most widely cultivated yam (Egesi *et al.*, 2003; Malapa *et al.*, 2005; Arnau *et al.*, 2009), its origin thus still remains unclear. Broader sampling, including specimens of *D. alata* from different archipelagos in Oceania, South-East Asia and New-Guinea, with wild and cultivated specimens of related species, will be necessary for further research on the origin of this species.

Each of the species investigated in our study displayed at least two haplotypes differing from each other by one to eight mutations, showing evidence for within-species divergence. The haplotypes identified in *D. alata* and *D. transversa* have diverged from a common ancestor. Each of the *D. alata* haplotypes differs by one mutation from their common ancestor, indicating a recent divergence that would be consistent with phylogeography. One of the *D. alata* haplotypes is carried by the three specimens from India and one from Vanuatu, whereas the other one is borne by the specimens from Vanuatu. However, our nuclear markers could not discriminate Indian from Vanuatu genepools. This pattern may result from our small sample size and needs to be confirmed further.

Dioscorea transversa and *D. hastifolia*, both originating from Australia, seem to share a common ancestor after their divergence from the common ancestor with *D. nummularia* and *D. alata*. This pattern is consistent with the geography. The two haplotypes of *D. transversa*, corresponding to specimens collected in the Torres Strait Islands and in eastern Australia, support the hypothesis of a geographical differentiation within this species. The two specimens of *D. hastifolia* are separated by a single mutation leading to two haplotypes, suggesting a recent divergence. These haplotypes probably ensue from the existence of distinct genepools within the species. With regard to the history of *D. hastifolia*, an Australian native species that was probably domesticated by the Aborigines well before the arrival of

