Title

Cowpea [Vigna unguiculata (L.) Walp.] maternal lineages, chloroplast captures, and wild cowpea evolution

Journal

Genetic Resources and Crop Evolution, 68(7)

ISSN

0925-9864

Authors

Pasquet, Remy S
Feleke, Yonas
Gepts, Paul

Publication Date

2021-10-01

DOI

10.1007/s10722-021-01155-y

Peer reviewed

Genet Resour Crop Evol
https://doi.org/10.1007/s10722-021-01155-y

RESEARCH ARTICLE

Cowpea [*Vigna unguiculata* (L.) Walp.] maternal lineages, chloroplast captures, and wild cowpea evolution

Remy S. Pasquet · Yonas Feleke · Paul Gepts

Received: 11 July 2020 / Accepted: 1 March 2021

Abstract *Vigna unguiculata* is the major grain legume in Africa and is also important in South-East Asia as a vegetable. The wild relatives, present only in Africa, are currently made of ten subspecies. However, while the different groups are well identified, the relationships between them are not firmly established. The present study of chloroplast diversity through restriction fragment length polymorphism confirms the different groups, as well as the split between savannah subspecies and forest subspecies. It justifies treating subsp. *stenophylla* as a subspecies independent from subsp. *protracta* and invalidates the subsp. *stenophylla sensu lato* concept. It suggests that the species *V. unguiculata* originated in southern Africa and that the annual subsp. *unguiculata* var. *spontanea* originated in eastern Africa. More importantly, it proves that three subspecies are of hybrid origin, between the annual var. *spontanea* and various perennial subspecies from southern Africa with left-twisted keel. Owing to its annual and inbred properties, var. *spontanea* spread across all African low- to medium-altitude savannahs, displacing the left-twisted keel savannah perennials from their warmer and drier areas.

Keywords Cowpea · *Vigna unguiculata* · CpDNA RFLP · Infraspecific phylogeny · Chloroplast capture

Introduction

Cowpea, *Vigna unguiculata (*L.) Walp., is one of the main tropical grain legumes. Most of the production originates in the Sudanian and Sahelian zones of sub-Saharan Africa from Senegal to Ethiopia and from Angola to Mozambique. Substantial quantities of cowpea are also produced in intertropical America (largely in semiarid northeastern Brazil) and dry South Asia. Cowpea has been grown in southern Europe during Antiquity and medieval times but, after the discovery of the New World, it was largely replaced by common bean, *Phaseolus vulgaris* L. Cowpea is also important in East Asia, where yard long bean (*V. unguiculata* cv.-gr. Sesquipedalis) green pods are considered as one of the top ten Asian vegetables.

Cowpea is typically grown in equatorial and subtropical lowlands but is replaced by common bean (*P. vulgaris*) at altitudes above 1300–1600 m. Cowpea

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10722-021-01155-y.

R. S. Pasquet (✉)
DIADE, IRD, Univ Montpellier, Montpellier, France
e-mail: remy.pasquet@ird.fr

Y. Feleke · P. Gepts
Department of Plant Sciences, Section of Crop and Ecosystem Sciences, University of California, Davis, CA, USA

Published online: 22 March 2021

is adapted to high temperatures and is tolerant to low fertility soils. Dry grain for human consumption is the main objective of cowpea growers but leaves, fresh grains, and fresh green pods (especially in East Asia) are also consumed. In earlier times, cowpea from cv.-gr. Textilis was grown for the fibers of its long floral peduncles. In addition, cowpea can be cultivated for green manure (in the USA and in Australia) and cowpea haulms provide important fodder to ruminants especially in the Sahel regions of West and Central Africa (Ehlers and Hall 1997; Pasquet and Baudoin 2001; Timko et al. 2007; Xu et al. 2010; Boukar et al. 2016).

Although considered an orphan crop for a long time, cowpea genomics has made considerable progress recently (e.g., Munoz-Amatriain et al. 2017) and genetically-modified lineages are now available (e.g. Citadin et al. 2013; Cruz and Aragao 2014; Bett et al. 2017). However, genetic resources from the wild gene pool have never been used in breeding (Boukar et al. 2016). There are basically two reasons for this situation. The first is the small seed-size associated with the wild forms. Since small seed-size is dominant to large seed size, several backcrosses are required in order to recover the desired seed-size of the domesticated type after a wild x cultivated hybridization. The second reason is that there is still no general consensus on a wild cowpea classification and nomenclature. Of course, taxonomic resolution below the species rank is often challenging (Hardion et al. 2017). However, the different *V. unguiculata* infraspecific groups have been clearly identified in the early nineties (Pienaar and Van Wyk 1992; Mithen and Kibblewhite 1993; Pasquet 1993a, b, 1997; Padulosi 1993), although the relationships between them are not well established. Pasquet and Padulosi (2012) provided a comprehensive review of the wild cowpea gene pool describing the major trends and comparing the different taxonomic nomenclatures. They also listed the points requiring some clarification.

The cowpea gene pool is currently made of 10 subspecies. Following Verdcourt's (1970) *Vigna unguiculata* infraspecific treatment, these subspecies can be split into two groups. The Mensensis group includes four subspecies from forest areas: subsp. *baoulensis* (A.Chev.) Pasquet in West Africa, subsp. *letouzeyi* Pasquet in Central Africa, subsp. *pawekiae* Pasquet in the highlands from Ethiopia to South Africa, and subsp. *aduensis* Pasquet from northern Ethiopia. The latter is related to subsp. *pawekiae* but whether it should be considered as a separate subspecies or a subsp. *pawekiae* variety is not settled yet. The Dekindtiana group includes 6 subspecies from savannah or grassland areas: subsp. *alba* (G.Don) Pasquet in coastal areas from Congo and Angola, subsp. *dekindtiana* (Harms) Verdc. *sensu stricto* in southern Angola highlands, subsp. *tenuis* (E.Mey.) Maréchal et al. with one part in Zimbabwe–Zambia–Malawi highlands (var. *parviflora* Padulosi nom. nud. or tenuis ZWE in Pasquet 1999) and one part in coastal South Africa and Mozambique (var. *tenuis* or tenuis ZAF in Pasquet 1999), subsp. *pubescens* (R.Wilczek) Pasquet mainly in coastal plains from Mozambique to Kenya, subsp *stenophylla* (Harv.) Maréchal et al. *sensu lato* with one part distributed around the Kalahari area (var. *kgalagadiensis* Mithen or stenophylla BWA in Pasquet 1999), one part mainly in the highveld around Pretoria (subsp. *stenophylla sensu stricto*), and one part found in eastern South Africa [var. *protracta* (E.Mey.) Verdc. or stenophylla ZAF in Pasquet 1999].

The last subspecies is subsp. *unguiculata*. It includes the domesticated forms (var. *unguiculata*) and their wild progenitor [var. *spontanea* (Schweinf.) Pasquet, also known as subsp. *dekindtiana sensu* Verdc. non Harms]. Var. *spontanea* is mainly made of annual and inbred plants but there are perennial var. *spontanea* in the Congo Bateke plateau as well as in the Indian Ocean coastal plains. Whether subsp. *tenuis* should be split into two subspecies and whether subsp. *stenophylla* should be split into three subspecies has yet to be decided. In the end, there is a group of var. *spontanea* accessions mainly from Botswana (spontanea BWA group in Pasquet 1999) that are characterized by very unusual isozymes alleles, including some encountered in subsp. *stenophylla sensu lato*, and some not encountered in any studied perennial subspecies (Pasquet and Padulosi 2012).

