1 Origin and diversification of winged bean (*Psophocarpus tetragonolobus* (L.) DC.; Fabaceae) a multi-
2 purpose underutilized legume [1]

3

4 Shuyi Yang[2,3], Aurélie Grall[4] and Mark A Chapman[2,5,6]

5

6 [2] Biological Sciences, University of Southampton, Life Sciences Building 85, Highfield Campus,
7 Southampton, SO17 1BJ, UK

8 [3] College of Forestry and Landscape Architecture, South China Agricultural University, Guangdong
9 510642, China

10 [4] Africa & Madagascar Team, Identification & Naming Department, Royal Botanic Gardens, Kew TW9
11 3AE, UK

12 [5] Centre for Underutilised Crops, University of Southampton, Highfield Campus, Southampton, SO17
13 1BJ, UK

14 [6] Author for correspondence: Tel: +44 (0)2380 594396, Email: m.chapman@soton.ac.uk

15

16 ABSTRACT

17 Premise of the study - For many crops, research into the origin and partitioning of genetic variation
18 is limited and this can slow or prevent crop improvement programs. Many of these underutilized
19 crops have traits which could be of benefit in a changing climate due to stress tolerance or
20 nutritional properties. Winged bean (*Psophocarpus tetragonolobus* (L.) DC.) is one such crop. All
21 parts of the plant can be eaten, from the roots to the seeds, and is high in protein as well as other
22 micronutrients. The goal of our study was to identify the wild progenitor and analyze the partitioning
23 of genetic variation in the crop.

24 Methods - We used molecular phylogenetic analyses (cpDNA and nuclear ITS sequencing) to resolve
25 relationships between all species in the genus, and population genetics (utilizing microsatellites) to
26 identify genetic clusters of winged bean accessions and compare this to geography.

27 Key results - We find that winged bean is genetically distinct from all other members of the genus.
28 We also provide support for four groups of species in the genus, largely, but not completely,
29 corresponding to the results of previous morphological analyses. Within winged bean, population
30 genetic analysis using 10 polymorphic microsatellite markers suggests four genetic groups; however
31 there is little correspondence between the genetic variation and the geography of the accessions.

32 Conclusions - The true wild progenitor of winged bean remains unknown (or is extinct). There has
33 likely been large-scale cross-breeding, trade and transport of winged bean and/or multiple origins of
34 the crop.

35

36

INTRODUCTION

With climate change and an increasing world's population, research is needed into nutritious and diverse crops with the ability to grow under stressful or sub-optimal conditions (Schmidhuber and Tubiello, 2007; Campbell et al., 2016). Half of the world's calorie intake comes from just three crops (wheat, rice and corn), yet these cereals are relatively poor in protein and some micronutrients. Legumes complement cereals in the diet well (Foyer et al., 2016); they supply considerable protein, essential amino acids and micronutrients, and serve as a substitute for meat when this is unavailable or unaffordable (Tharanathan and Mahadevamma, 2003). Several legumes are grown for forage, with the high protein content being of benefit to the livestock.

In addition to the nutrient content, legumes have evolved the ability to fix atmospheric nitrogen meaning that fertilizer input can be lower, legumes can often tolerate poor soils, and they can improve the soil for subsequent crops (Peoples et al., 1995). This poses advantages for resource-poor farmers who cannot obtain or afford fertilizers (Siddique et al., 2012). However legumes remain understudied relative to cereal crops (Foyer et al., 2016).

Several legumes are described as underutilized (or orphan) crops; i.e., crops which are locally important but essentially absent from international trade and having received little research attention (Cullis and Kunert, 2017). These underutilized legumes are typically grown on small scale for feeding families, especially in times of harsh environmental conditions, therefore are vitally important despite their lack of international recognition. Although these crops have received comparatively little research and funding, their importance has been recognized for decades (e.g., NAS, 1975b). In the past decade several genome and transcriptome sequences have been published for underutilized legumes, for example the pigeon pea *Cajanus cajan* (L.) Huth genome (Varshney et al 2012) and lablab *Lablab purpureus* (L.) Sweet, winged bean *Psophocarpus tetragonolobus* (L.) DC., Bambara groundnut *Vigna subterranea* (L.) Verdc. and grasspea *Lathyrus sativus* L. transcriptomes (Chapman, 2015; Vatanparast et al., 2016) with the ability to identify genes and variants associated with adaptive phenotypes (e.g., Varshney et al., 2017). A recent initiative to sequence 100 genomes of 100 orphan crops (http://africanorphancrops.org/) has the potential to increase this amount of genomic information further.