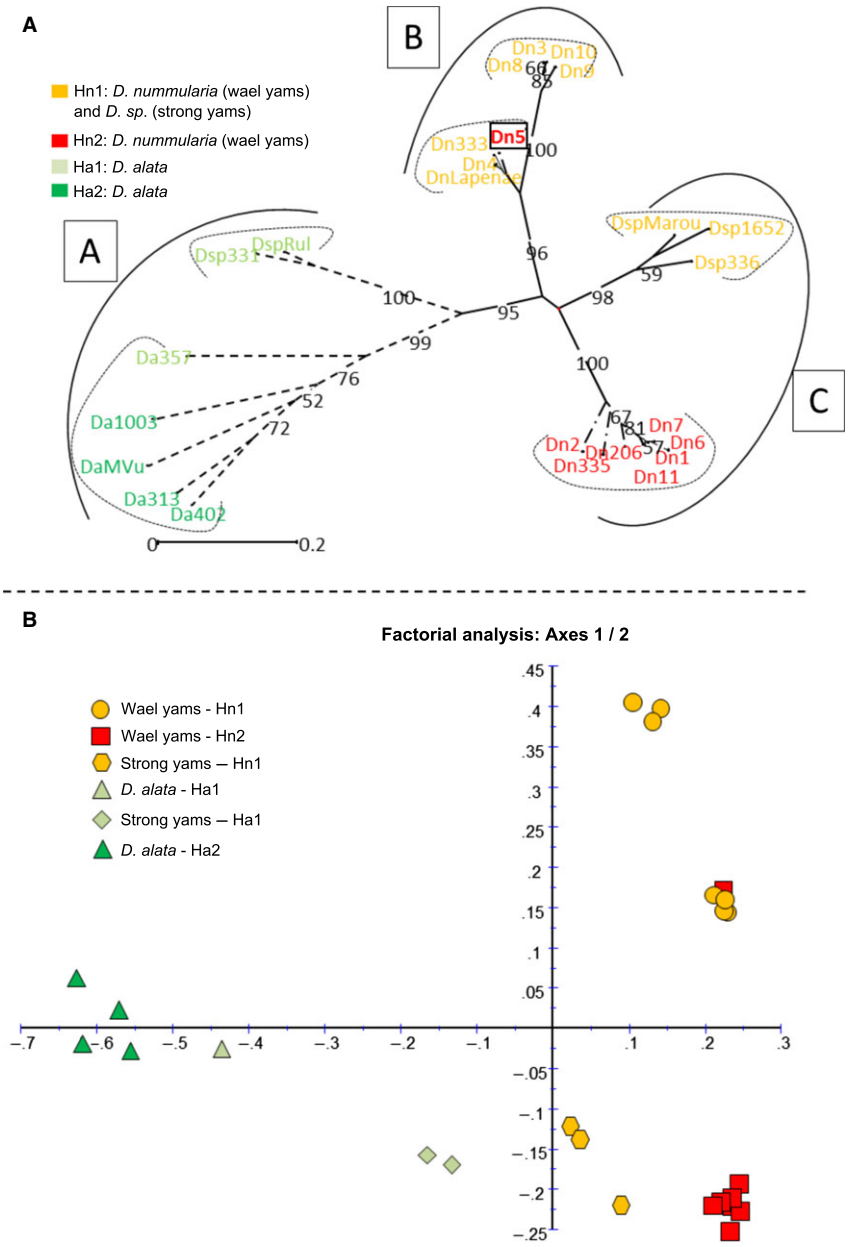

Figure 3. A, unrooted NJ tree based on nine microsatellite markers, using dice distance implemented in the Darwin 5 program, showing the genetic relationships among yam specimens. The tree is based on 500 bootstraps on individuals. Only bootstrap values > 50% are shown. Each branch is coded according to membership into the haplotypes identified previously by plastid sequences. B, PCoA plot of the 25 yam specimens based on microsatellite data. Boxed A, B and C indicate the clusters identified.

Europeans in the 18th century (Walter & Lebot, 2003), the identification of a correlation between the genetic divergence and geography is hazardous. The domestication and diffusion of this species have been poorly investigated even though anthropological studies report its past cultivation. Our data are thus insufficient to further explain the presence of two haplotypes in this species.

Is *Dioscorea nummularia* still obscure?

According to Lebot (2009), *D. nummularia*, a high-yielding species with agronomic potential, is a polymorphic species that remains obscure and has not been studied thoroughly. The plastid sequences used in the present study allow the identification of two main haplotypes among the *D. nummularia*

specimens. Hn1 encompassed wael yams and all strong yams except Rul and Dsp331. Haplotype Hn2 includes only wael yams and the herbarium specimen of *D. nummularia* collected in Papua New-Guinea. This is the first time that two haplotypes have been identified in the semi-perennial plants of this species. It has been suggested that misidentified species could be included under this binomial (Lebot, 2009), but the main dichotomy described previously in *D. nummularia* was related to the annual and semi-perennial split, i.e. strong yams and wael yams, reported in the Pacific Island farming systems such as Futuna (Kirch, 1994) and Vanuatu (Malapa, 2005; Sardos, 2008). Such a split is not confirmed by our molecular data, but our results suggest a much more complex pattern. The nuclear markers have shown that the Hn1 haplotype was split into three subclusters with one formed only of strong yams, whereas the two others contained only wael yams. One of them included one Hn2 specimen (Dn5). The position of the latter among Hn1 haplotype specimens and the intermediate position of a small group of *D. nummularia*, including Dn5, between Hn1 and Hn2 on the PCoA suggests the occurrence of gene flow between the two haplotypes. Thus, the wael yam specimens, Hn1 or Hn2 haplotypes, are probably interfertile and hybridization between these groups seems possible. The presence of fertile wild forms of *D. nummularia* in Vanuatu was reported previously by Lebot (2002), supporting our results. By contrast, the Hn2 specimens formed one distinguishable subgroup with seven specific alleles. Moreover, the fact that both the plastid and the nuclear markers are able to discriminate the Hn2 group suggests that gene flow between specimens of Hn2 and Hn1 haplotypes is not important enough to avoid their genetic differentiation even though they evolve in the same environment. Therefore, wael yams, classified as belonging to *D. nummularia*, seem heterogeneous. This dichotomy was supported by both plastid and nuclear markers. It is clear that in Vanuatu, two different genepools, if not two different taxonomic entities coexist under the same name and are managed by the same farmers in the same forests and gardens. However, the partition reported previously in *D. nummularia* is stated here not based on the cultivation cycle but rather on the haplotypes and genotypes. More investigation including cytology, plastid and nuclear molecular identification backed up with farmers' knowledge, documentation and botanical comparison between the different haplotypes identified, similar to studies that have been conducted on other *Dioscorea* spp. (Wilkin *et al.*, 2009), will be necessary for further determination of their respective taxonomic status.