The present study of chloroplast DNA polymorphism was undertaken by Pasquet in 1996–1997 and continued by Feleke in 2002–2003. The initial goal was to get a clear infraspecific phylogeny based on chloroplast DNA and to clarify some of the disputed points in the taxonomy. In the end, this study presents interesting contributions that complement the Pasquet (1999) isozyme study of the wild cowpea gene-pool and sheds new light on cowpea evolution, especially on the origin of several subspecies.

Materials and methods

We studied 228 accessions (Supplementary material), 52 domesticated and 176 wild from 12 infraspecific taxa: the subsp. *unguiculata* var. *spontanea*, spontanea BWA group, subsp. *pubescens*, subsp. *alba*, tenuis ZAF group (var. *tenuis*), tenuis ZWE group (var. *parviflora* nom. nud.), stenophylla BWA group (var. *kgalagadiensis*), stenophylla ZAF group (var. *protracta*), stenophylla STE group (or subsp. *stenophylla sensu stricto*), subsp. *baoulensis*, subsp. *letouzeyi*, and subsp. *pawekiae* groups. With the exception of subsp. *stenophylla sensu stricto*, all the taxa were present in Pasquet's (1999) work. *Vigna monantha* Thulin, subsp. *dekindtiana sensu stricto*, and subsp. *aduensis* Pasquet remained unavailable at the time of this study. Accession numbers are the ones used in Pasquet (1999, 2000) and Feleke et al. (2006). All the wild accessions are available from the Meise Botanical Garden (Belgium, http://www.br.fgov.be/research/collections/living/phaseolus/) or IITA (Ibadan, Nigeria, https://www.iita.org/research/genetic-resources/) where passport data can be retrieved.

Total DNA was isolated from fresh or silica dried leaves following the modified CTAB method (Doyle and Doyle 1987). Five to seven micrograms of an aliquot of total DNA from the different cowpea accessions were digested with a series of 27 four- or six-cutter restriction enzymes following the manufacturer's recommendations. The enzymes were *Afl*II, *Alu*I, *Bam*HI, *Ban*I, *Bcl*I, *Bgl*II, *Bst*BI, *Bst*UI, *Dra*I, *Eco*0190, *Eco*RI, *Eco*RV, *Hae*III, *Hha*I, *Hind*III, *Hinf*I, *Hpa*II, *Mbo*I, *Nci*I, *Nsi*I, *Pflm*I, *Pst*I, *Ssp*I, *Taq*I, *Xba*I, *Xho*I, *Xmn*I (NE Biolabs). Digested DNA fragments were separated on 1–1.2% agarose gel in TAE running buffer for 16–18 h at 1 V/cm of gel. The DNA fragments were transferred to Zetabind nylon membranes (AMF-CUNO, Meriden, Conn, USA) following Southern blot (Southern 1975). Finally, transferred DNA fragments were fixed in position by UV cross-linking (GS Gene linker, Bio-Rad) to the membrane.

Eight *Pst*I and three *Sal*I clones (provided by J.D. Palmer, Indiana University) representing almost the entire chloroplast genome of *Vigna radiata* (L.) R.Wilczek were used as probes (Palmer and Thompson 1981). Whole plasmids containing cpDNA insert were radiolabeled following the method of Feinberg and Vogelstein (1983). Hybridizations with ^{32}P-labelled probes were performed at 65 °C overnight in hybridization buffer (5 × SSPE, 7.5% SDS). Sequential washes were carried out at 65 °C for 30 min each in 2 × SSC, 0.5% SDS at 65 °C, 1 × SSC, 0.5% SDS at room temperature, and 0.5 × SSC, 0.5% SDS at 60 °C. The washed nylon membranes were exposed to X-Omat X-Ray film for 1 h to 7 days at − 80 °C. The X-ray film was developed using an automatic X-ray film processor (Konica Medical Imaging Inc., Wayne, NJ, USA). The membranes were re-used up to 8 times after stripping the previous probe by a series of washes in 0.4 N NaOH and 0.1 × SSC, 0.5% SDS, 0.2 M Tris–HCl (pH 7.5) at 42 °C for 30 min each.

Results

Only mutations without missing data were considered in the analysis. Therefore, 38 mutations were considered (Table 1). These mutations were all restriction site mutations and not length mutations because similar changes in fragment size would have been observed with other restriction enzymes when probing the same region.

After a first screening (involving a set of 14 restriction enzymes screened with most of the probes), 38 mutations were identified but the phylogeny still consisted of a large polytomy with seven clusters and a group of accessions not belonging to any cluster. Thus, we decided to try additional restriction enzymes on a subset of 38 accessions to resolve this polytomy. The additional restriction enzymes yielded a total of 93 mutations (Table 2). However, none of the 55 additional mutations was able to reduce the polytomy. Therefore, none of the additional mutations was screened over the 228 accessions and none of them was included in the final analysis. Most of the 93 mutations were observed in the Long Single Copy section, 10 to 14 in the Short Single Copy Section and four to eight in the inverted repeat (Table 2).

Among the mutations not included in the analysis is Vaillancourt and Weeden's (1992) mutation 20 (*Dra*I, s10.6). It was observed in most of the clades (5 out of 20 in clade A, 3 out of 3 in clade B, 2 out of 2 in clade C, in all clade D, 24 out of 26 in clade E, 24 out of 25 in haplotype 12, 23 out of 24 in clade F, 38 out of 71 in clade G). However, with the data available (175 accessions out of 228), mutation 20 introduced

Table 1 Mutations used in the analysis

	IR	IR		1/2 IR	IR								
	p16.2		p 9.7	p18.8		p 7.5	p11.1	p 7.0	p 5.6	p 7.8	s16.5	s13.3	s10.6
Afl II			4			3							
BamH I										6	19	24, 37	
Bcl I									22, 27				
Bgl II													1
BstB I				29			28						18
Dra I	9, 21		36						2, 16	16, 35		5, 13	
EcoR I	25		23										
Hae III									34				
Hinf I												14, 38	
Hpa II									33	33			20
Nci I												7, 8, 17	7
Nsi I			31										
Ssp I							26, 15			10, 12		11, 30	
Xmn I	32												

Shaded areas of the chloroplast were not included in the probes. Half of p18.8 is in the Short Single Copy Region, the other half is in the Inverted Repeat. p7.8 and s16.5 do overlap

numerous homoplasies and reverse mutations. Furthermore, it was not reducing the polytomy. The tree with mutation 20 at the base of a clade joining the current clades A and G was less parsimonious than the one presented in Fig. 1. Therefore, we did not attempt to get a complete set of data and mutation 20 was not used in the final analysis.

The set of 38 mutations generated 25 haplotypes. Four haplotypes were present in a single accession. A single most parsimonious tree was generated (Fig. 1). This unique tree had a consistency index of 1.0, a homoplasy index of 0.0, and a retention index of 1.0. It consisted of a seven-clade polytomy. Forty accessions did not show any of the 38 mutations and made up haplotype 12. The geographic distribution of the accessions from the different clades is shown in Figs. 2 and 3.