One example of an underutilized legume is the winged bean (*Psophocarpus tetragonolobus* (L.) DC.), which has been recognized by a number of authors and institutes as having much promise for nutritional security in the coming decades. Winged bean is an annual or perennial vine which thrives in hot and humid tropical conditions. It is grown widely, but mainly on a local scale, throughout much of Asia, especially India, Southeast Asia, Indonesia and Papua New Guinea, as well as in some parts of Africa (Verdcourt and Halliday, 1978). Considerable morphological variation exists in winged bean especially in Southeast Asia, Indonesia and Papua New Guinea (Khan, 1976; Harder and Smartt, 1992).

One notable feature of the winged bean is the potential for almost all parts of the plant to be eaten, from the seeds, pods and flowers, to the leaves and tuberous roots (Haq 1982; Khan 1976), with the stems and leaves used as fodder. It also serves well in crop rotation due to the nitrogen fixation capability (Rahman et al., 2014). As with several underutilized crops, however, there are anti-nutrition factors in winged bean (specifically trypsin inhibitors), requiring thorough soaking, rinsing

79 and cooking of the dried beans (NAS, 1975a), which may deter more widespread adoption of winged
80 bean.

81 Protein content of the winged bean seeds is high (30-39%, NAS, 1975a; 38-45%, Prakash et al., 1987)
82 and comparable to that of soybean. Content of essential amino acids is generally as high as, or
83 higher than, soybean (Prakash et al., 1987). It is also relatively high in vitamins A and C, calcium and
84 iron (USDA Food Composition Databases, 2015). In countries where protein deficiency is high, or
85 access to meat protein is low, winged bean is a candidate for helping to diversify diets and improve
86 nutrition (NAS, 1975a). Three wild relatives of the winged bean are cultivated, but on much smaller
87 scales, two in Eastern Africa (*P. palustris* Desv. and *P. grandiflorus* R.Wilczek) and one in Southeast
88 Asia, Brazil and Jamaica (*P. scandens* (Endl.) Verdc.) (Zeven and de Wet, 1982; Harder and Smartt,
89 1992). Knowing more about the relationships between the species of *Psophocarpus* is important to
90 inform future winged bean breeding and improvement.

91 The wild progenitor of winged bean has remained somewhat enigmatic given the absence of wild
92 *Psophocarpus* in Asia, leading to one suggestion that the true wild progenitor is now extinct
93 (Verdcourt and Halliday, 1978). Morphological phylogenetic analyses of the nine species in the genus
94 have come to varied conclusions. In particular, the position of *P. grandiflorus*, a wild species found
95 throughout East Africa from the Democratic Republic of the Congo to Ethiopia, is inconsistent. Some
96 studies suggest this species is closely related to, if not the progenitor of, winged bean (Verdcourt
97 and Halliday, 1978; Smartt, 1980; Maxted, 1990; Harder and Smartt, 1992), whereas others place *P.
98 grandiflorus* more distantly related in the genus (Nur Fatihah et al., 2012). The most recent
99 morphological analysis places winged bean alongside *P. scandens* and *P. palustris* in subgenus
100 *Psophocarpus*, with *P. scandens* the closest wild species to winged bean (Nur Fatihah et al., 2012).
101 Few attempts to cross winged bean with other members of the genus have been reported, however
102 one successful cross between winged bean and *P. scandens* has been made following several
103 attempts (Nazmul Haq, University of Southampton, unpublished data).

104 As yet a molecular phylogenetic analysis is lacking. Identifying the true wild progenitor(s) may assist
105 in the breeding of winged bean and would be necessary to understand the genetic changes
106 associated with domestication. Relatively few studies have investigated the domestication genetics
107 of legumes, except for studies of common bean (e.g., Bitocchi et al., 2012; Schmutz et al., 2014), and
108 therefore little is known about the genes and alleles that were under selection by early farmers. This
109 may contribute to the observation that genetic enhancement of legumes remains slow relative to
110 other crops (Foyer et al., 2016). As a first step to remedy this lack of investigation, more work is
111 needed to identify wild ancestors of legume crops. Breeding with crop wild relatives has the
112 potential to introduce novel variation for crop improvement (reviewed in Dempewolf et al., 2017).