When dealing with cultivated species, it is extremely difficult to assess their origin clearly: numerous human migrations led to the dispersion of the genepools across wide geographical distances. In Oceania, domestication of traditional crops is believed to have occurred in New Guinea during the early and mid-Holocene and to have been further followed by dispersal throughout Oceania as settlers colonized the Pacific islands (Lebot, 1999; Bird, Hope & Taylor, 2004). Whether local genepools of these crops, including yams, existed in the islands prior to human settlement is not clearly assessed and probably depends on which crop species are being considered. Molecular investigation revealed that the genepool of local Micronesian breadfruit (*Artocarpus altilis* Fosberg; Moraceae) has probably contributed to its current diversity (Zerega, Ragone & Motley, 2004), whereas the decrease of taro genetic diversity from Melanesia in the west to Polynesia in the east suggests an introduction in Oceania from a single Papuan genepool (Mace *et al.*, 2010). Given the presence of the New Guinea herbarium specimen of *D. nummularia* in haplotype Hn2, we may assume that it is native to this large island and was introduced to the Pacific islands by human settlers. This is supported by the closeness in the network of Hn2 and *D. transversa*, originating from Australia, with the common ancestor of the species studied. It raises the question of the origin of Hn1 which seems, according to the network, to have evolved from Hn2 independently of *D. alata*, *D. hastifolia* and *D. transversa*. Whether Hn1 has been introduced, by humans, in Vanuatu from an exotic genepool or has locally evolved from a common ancestor with Hn2 that would have naturally been introduced in the archipelago is difficult to assess.

Independently of the origin of their occurrence in Vanuatu, whether the two *D. nummularia* haplotypes identified in our study represent different genepools of the same species or belong to different taxa is unclear. A larger sampling of specimens classified as *D. nummularia* across its distribution range including Vietnam, Philippines, New Guinea and Pacific islands (http://e-monocot.org), supported by a systematic morphological description and documentation of farmers' practices, should contribute to shed light on the taxonomic position and origin of haplotypes Hn1 and Hn2.

Farmers' use of natural hybridization between *D. alata* and *D. nummularia*

In Vanuatu, farmers do not classify their yams according to the Linnean taxonomy, but rather according to their vegetative and tuber morphologies and uses. Architecture (number and colour of the stems, spinescence) and morphology of the aerial

organs (shape, size, texture and leaf colour) are sufficient for farmers to distinguish between different cultivated and spontaneous forms (Malapa, 2005). Nevertheless, the farmers' classification is often congruent with taxonomy (Sardos, 2008). Strong yams are cultivated yams producing tubers with high dry matter content compared with the most widely cultivated species *D. alata*. Farmers are thus able to distinguish them based on their crop cycle and tuber quality. However, confusion remains regarding their formal taxonomic classification. They are classified either as *D. nummularia* (Thaman, 1988; Kirch, 1994) or *D. transversa* (Malapa *et al.*, 2006) but never as *D. alata*, whereas in our study, strong yams displayed two haplotypes, Hn1 and Ha1. Nuclear markers revealed that strong yams share alleles with both *D. alata* and *D. nummularia* (haplotype Hn1 or Hn2), whereas they display an intermediate position between *D. alata* and wael yams (Hn2) in the PCoA. This suggests that they may have emerged through natural hybridization between wael yams and *D. alata*. Although *D. alata* was introduced in these islands, it was observed that it flowers profusely in Vanuatu. In addition, the presence of *D. nummularia* and *D. alata* haplotypes in strong yams suggests that gene flow probably occurs in both directions. Therefore, confusion reported in the classification of strong yams, such as the erroneous assignation of the cultivar 'Marou' to *D. transversa*, is probably related to its potential hybrid status. Natural hybridization among closely related species in sympatric populations commonly produces complex patterns of morphological variation (Lopez-Caamal *et al.*, 2013), as reported for other species (Jiang *et al.*, 2013; Kim *et al.*, 2014).