Within clade G, 37 of the 52 domesticated accessions were showing haplotype 25. This haplotype 25 was including accessions from all the cultivar-groups. Nine of 10 cv.-gr. Textilis were in clade 25. On the other hand, haplotype 24 included almost exclusively accessions from cv.-gr. Biflora from Ethiopia and cv.-gr. Unguiculata.

Discussion

Comparison with Vaillancourt and Weeden (1992)

Although we studied a much larger set of accessions and mutations, our results remain rather similar to the Vaillancourt and Weeden (1992) results in the end. The clade with their haplotypes 1 to 4 is our clade G. The clade with their haplotypes 5, 6, 16 and 19 is our clade F. We did not observe the clade with their haplotype 7, 9 and 10. MT 320 and MT 365 were not among the accessions we screened and we did not observe their mutation 16 (*Dra*I p16.2) with our MT 99 and MT 102, which are both within our haplotype 12. Their haplotype 18 with MT 55 fits our haplotype 12. The clade with their haplotype 12, 13 and 17 is our clade E. Their haplotype 11 with NI 794 is our clade A and their haplotype 8 with MT 53 is our clade B. In the end, 5 of our clades as well as haplotype 12 were observed in their work, although important clades A and B were each represented by a single accession. However, they had no accession for detecting our clades C and D.

Current status of the *Vigna unguiculata* phylogeny

The main result of this study is a seven-clade polytomy (Fig. 1), which suggests a rapid diversification event

Table 2 Number of mutations provided by each enzyme-probe combination

	IR	IR		1/2 IR	IR							
	p16.2		p9.7	p18.8		p7.5	p11.1	p7.0	p5.6	p7.8/s16.5	s13.3	s10.6	Total
Afl II			1	1		1					1		4
BamH I										4	2		6
Bcl I									2	2			4
Bgl II												1	1
BstB I			1	1			1		2			2	7
Dra I	2		2			1			4	4	1	2	16
EcoR I	1		2					1		1			5
Hae III								1	1				2
Hinf I						1				1	2		4
Hpa II									1	2		1	4
Nci I										3	3	1	7
Nsi I			1			1							2
Ssp I							2			3	2		7
Xmn I	1		1						1	2			5
Alu I													0
Ban I													0
BstU I												1	1
EcoO190									1	2			3
EcoR V										1			1
Hha I			2	1			2						5
Hind III													0
Pflm								1					1
Pst I							1			1			2
Taq I										2			2
Xba I								1	1				2
Xho I				1			1						2
Total	4		10	4		4	7	4	13	28	11	8	

Shaded areas of the chloroplast were not included in the set of probes. Half of p18.8 is in the Short Single Copy Section, the other half is in the Inverted Repeat. p7.8 and s16.5 do overlap

with two forest clades, five savannah clades, and 40 accessions without any cpDNA mutation not belonging to any of these clades (haplotype 12). The distribution of these 40 accessions at the base of the cladogram (Fig. 3) reinforces the idea that the species *Vigna unguiculata* originated in the southern part of Africa (Padulosi 1993; Padulosi and Nq 1997). Isozyme data were suggesting a diversification of forest taxa preceding the diversification of savannah taxa (Pasquet 1999). This is not confirmed here. However, although the Mensensis group is split between two clades, it is well isolated. These two forest clades include only accessions from the Mensensis group and no Mensensis group accession appears elsewhere in the tree. Within the Mensensis group, the three subspecies are monophyletic.

Discontinuity between subsp. *baoulensis* and subsp. *letouzeyi* (Pasquet and Padulosi 2012) is a well known discontinuity between Lower Guinea (subsp. *baoulensis*) and Congolia (subsp. *letouzeyi*) forest centres of endemism around the Cameroonian volcanic line (White 1983; Linder et al. 2012). The discontinuity between Congolia (subsp. *letouzeyi*) and Afromontane (subsp. *pawekiae*) forest centres of endemism (White 1983; Fayolle et al. 2014) is also appearing here (Fig. 2), even if there is no subsp. *letouzeyi* accession from DRC available. Interestingly, the fact that there is no forest taxon linked to the Indian

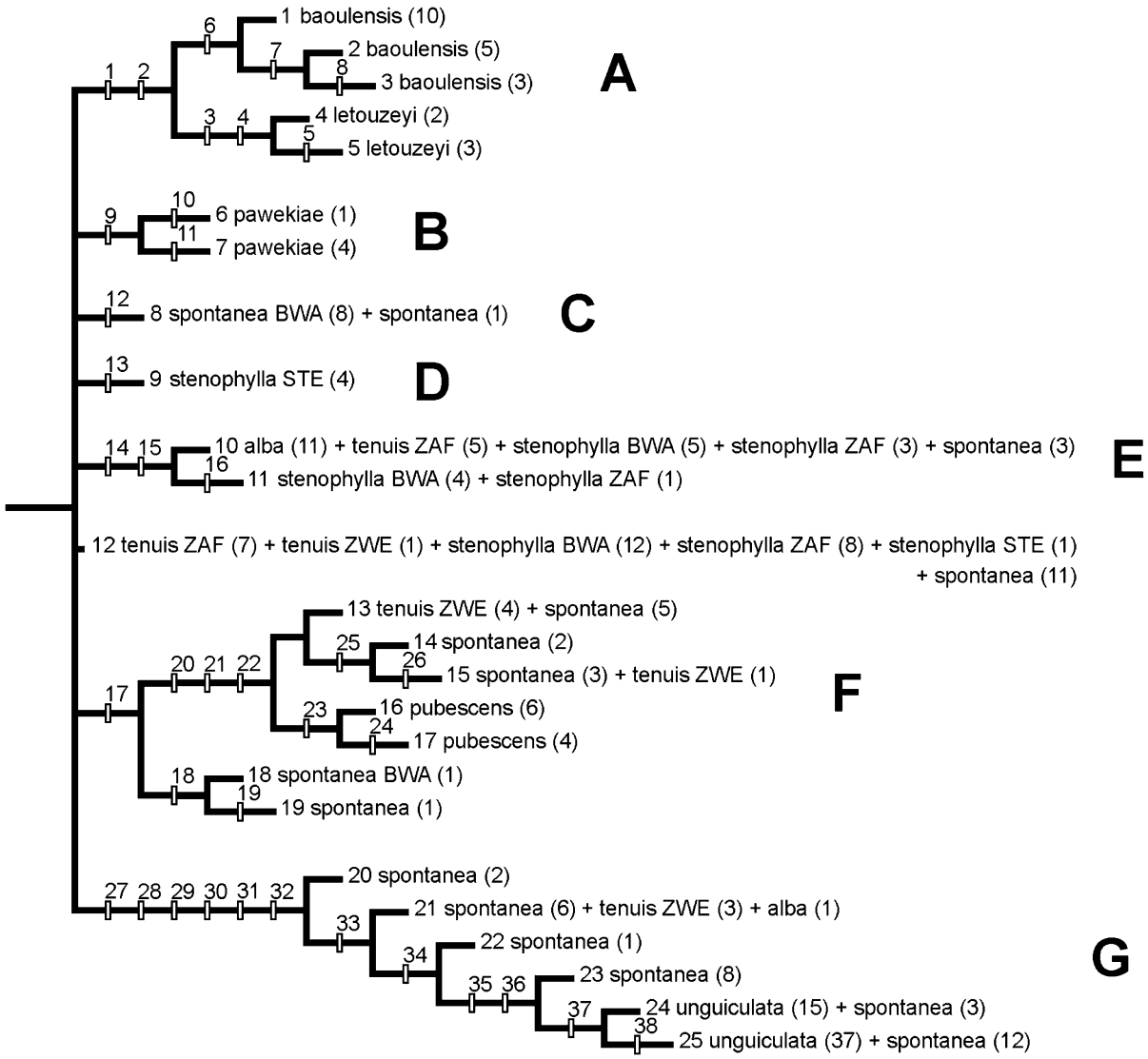

Fig. 1 Cladogram of *Vigna unguiculata* plastome types. Cross-bars in each branch represent different mutations. The first number in the branch labels is the haplotype number. Numbers in parentheses represent the number of entries in each haplotype/taxon

Ocean coastal forests would confirm this recent diversification. The polytomy looks definitely younger than the Miocene and may date back to the Pliocene (Burgess et al. 1998; Couvreur et al. 2008).

Although there is some overlap, the geographic split between clades C, D, E and haplotype 12, on the one hand, and clades F and G (considering only the oldest branches of clade G), on the other hand, is the second main result of this study. The split is clear with no subsp. *stenophylla sensu lato* encountered outside clade D, E and haplotype 12. Such a regional phylogenetic split between East Africa (Zambesian region) and South Africa (Kalahari/Highveld or South African Region) has been highlighted previously (White 1983; Linder et al. 2012) and is reflected in several studies of intraspecific diversity, for example in Lorenzen's et al. (2010) work on arid-adapted mammals. In their study, a number of species have shown varying degrees of mitochondrial lineage differentiation between East and South Africa, a split dated to the Pliocene–Pleistocene transition. With these arid-adapted mammals, the southern region showed a lower level of haplotype structuring than the eastern region. This is verified here with 5

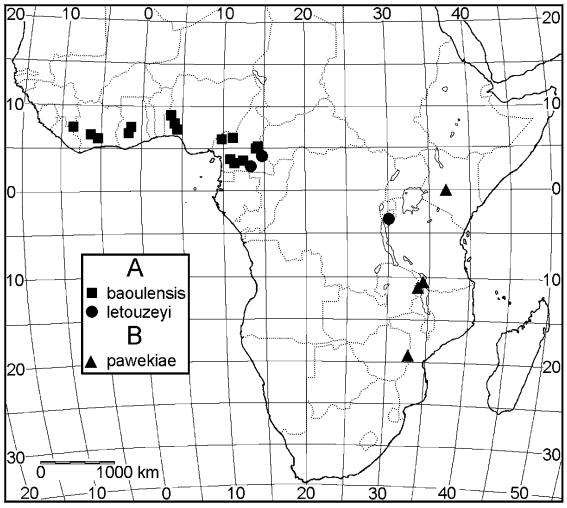

Fig. 2 Geographic distribution of *Vigna unguiculata* A and B clades

haplotypes for clade C, D, and E (southern Africa) and 13 haplotypes for clade F and G (eastern Africa), although this may be also caused by the switch from perennial to annual in clade G. Clade G alone includes one third of the mutations observed (12 out of 38) while var. *spontanea* is a mainly annual taxon. During the same period, the annual var. *spontanea* should have provided more generations than any of the other perennial subspecies, therefore more opportunities for mutations. It is well known that mutations accumulate more in annual plants than in perennial plants with longer generation times (e.g., Eyre-Walker and Gaut 1997; Ma et al. 2012) and that the shift of life history to the annual habit is associated with increases in diversification rates in a number of lineages (Azani et al. 2019).

Lorenzen et al. (2010) explained their results by a gradient between a climatically rather stable southern Africa and a climatically more unstable eastern Africa. Since the shift to the annual habit is a common adaptive strategy in dry and fast-changing environments, perennials tend to occur in cooler and wetter areas, often in mountainous regions, while annuals favor hotter and dryer conditions (Azani et al. 2019). This explains why the shift to annual habit occurred in eastern African cowpea, according to the accessions at the bottom of clade G. This makes clade G the northernmost clade of the savanna clades. If the species *Vigna unguiculata* originated in southern Africa, the annual var. *spontanea* originated in eastern Africa.

Discrepancies between nuclear and chloroplast data

Since there are no strong genetic barriers between the Dekindtiana group subspecies (Echikh 2000), relationships in this group may be more difficult to identify. Indeed, there are numerous discrepancies between nuclear (Pasquet 1999) and the present cpDNA data, indicating either incomplete lineage sorting (ILS) and/or intergroup gene flow leading to chloroplast capture (Rieseberg and Soltis 1991; Maddison 1997; Naciri and Linder 2015).

In ILS, some of the haplotype diversity within a taxon predates the origin of that taxon if it starts out with multiple haplotypes, because of retention of ancestral polymorphism (Maddison 1997). ILS can cause different taxa to share the same chloroplast haplotypes even in the absence of hybridization and usually leads to a general mixture across more or less the entire range of the taxa (e.g., Comes and Abbott 2001 or Wu and Campbell 2005). It seems that the existence of stenophylla BWA and stenophylla ZAF groups sharing haplotypes 10 to 12 should be the result of ILS.

However, all the other phylogenetic incongruences between cpDNA and nuclear DNA results are more likely due to hybridization than to ILS. Indeed, chloroplast capture is expected to be more commonly observed than nuclear gene exchange in plants, due to the smaller effective population size of the haploid plastid genome, making it more likely that a foreign chloroplast genotype will become fixed in populations of the recipient species (Rieseberg and Soltis 1991; Petit et al. 2003; Currat et al. 2008). In addition, cpDNA shuffling among closely related plant species appears to be quite a common phenomenon (e.g. Terry et al. 2000; Cannon and Manos 2003). Phylogeographic studies that cover a plant species complex often show a geographic distribution of the haplotypes irrespective of the constituent taxa (e.g. Byrne et al. 2002; Petit et al. 2002; Durovic et al. 2017; Wan et al. 2017).

First, there are examples of chloroplast captures at the accession level. Some accessions are in clades different from the clade in which the majority of the accessions is. The best example is the subsp. *alba*