Our results suggest strongly the occurrence of natural interspecific hybridization between *D. alata* and *D. nummularia* in Vanuatu, even though high genetic differentiation between the two groups of yams was found. Indeed, only three alleles are shared between *D. alata* and wael yams whether they have Hn1 or Hn2 haplotypes. Such natural hybridization would be exceptional, but the hybrids seem to have been identified, valued and selected by farmers. Our results, if confirmed, are of great interest in clarifying the evolution and taxonomy of yams grown in traditional agrosystems. In addition, they could be useful for crop improvement programmes as *D. alata* is one of the most important yam species for food security in developing countries.

ACKNOWLEDGEMENTS

This work was financially supported by Agropolis Foundation under the project 'Sunda or Sahul? Greater yam origin' (no. 1403-023). We thank the two reviewers for their valuable comments. We thank the National Herbarium of Australia (CANB) for providing herbarium specimens of *D. nummularia*, *D. hastifolia* and *D. transversa*. We thank Valentin Guignon for assistance with informatics. Our greatest thanks go to the farmers in Vanuatu for their interest in participating in our study and providing us with valuable information and plant materials.

REFERENCES

Andris M, Aradottir GI, Arnau G, Audzijonyte A, Bess EC, Bonadonna F, Bourdel G, Bried J, Bugbee GJ, Burger PA, Chaïr H, Charruau PC, Ciampi AY, Costet L, Debarro J, Delatte H, Dubois MP, Eldridge MDB, England PR, Enkhbileg D, Fartek B, Gardner MG, Gray KA, Gunasekera RM, Hanley SJ, Havil N, Hereward JP, Hirase S, Hong Y, Jarne P, Jianfei Q, Johnson RN, Kanno M, Kijima A, Kim HC, Kim KS, Kim WJ, Larue E, Lee JW, Lee JH, Li C, Liao M, Lo N, Lowe AJ, Malausa T, Malé PJG, Marko MD, Martin JF, Messing R, Miller KJ, Min BW, Myeong JI, Nibouche S, Noack AE, Noh JK, Orivel J, Park CJ, Petro D, Prapayotin-Riveros K, Quilichini AE, Reynaud B, Riginos C, Risterucci AM, Rose HA, Sampaio I, Silbermayr K, Silva MB, Tero N, Thum RA, Vinson CC, Vorsino A, Vossbrinck CR, Walzer C, White JC, Wieczorek A, Wright M. 2010. Permanent genetic resources added to molecular ecology resources database 1 June 2010–31 July 2010. *Molecular Ecology Resources* **10:** 1106–1108.

Arnau G, Nemorin A, Maledon E, Abraham K. 2009. Revision of ploidy status of *Dioscorea alata* L. (Dioscoreaceae) by cytogenetic and microsatellite segregation analysis. *Theoretical and Applied Genetics* **118:** 1239–1249.

Ayensu ES. 1972. *Anatomy of the monocotyledons. VI. Dioscoreales*. Oxford: Oxford University Press.

Bandelt HJ, Forster P, Rohl A. 1999. A median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* **16:** 37–48.

Bedford S. 2006. *Pieces of the Vanuatu puzzle: archaeology of the north, south and centre. Terra Australis (23)*. Acton: The Australian National University Press.

Benson G. 1997. Sequence alignment with tandem duplication. *Journal of Computational Biology* **4:** 351–367.