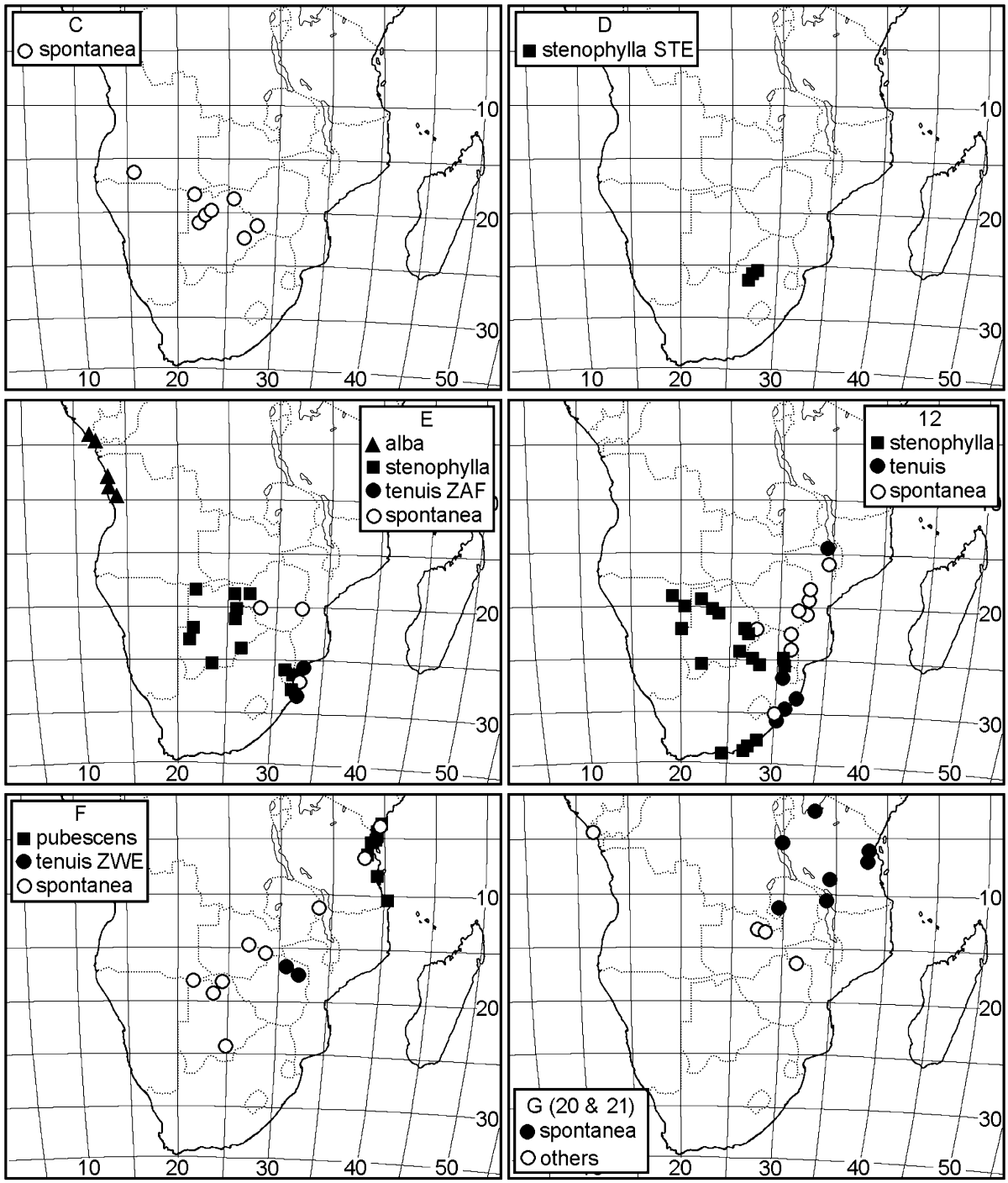

Fig. 3 Geographic distribution of *Vigna unguiculata* C to G clades. Plastome 12 is the one without any mutation. Plastomes 20 and 21 are the ones at the base of clade G. Plastomes 22 to 25 are not mapped

accession SP 173 showing haplotype 21. On the other hand, there are several var. *spontanea* accessions with haplotype 10 and 12 in areas where var. *spontanea* overlaps with subsp. *stenophylla sensu lato*.

More striking is clade C, which is made of accessions from spontanea BWA group (including MT 621 which shows few spontanea BWA alleles only) with an isozyme profile intermediate between var. *spontanea* and an unknown perennial taxon, supposed to be subsp. *dekindtiana sensu stricto* (Pasquet 1999). Very likely, subsp. *dekindtiana* chloroplast was captured in this area (right South of the actual subsp. *dekindtiana sensu stricto* area) by var. *spontanea* plants, and this unique clade C haplotype should be subsp. *dekindtiana sensu stricto* haplotype.

However, beyond these "accession or population" chloroplast captures, some groups may be of hybrid origin and their position in the phylogeny likely due to old chloroplast captures. Since chloroplast capture can happen between either recently diverged or more distantly related species, even genera, it can cause different taxa to share the same chloroplast haplotypes even in the absence of present hybridization. This is especially the case when one species invaded the area occupied by another one until the contact zone arrived at a place where the second lineage had a competitive advantage over the first. If there are no areas in which the second had the competitive advantage, it could have been driven to extinction by the first species (Endler 1977; Rieseberg and Soltis 1991; Tsitrone et al. 2003; Petit et al. 2003).

Groups of hybrid origin

Regarding chloroplast capture at the taxon level, the case of tenuis ZAF group is the easiest to deal with. Geographically, the taxon is encountered in the coastal lowlands from Natal to mid Mozambique, therefore in a buffer area between stenophylla ZAF and var. *spontanea* (Pasquet and Padulosi 2012). Isozyme data show that tenuis ZAF is very close to var. *spontanea* but the alleles with a high frequency in tenuis ZAF and a low frequency in var. *spontanea* are all present in stenophylla ZAF. With all the accessions studied here having a stenophylla ZAF chloroplast, it seems that tenuis ZAF is of hybrid origin, arising from a stenophylla ZAF chloroplast captured by var. *spontanea* and that var. *spontanea*, thanks to tenuis ZAF group, displaced stenophylla ZAF group from its warmer localities.

Regarding subsp. *alba*, all accessions except one are showing haplotype 10, a haplotype dominated by subsp. *stenophylla sensu lato* while isozyme data are showing that subsp. *alba* is closer to var. *spontanea* than to subsp. *stenophylla sensu lato*. Although less obvious than the tenuis ZAF case, this is clearly a second case of chloroplast capture. Within clade F, subsp. *pubescens* is also a chloroplast capture case. It is the closest subspecies to var. *spontanea* while it has predominantly a clade F chloroplast.

Among the subspecies and the groups previously discussed, tenuis ZWE group is also closer to var. *spontanea* than to any other isozyme data group while it shows chloroplast from various clades. The accessions from Harare vicinity belong to clade F, while 3 accessions from Zimbabwe and Zambia belong to clade G and one from Malawi shows haplotype 12. Although not morphologically identified as a subsp. *tenuis* but clustering with subsp. *tenuis* with isozyme data, MT 55 (Zimbabwe) shows haplotype 12. Obviously, tenuis ZWE is of hybrid origin but with various female parents, i.e. subsp. *stenophylla sensu lato*, a clade F taxon, and var. *spontanea*.

Regarding clade F, two thirds of clade F is made up of two taxa of hybrid origin (tenuis ZWE group and subsp. *pubescens*), the remaining third of clade F is made of annual var. *spontanea* accessions having captured a clade F chloroplast, with some perennial var. *spontanea* also clearly from hybrid origin. Obviously, clade F was overwhelmed by var. *spontanea* and the original taxon may be now extinct. Since tenuis ZWE group shows numerous subsp. *stenophylla sensu lato* alleles (as well as most var. *spontanea* accessions from clade F), we can hypothesize that this initial clade F taxon was similar to subsp. *stenophylla sensu lato*. As the two annual accessions at the bottom of clade F are coming from Botswana (MT 612) and Namibia (SP 160), i.e., from the stenophylla BWA area, such a hypothesis appears likely. Hopefully, future cpDNA phylogenetic studies based on whole chloroplast sequences should help. In the meantime, given that clade F original taxon may be considered extinct, analysis of more material may highlight other extinct taxa. The case of the Vaillancourt and Weeden's (1992) clade stemming from their mutation 16 would justify the examination of more annual material from southern Africa.

If we summarize all these cases of chloroplast captures, all of them involve var. *spontanea* on the one hand and perennial taxa from the southern and western parts of southern Africa on the other hand, i.e. subsp. *stenophylla sensu lato*, subsp. *dekindtiana*, and the unknown initial clade F taxon. Indeed, although crosses are not always easy, perennial subspecies can be fertilized by subsp. *unguiculata* pollen and can yield fertile F_1 (Aliyu 2005; Pasquet, unpublished data). Obviously, var. *spontanea* displaced these perennial taxa from their northern and eastern original areas (coastal Congo, Zambia, Zimbabwe, South Mozambique and Indian Ocean coastal areas of South Africa). Since the shift to the annual habit is a common adaptive strategy in dry and fast-changing environments, annuals have often advantages over perennials for colonization and survival in novel habitats (Azani et al. 2019). Indeed, var. *spontanea*'s annual and inbred status provided a large competitive advantage in the warm tropical savannas. Compared to outcrossed perennials easily fertilized by foreign pollen, annual inbred plants are more inbred and are prolific seeders. Inbred annuals can easily outnumber outcrossed perennials when they are sympatric in environments that are not unfavorable to them. In addition, since it is the seeds swallowed by mammals or birds that allow long distance dispersal (Pasquet, unpublished data; Ariani et al. 2018), the prolific seeder var. *spontanea* would be better able to invade perennial territories.

There are numerous additional chloroplast capture cases, mostly by var. *spontanea* capturing perennial chloroplasts and a large number of introgressed var. *spontanea*. These include the perennial var. *spontanea* from the Bateke Plateau and the Indian Ocean coastal plains, and, in southern Africa, some var. *spontanea* with rootstock like MT 55, the numerous annual var. *spontanea* accessions showing 1–2 noded inflorescence with a rather high ovule number (MT 102 for example) or a low ovule number with a multinoded inflorescence (NI 1381 = TVNU 267 for example), which cannot be considered as subsp. *tenuis* plants. In the end, this border area between perennial taxa with left-twisted keel and var. *spontanea*, from Atlantic Ocean to Indian Ocean, is lined with right-twisted keel taxa of hybrid origin (subsp. *alba*, tenuis ZWE group, and tenuis ZAF group) interspersed with numerous var. *spontanea* plants having captured a perennial chloroplast and annual var. *spontanea* plants with clade G chloroplast but showing signs (either through morphology or nuclear DNA) of introgression from left-twisted keel perennials.

Nomenclature consequences

Regarding nomenclature, the synonymy between subsp. *letouzeyi*—subsp. *burundiensis* Pasquet proposed after examination of unpublished ITS sequence data (Pasquet and Padulosi 2012) is confirmed here. The tenuis ZWE and tenuis ZAF groups are two different taxa, with different origins. However, if tenuis ZWE has smaller flowers and a smaller ovule number (Padulosi 1993), there is a morphological overlap and not a sharp difference between the two groups. Therefore, splitting subsp. *tenuis* could be unjustified at this stage, especially since both groups are of hybrid origin with var. *spontanea* being a parent of both groups. Subsp. *stenophylla sensu lato* needs to be split between subsp *stenophylla sensu stricto* and subsp. *protracta* (E.Mey.) Pienaar since all highland subsp. *stenophylla sensu stricto* accessions constitute clade D with the exception of accession SP 362 in clade E, which is a lowland accessions (70 km SE of Tzaneen) that is obviously introgressed. The subsp. *stenophylla sensu lato* concept (Pasquet 1993a) does not seem to be justified any more. However, the status of var. *protracta* and var. *kgalagadiensis* based on pod pubescence difference would require more data.

Cowpea evolution hypothesis

More importantly, a new hypothesis of cowpea evolution (Fig. 4) can be drawn up based on data from the current geographic distribution of the

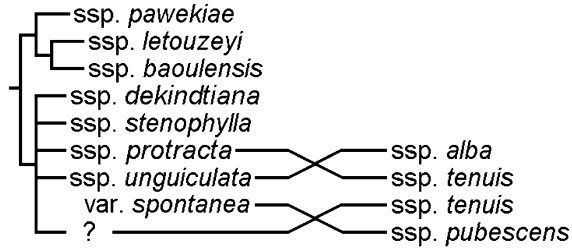

Fig. 4 *Vigna unguiculata* evolution hypothesis. Subsp. *alba*, subsp. *pubescens*, and subsp. *tenuis* are all subspecies of hybrid origin between annual subsp. *unguiculata* var. *spontanea* and various perennial subspecies. The question mark represents the hypothetical perennial taxon leading to clade F

different taxa (Pasquet and Padulosi 2012), the Pasquet (1999) nuclear data, and the present chloroplast data:

1. Since isozyme data (Pasquet 1999) showed the Mensensis group subspecies separated and more diversified than the Dekindtiana group subspecies, the first step should be a split between a forest taxon leading to the Mensensis group subspecies and a savanna taxon leading to the Dekindtiana group subspecies.
2. Then, the present data suggest that the montane forest subsp. *pawekiae* separated from the lowland forest taxon, which was followed by the split between left-twisted keel subsp. *baoulensis* and right-twisted keel subsp. *letouzeyi*.
3. At the same time, the present data suggest that the drier savanna taxon would have split into at least five taxa, one of them becoming inbred and annual (subsp. *unguiculata* var. *spontanea*).
4. In the end, the present data suggest that var. *spontanea* invaded the warmer and dryer areas of the perennial savanna taxa, driving the original clade F taxon to extinction and displacing the other taxa to areas too cold or too dry for var. *spontanea*. In the process, hybridizations gave birth to three hybrid taxa more or less lining the border area between savanna perennials area and var. *spontanea* area, i. e. subsp. *alba*, subsp. *tenuis*, and subsp. *pubescens*.

Conclusion

This work sheds new lights on wild cowpea evolution as it shows that three wild cowpea subspecies are of hybrid origin. In the future, complete sequences of chloroplast from each of the eight clades (considering haplotype 12 as an additional clade) should clear up the polytomy and allow a hierarchy of the seven main taxa. This should allow dating of the different steps of wild cowpea evolution and an assessment whether the spread of annual var. *spontanea* across southern Africa is anterior or contemporary to the spread of cattle rearing by southern Africa first herders. Since it is now possible to use DNA from herbarium material (e.g., Staats et al. 2013), the problem of missing taxa (subsp. *dekindtiana sensu stricto*, subsp. *aduensis*, and *Vigna monantha* from Somalia) should disappear. In the end, this work is a first step forward subsequent to Pasquet's (1999) work but future work should bring a more precise confirmation, including timing, of the present results.