Bird M, Hope G, Taylor D. 2004. Populating PEP II the dispersal of humans and agriculture through Austral-Asia and Oceania. *Quaternary International* **118–119:** 145–163.

Borsch T, Quandt D. 2009. Mutational dynamics and phylogenetic utility of noncoding chloroplast DNA. *Plant Systematics and Evolution* **282:** 169–199.

Bourret D. 1973. Etude ethnobotanique des Dioscoréacées alimentaires. Ignames de Nouvelle Calédonie. D. Phil. Thesis, Paris University.

Bousalem M, Viader V, Mariac C, Gomez RM, Hochu I, Santoni S, David J. 2010. Evidence of diploidy in the wild

Amerindian yam, a putative progenitor of the endangered species *Dioscorea trifida* (Dioscoreaceae). *Genome* **53:** 371–383.

Burkill IH. 1960. The organography and evolution of the Dioscoreaceae, the family of yams. *Botanical Journal of Linnean Society* **56:** 319–420.

Caddick LR, Rudall PJ, Wilkin P, Hedderson TAJ, Chase MW. 2002. Phylogenetics of Dioscoreales based on combined analyses of morphological and molecular data. *Botanical Journal of Linnean Society* **138:** 123–144.

Chaïr H, Perrier X, Agbangla C, Marchand JL, Dainou O, Noyer JL. 2005. Use of cpSSRs for the characterisation of yam phylogeny in Benin. *Genome* **48:** 674–684.

Chaïr H, Cornet D, Deu M, Baco MN, Agbangla A, Duval MF, Noyer JL. 2010. Impact of farmer selection on yam genetic diversity. *Conservation Genetics* **11:** 2255–2265.

Clarke PA. 2007. *Aboriginal people and their plants*. Dural Delivery Centre: Rosenberg Publishing.

Coursey DG. 1976. The origins and domestication of yams in Africa. In: Harlan JR, ed. *Origins of African plant domestication*. The Hague: Mouton, 383–408.

Cruz VMV, Ramirez DA. 1999. Variation in morphological characters in a collection of yams (*Dioscorea* spp.). *Philippines Journal of Crop Science* **24:** 27–35.

Davidovich O, Kimmel G, Shamir R. 2007. GEVALT: an integrated software tool for genotype analysis. *BMC Bioinformatics* **8:** 36.

Denham T. 2008. Traditional forms of plant exploitation in Australia and New Guinea: the search for a common ground. *Vegetation History and Archaeobotany* **17:** 245–248.

Dice L. 1945. Measures of the amount of ecologic association between species. *Ecology* **26:** 297–302.

Dounias E. 2001. The management of wild yam tubers by the Baka pygmies in southern Cameroon. *African Study Monographs Supplement* **26:** 135–156.

Egesi CN, Asiedu R, Egunjobi JK, Bokanga M. 2003. Genetic diversity of organoleptic properties in water yam (*Dioscorea alata* L). *Journal of the Science of Food and Agriculture* **83:** 858–865.

Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39:** 783.

Gammage W. 2009. Galahs. *Australian Historical Studies* **40:** 275–293.

Girma G, Hyma KE, Asiedu R, Mitchell SE, Gedil M, Spillane C. 2014. Next-generation sequencing based genotyping, cytometry and phenotyping for understanding diversity and evolution of guinea yams. *Theoretical and Applied Genetics* **127:** 1783–1794.

Govaerts R, Wilkin P, Saunders RMK. 2007. *World checklist of Dioscoreales, yams and their allies*. Kew: Kew Publishing.

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

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology* **59:** 307–312.

Hahn SK. 1995. Yams. In: Smartt J, Simmonds NW, eds. *Evolution of crop plants*. London: Longman Group, 112–120.

Hall AT. 1999. BioEdit: a user-friendly biological sequence aligment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41:** 95–98.

Hallam SJ. 1975. *Fire and earth: a study of Aboriginal usage and European usurpation in south-western Australia*. Canberra: Australian Institute of Aboriginal Studies.

Hodkinson TR, Chase MW, Takahashi C, Leitch IJ, Bennett MD, Renvoize SA. 2002. The use of DNA sequencing (ITS and *trnL-F*), AFLP and fluorescent *in-situ* hybridisation to study allopolyploid *Miscanthus* (Poaceae). *American Journal of Botany* **89:** 279–286.

Hsu KM, Tsai1 JL, Chen MY, Ku HM, Liu SC. 2013. Molecular phylogeny of *Dioscorea* (*Dioscoreaceae*) in East and Southeast Asia. *Blumea* **58:** 21–27.

Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17:** 754–755.

Jiang J, Zhu M, Ai X, Xiao L, Deng G, Yi Z. 2013. Molecular evidence for a natural diploid hybrid between *Miscanthus sinensis* (Poaceae) and *M. sacchariflorus*. *Plant Systematics and Evolution* **299:** 1367–1377.

Kelchner S. 2000. The evolution of non-coding chloroplast DNA and its application in plant systematics. *Annals of Missouri Botanical Garden* **87:** 482–498.

Kim JO, Tamura MN, Fuse S, Sook LN. 2014. Taxonomic status and phylogeny of *Veratrum* section *Veratrum* (Melanthiaceae) in Korea and Japan based on chloroplast and nuclear sequence data. *Plant Systematics and Evolution* **300:** 75–89.

Kirch PV. 1994. *The wet and the dry: irrigation and agriculture intensification in Polynesia*. Chicago: University of Chicago Press.

Kirch PV. 2000. *On the road of the winds. An archaeological history of the Pacific Islands before European contact*. Berkeley: University of California Press.

Knuth R. 1924. Dioscoreaceae. In: Engler HGA, ed. *Das Pflanzenreich 87 (IV. 43)*. Leipzig: H. R. Engelmann (J. Cramer), 1–387.

Lanfear R, Calcott B, Ho SYW, Guindon S. 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetics analyses. *Molecular Biology and Evolution* **29:** 1695–1701.

Lebot V. 1999. Biomolecular evidence for plant domestication in Sahul. *Genetic Resources and Crop Evolution* **46:** 619–628.

Lebot V. 2002. La domestication des plantes en Océanie et les contraintes de la voie asexuée. *Journal de la Société des Océanistes* **114–115:** 45–61.

Lebot V. 2009. *Tropical root and tuber crops: cassava, sweet potato, yams and aroids*. Wallingford, UK: CABI.

Lebot V, Trilles B, Noyer JL, Modesto J. 1998. Genetic relationships between *Dioscorea alata* L. cultivars. *Genetic Resources and Crop Evolution* **45:** 499–508.

Lopez-Caamal A, Mussali-Galante P, Valencia-Cuevas L, Jiménez Ramirez J, Vega Flores K, Tovar-Sanchez E. 2013. Transgressive character expression in hybrid

zones between the native invasives *Tithonia tubaeformis* and *Tithonia rotundifolia* (Asteraceae) in Mexico. *Plant Systematics and Evolution* **299:** 1781–1792.

Mace ES, Mathur PN, Godwin ID, Hunter D, Taylor MB, Singh D, Delacy IH, Jackson GVH. 2010. Development of a regional core collection (Oceania) for taro, *Colocasia esculenta* (L.) Schott, based on molecular and phenotypic characterization. In: Ramanatha RV, Matthews PJ, Eyzaguire PB, Hunter D, eds. *The global diversity of taro: ethnobotany and conservation*. Rome: Bioversity International, 185–201.

Malapa R. 2005. Description de la diversité de *Dioscorea alata* L. au Vanouatou à l'aide de marqueurs agromorphologiques et moléculaires (AFLP). Relations avec les autres espèces de la section Enantiophyllum. D. Phil. Thesis, Montpellier University.