Identification key to the wild subspecies of *Vigna unguiculata*

1a. Keel twisted toward left (Fig. 5)	2
2a. Keel with a marked beak 6–8 mm long (Fig. 5), calyx-lobe 2–5 mm long, flower 24–33 mm long, pod 13–15 ovuled	**subsp. dekindtiana**
2b. Keel with a short beak or without beak (Fig. 5)	3
3a. Calyx-lobe 0.5–2 mm, flower 24–38 mm, pod 16–20 ovuled, pod black and smooth	**subsp. baoulensis**
3b. Pod scabrous or pubescent	4
4a. Calyx-lobe 5–15 mm long, flower 20–32 mm long, pod scabrous 15–18 ovuled	5
5a. Stipule 12–27 × 7–10 mm	**subsp. aduensis**
5b. Stipule 9–16 × 3–7 mm	**subsp. pawekiae**
4b. Calyx-lobe 1.5–6 mm long, flower 14–25 mm long, pod 10–15 ovuled, scabrous or pubescent	6
6a. Stem and pod scabrous or pubescent, leaflet rhomboid or with conspicuous lobes, wider than 8 mm	**subsp. protracta**
6b. Stem and pod scabrous, never pubescent, leaflet linear, up to 8 mm wide, not or slightly lobed	**subsp. stenophylla**
1b. Keel twisted toward right with a short beak up to 3 mm long (Fig. 5)	
7a. Pubescent stem, leaflet, and pod, long inflorescence internodes, calyx-lobe 1.5–5 mm, flower 17–24 mm, pod 13–17-ovuled	**subsp. pubescens**
7b. Scabrous or smooth stem and pod, short inflorescence internodes	8
8a. Inflorescence 1–2-noded, calyx-lobe 1–4 mm, flower 14–22 mm, pod 12–17-ovuled	**subsp. tenuis**
8b. Inflorescence multinoded	9
9a. Seed 3–6 mm long, calyx-lobe 4–15 mm, flower 23–30 mm, pod 17–21-ovuled	**subsp. letouzeyi**
9b. Calyx-lobe 0.5–4.5 mm, flower 15–23 mm	10

10a. Seed 2–3 mm long, calyx-lobe 0.5–4.5 mm, flower 17–23 mm, pod 16–22-ovuled	**subsp. alba**
10b. Seed 3–5 mm long, calyx-lobe 1.5–4 mm, flower 15–23 mm, pod 10–18-ovuled	**subsp. unguiculata var. spontanea**

Fig. 5 *Vigna unguiculata* keel morphology. Subsp. *dekindtiana* long keel beak from specimen Welwitsch 2264 (BM) (top left), subsp *unguiculata* var. *spontanea* short keel beak from accession SP 87 (top center), subsp. *baoulensis* without keel beak from accession SP 136 (top right), subsp. *alba* keel twisted to the right from accession SP 74, the keel beak hide the stigmatic surface (bottom left), subsp. *baoulensis* keel twisted to the left from accession SP 136, the absence of keel beak makes the stigmatic surface visible (bottom right)

Funding This study was completed with the help of grant 2000 GI 087 from the Rockefeller Foundation.

Compliance with ethical standards

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

References

Aliyu B (2005) Crossability of *Vigna rhomboidea* Burtt. Davy with cowpea (*V. unguiculata* (L.) walp.). Genet Resour Crop Evol 52:447–453

Ariani A, Berny Mier y Teran J, Gepts P (2018) Spatial and temporal scales of range expansion in wild *Phaseolus vulgaris*. Mol Biol Evol 35:119–131

Azani N, Bruneau A, Wojciechowski MF, Zarre S (2019) Miocene climate change as a driving force for multiple origins of annual species in *Astragalus* (Fabaceae, Papilionoideae). Mol Phylogenet Evol 137:210–221

Bett B, Gollasch S, Moore A, James W, Armstrong J, Walsh T, Harding R, Higgins TJV (2017) Transgenic cowpeas (*Vigna unguiculata* L. Walp) expressing Bacillus thuringiensis Vip3Ba protein are protected against the Maruca pod borer (*Maruca vitrata*). Plant Cell Tissue Organ Cult 131:335–345

Boukar O, Fatokun CA, Huynh BL, Roberts PA, Close TJ (2016) Genomic Tools in Cowpea Breeding Programs: Status and Perspectives. Front Plant Sci 7:757

Burgess ND, Clarke GP, Rodgers WA (1998) Coastal forests of eastern Africa: status, endemism patterns and their potential causes. Biol J Linnean Soc 64:337–367

Byrne M, MacDonald B, Coates D (2002) Phylogeographical patterns in chloroplast DNA variation within the *Acacia acuminata* (Leguminosae: Mimosoideae) complex in Western Australia. J Evol Biol 15:576–587

Cannon CH, Manos PS (2003) Phylogeography of the Southeast Asian stone oaks (*Lithocarpus*). J Biogeogr 30:211–226

Citadin CT, Cruz ARR, Aragao FJL (2013) Development of transgenic imazapyr-tolerant cowpea (*Vigna unguiculata*). Plant Cell Rep 32:537–543

Comes HP, Abbott RJ (2001) Molecular phylogeography, reticulation, and lineage sorting in Mediterranean *Senecio* sect. *Senecio* (Asteraceae). Evolution 55:1943–1962

Couvreur TLP, Chatrou LW, Sosef MSM, Richardson JE (2008) Molecular phylogenetics reveal multiple tertiary vicariance origins of the African rain forest trees. BMC Biol 6:54

Cruz ARR, Aragao FJL (2014) RNAi-based enhanced resistance to Cowpea severe mosaic virus and Cowpea aphid-borne mosaic virus in transgenic cowpea. Plant Pathol 63:831–837

Currat M, Ruedi M, Petit RJ, Excoffier L (2008) The hidden side of invasions: massive introgression by local genes. Evolution 62:1908–1920

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15

Durovic S, Schonswetter P, Niketic M, Tomovic G, Frajman B (2017) Disentangling relationships among the members of the *Silene saxifraga* alliance (Caryophyllaceae): Phylogenetic structure is geographically rather than taxonomically segregated. Taxon 66:343–364

Echikh N (2000) Organisation du pool genique de formes sauvages et cultivees d'une legumineuse alimentaire, *Vigna unguiculata* (L.) Walp. These de doctorat, Fac Univ Sci Agron, Gembloux

Ehlers JD, Hall AE (1997) Cowpea (*Vigna unguiculata* L Walp). Field Crop Res 53:187–204

Endler JA (1977) Geographic variation, speciation and clines. Princeton University Press, Princeton

Eyre-Walker A, Gaut BS (1997) Correlated rates of synonymous site evolution across plant genomes. Mol Biol Evol 14:455–460

Fayolle A, Swaine MD, Bastin JF, Bourland N, Comiskey JA, Dauby G, Doucet JL, Gillet JF, Gourlet-Fleury S, Hardy OJ, Kirunda B, Kouame FN, Plumptre AJ (2014) Patterns of tree species composition across tropical African forests. J Biogeogr 41:2320–2331

Feinberg AP, Vogelstein B (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132:9–13

Feleke Y, Pasquet RS, Gepts P (2006) Development of PCR-based chloroplast DNA markers to assess gene flow between wild and domesticated cowpea (*Vigna unguiculata*). Plant Syst Evol 262:75–87

Hardion L, Verlaque R, Vorontsova M, Chen CW, Takamizo T, Combroux I, Vila B (2017) Does infraspecific taxonomy match species evolutionary history? A case study of phylogeography in *Arundo formosana* (Poaceae). Bot J Linnean Soc 183:236–249

Linder HP, de Klerk HM, Born J, Burgess ND, Fjeldsa J, Rahbek C (2012) The partitioning of Africa: statistically defined biogeographical regions in sub-Saharan Africa. J Biogeogr 39:1189–1205

Lorenzen ED, Masembe C, Arctander P, Siegismund HR (2010) A long standing Pleistocene refugium in southern Africa and a mosaic of refugia in East Africa: insights from mtDNA and the common eland antelope. J Biogeogr 37:571–581

Ma PF, Guo ZH, Li DZ (2012) Rapid sequencing of the bamboo mitochondrial genome using Illumina technology and parallel episodic evolution of organelle genomes in grasses. PloSOne 7:e30297

Maddison WP (1997) Gene trees in species trees. Syst Biol 46:523–536

Mithen R, Kibblewhite H (1993) Taxonomy and ecology of *Vigna unguiculata* (Leguminosae-Papilionoideae) in South-Central Africa. Kirkia 14:100–113

Munoz-Amatriain M, Mirebrahim H, Xu P, Wanamaker SI, Luo MC, Alhakami H, Alpert M, Atokple I, Batieno BJ, Boukar O, Bozdag S, Cisse N, Drabo I, Ehlers JD, Farmer A, Fatokun C, Gu YQ, Guo YN, Huynh BL, Jackson SA, Kusi F, Lawley CT, Lucas MR, Ma YQ, Timko MP, Wu JJ, You F, Barkley NA, Roberts PA, Lonardi S, Close TJ (2017) Genome resources for climate-resilient cowpea, an essential crop for food security. Plant J 89:1042–1054

Naciri Y, Linder HP (2015) Species delimitation and relationships: the dance of the seven veils. Taxon 64:3–16

Padulosi S (1993) Genetic diversity, taxonomy and ecogeographic survey of the wild relatives of cowpea (*Vigna unguiculata* (L.) Walpers). PhD Thesis, Université catholique Louvain la Neuve

Padulosi S, Ng NQ (1997) Origin, taxonomy, and morphology of *Vigna unguiculata* (L.) Walp. In: Singh BB, Mohan Raj DR, Dashiell KE, Jackai LEN (eds) Advances in cowpea research. IITA-JIRCAS, Ibadan, pp 1–12

Palmer JD, Thompson WF (1981) Clone banks of the mung bean, pea and spinach chloroplast genomes. Gene 15:21–26

Pasquet RS (1993a) Classification infraspécifique des formes spontanées de *Vigna unguiculata* (L.) Walp. à partir de données morphologiques. Bull Jard Bot Nat Belg 62:127–173

Pasquet RS (1993b) Two new subspecies of *Vigna unguiculata* (L.) Walp. *(Leguminosae: Papilionoideae)*. Kew Bull 48:805–806

Pasquet RS (1997) A new subspecies of *Vigna unguiculata* (*Leguminosae-Papilionoideae*). Kew Bull 52:840

Pasquet RS (1999) Genetic relationships among subspecies of *Vigna unguiculata* (L.) Walp. based on allozyme variation. Theor Appl Genet 98:1104–1119

Pasquet RS (2000) Allozyme diversity of cultivated cowpea *Vigna unguiculata* (L.) Walp. Theor Appl Genet 101:211–219

Pasquet RS, Baudoin JP (2001) Cowpea. In: Charrier A, Jacquot M, Hamon S, Nicolas D (eds) Tropical plant breeding. Science Publishers, Enfield - CIRAD, Montpellier, pp 177–198

Pasquet RS, Padulosi S (2012) Genus *Vigna* and cowpea (*V. unguiculata* [L.] Walp.) taxonomy: current status and prospects. In: Boukar O, Coulibaly O, Fatokun CA, Lopez K, Tamo M (eds) Innovative research along the cowpea value chain. IITA, Ibadan, pp 66–87

Petit RJ, Csaikl UM, Bordacs S, Burg K, Coartd E, Cottrell J, van Dam B, Deans JD, Dumolin-Lapegue S, Fineschi S, Finkeldey R, Gillies A, Glaza I, Goicoechea PG, Jensen JS, Konig AO, Lowe AJ, Madsen SF, Matyas G, Munro RC, Olalde M, Pemonge MH, Popescu F, Slade D, Tabbener H, Taurchini D, de Vries SGM, Ziegenhagen B, Kremer A (2002) Chloroplast DNA variation in European white oaks—phylogeography and patterns of diversity based on data from over 2600 populations. For Ecol Manag 156:5–26

Petit RJ, Bodenes C, Ducousso A, Roussel G, Kremer A (2003) Hybridization as a mechanism of invasion in oaks. New Phytol 161:151–164

Pienaar BJ, Van Wyk AE (1992) The *Vigna unguiculata* complex (Fabaceae) in southern Africa. S Afr J Bot 58:414–429

Rieseberg LH, Soltis DE (1991) Phylogenetic consequences of cytoplasmic gene flow in plants. Evol Trends Plant 5:65–84

Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 7:503–517

Staats M, Erkens RHJ, van de Vossenberg B, Wieringa JJ, Kraaijeveld K, Stielow B, Geml J, Richardson JE, Bakker FT (2013) Genomic treasure troves: complete genome sequencing of herbarium and insect museum specimens. PLoS ONE 8:e69189

Terry RG, Nowak RS, Tausch RJ (2000) Genetic variation in chloroplast and nuclear ribosomal DNA in Utah Juniper (*Juniperus osteosperma*, Cupressaceae): evidence for interspecific gene flow. Am J Bot 87:250–258

Timko MP, Ehlers JD, Roberts PA (2007) Cowpea. In: Kole C (ed) Genome mapping and molecular breeding in plants, Vol 3, Pulses, sugar and tuber crops. Springer, Berlin, pp 49–67

Tsitrone A, Kirkpatrick M, Levin DA (2003) A model for chloroplast capture. Evolution 57:1776–1782

Vaillancourt RE, Weeden NF (1992) Chloroplast DNA polymorphism suggests nigerian center of domestication for the cowpea, *Vigna unguiculata* (Leguminosae). Am J Bot 79:1194–1199

Verdcourt B (1970) Studies in the Leguminosae-Papilionoideae for the "Flora of Tropical East Africa": IV. Kew Bull 24:507–569

Wan Q, Zheng Z, Huang K, Guichoux E, Petit RJ (2017) Genetic divergence within the monotypic tree genus *Platycarya* (Juglandaceae) and its implications for species' past dynamics in subtropical China. Tree Genet Genomes 13:73

White F (1983) The vegetation of Africa: a descriptive memoir to accompany the UNESCO/AETFAT/UNSO vegetation map of Africa. Natural Resources Research no. 20. UNESCO, Paris

Wu CA, Campbell DR (2005) Cytoplasmic and nuclear markers reveal contrasting patterns of spatial genetic structure in a natural Ipomopsis hybrid zone. Mol Ecol 14:781–792

Xu P, Wu XH, Wang BG, Liu YH, Qin DH, Ehlers JD, Close TJ, Hu TT, Lu ZF, Li GJ (2010) Development and polymorphism of *Vigna unguiculata* ssp *unguiculata* microsatellite markers used for phylogenetic analysis in asparagus bean (*Vigna unguiculata* ssp *sesquipedialis* (L.) Verdc.). Mol Breed 25:675–684