Malapa R, Arnau G, Noyer JL, Lebot V. 2005. Genetic diversity of the greater yam (*Dioscorea alata* L.) and relatedness to *D. nummularia* Lam. and *D. transversa* Br. as revealed with AFLP markers. *Genetic Resources and Crop Evolution* **52:** 919–929.

Malapa R, Noyer JL, Marchand JL, Lebot V. 2006. Genetic relationship between *D. alata* and *D. nummularia* as revealed by AFLP. In: Motley TJ, Zerega N, Cross H, eds. *Darwin's harvest*. New York: Columbia University Press, 239–265.

Perrier X, Jacquemoud-Collet JP. 2006. DARwin software. Available at: http://darwin.cirad.fr/darwin.

Posada D. 2008. jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution* **25:** 1253–1256.

Risterucci AM, Grivet L, N'goran JAK, Pieretti I, Flament MH, Lanaud C. 2000. A high density linkage map of *Theobroma cacao* L. *Theoretical and Applied Genetics* **101:**948–955.

Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4:** 406–425.

Sardos J. 2008. Conditions du maintien de la biodiversité des plantes à racines et tubercules cultivées au Vanouatou: diversités spécifiques et variétales et gestion de ces diversités. D. Phil Thesis,Montpellier University.

Särkinen T, Staats M, Richardson JE, Cowan RS, Bakker FT. 2012. How to open the treasure chest? Optimising DNA extraction from herbarium specimens. *PLoS ONE* **7:** e43808.

Sastrapradja S. 1982. *Dioscorea alata*: its variation and importance in Java, Indonesia. In Miege J, Lyonga SN, eds. *Yams. Ignames*. Oxford: Clarendon Press, 45–49.

Scarcelli N, Tostain S, Vigouroux Y, Agbangla C, Daïnou O, Pham JL. 2006. Farmers' use of wild relative and sexual reproduction in a vegetatively propagated crop. The case of yam in Benin. *Molecular Ecology* **15:** 2421–2431.

Shaw J, Lickey EB, Schilling EE, Small RL. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *American Journal of Botany* **94:** 275–288.

Taberlet P, Gielly L, Pautou G, Bouvet J. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* **17:** 1105–1109.

Takahashi M, Matsuda F, Margetic N, Lathrop M. 2003. Automated identification of single nucleotide polymorphisms from sequencing data. *Journal of Bioinformatics and Computational Biology* **1:** 253–265.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution* **30:** 2725–2729.

Thaman RR. 1988. Fijian agroforestry: trees, people, and sustainable polycultural development. In: Overton J, ed. *Rural Fiji*. Suva: IPS/USP, 31–58.

Untergasser A, Nijveen H, Rao X, Bisseling T, Geurts R, Leunissen JAM. 2007. Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Research* **35:** W71–W74.

Viruel J, Catalàn P, Segarra-Moragues JG. 2012. Disrupted phylogeographical microsatellite and chloroplast DNA patterns indicate a vicariance rather than long distance dispersal origin for the disjunct distribution of the Chilean endemic *Dioscorea biloba* (Dioscoreaceae) around the Atacama Desert. *Journal of Biogeography* **39:** 1073–1085.

Walter A, Lebot V. 2003. *Jardins d'Océanie*. Paris: IRD.

Wilkin P, Schols P, Chase MW, Chayamarit K, Furness CA, Huysmans S, Rakotonasolo F, Smets E, Thapyai C. 2005. A plastid gene phylogeny of the yam genus, *Dioscorea*: roots, fruits and Madagascar. *Systematic Botany* **30:** 736–749.

Wilkin P, Muasya AM, Banks H, Furness CA, Vollesen K, Weber O, Demissew S. 2009. A new species of yam from Kenya, *Dioscorea kituiensis*: pollen morphology, conservation status, and speciation. *Systematic Botany* **34:** 652–659.

Zerega NJC, Ragone D, Motley TJ. 2004. Complex origins of breadfruit: implications for human migrations in Oceania. *American Journal of Botany* **91:** 760–766.