113 Crop improvement can also come from selectively breeding different varieties with complementary
114 qualities, for example morphological characteristics, nutrition, abiotic tolerance and pest resistance.
115 An understanding of the partitioning of genetic variation in crops is therefore important to begin to
116 identify and locate genetically distinct varieties and potentially understand the genetic basis of traits
117 of interest (Hawkes, 1991; Mickelbart et al., 2015).

118 In this investigation we addressed the following objectives: First, using a molecular phylogenetic
119 approach we investigate the relationships between members of the genus *Psophocarpus*, in

120   particular focusing on identifying the close relative(s) of winged bean. Second, employing a
121   population genetic approach we use microsatellite markers to resolve the partitioning and levels of
122   genetic variation in the crop.

123

124   MATERIALS AND METHODS

125   Seeds from 53 accessions of winged bean (*Psophocarpus tetragonolobus*) were procured from the
126   USDA (https://www.ars-grin.gov/; 23 accessions), IITA (http://my.iita.org/accession2/; 14
127   accessions), and NARO (http://www.gene.affrc.go.jp/databases-plant_search_en.php; 16 accessions)
128   (Appendix S1A; see Supplemental Data with this article). Seed from 10 accessions of wild species of
129   *Psophocarpus* were obtained from the USDA (2 accessions), ILRI
130   (http://192.156.137.110/forage/frgdsearch.asp; 6 accessions) and the Millennium Seed Bank
131   Partnership (http://apps.kew.org/seedlist/; 2 accessions), and commercially for a single accession of
132   *Erythrina crista-galli* L., the outgroup (based on an *rbcL* phylogeny; Kajita et al., 2001; Appendix S1B).
133   Taxonomic status of two of the accessions from ILRI was uncertain. Seed were scarified using a razor
134   blade or sandpaper and placed in a 2:1 mix of Levington's F2+S: vermiculite. Pots were placed in the
135   greenhouse at the University of Southampton under 16h days (supplemented by fluorescent light)
136   and watered daily. Seed was not available for several species in the genus and so material from
137   herbarium sheets was obtained from Kew (5 accessions) and Paris (3 accessions) (Appendix S1B).

138   After 3 months a trifoliate leaf from half way up the main stem was removed from all of the winged
139   bean plants and dried flat. Leaves were scanned and the area and perimeter of the terminal leaflet
140   measured using ImageJ (Rasband, 1997-2013). Most plants did not flower and so we are unable to
141   analyze characteristics associated with the flowers or pods. After identifying genetic clusters (see
142   below) we tested for significant differences in leaf morphology (apical leaflet area and perimeter)
143   between the groups (1-way ANOVA with *post hoc* Tukey test).

144   *DNA extraction*

145   After 2-4 weeks of growth a leaf sample (ca. 100 mg) was removed from each accession for DNA
146   extraction following standard CTAB protocol (Doyle and Doyle, 1990) with minor modification
147   (Chapman and Burke, 2012). For the herbarium samples about 10mg of tissue was extracted. DNA
148   samples were quantified using a Nanodrop (Nanodrop Products, Wilmington, DE, USA) and by
149   running a sample on an agarose gel stained with GelRed (Biotium, Hayward, California, USA).

150   *PCR sequencing and phylogenetic analysis*

151   Twelve universal primer pairs for amplifying intergenic and non-coding regions of the chloroplast (cp)
152   genome (Appendix S2A) and one pair of primers for the nuclear internal transcribed spacer (ITS)
153   region (White et al., 1990) were tested for amplification on five DNA samples (two *P. palustris*, two *P.*
154   *scandens* and one *P. lancifolius* Harms). PCR amplification took place using a touchdown program
155   with 55°C final annealing temperature (see Chapman and Burke, 2012 for details). PCR products
156   were checked for amplification on agarose gels stained with GelRed (Biotium, Hayward, California,
157   USA). Successful and clean PCR products were purified using standard protocols and sequenced
158   using BigDye v3 reagents (see Chapman et al., 2008) on an ABI3730xl at the Department of Zoology,
159   University of Oxford.

160 Three cpDNA regions (the *psbC-trnS* and *psbM2-trnD*-GUC intergenic regions and the *rpS16* intron;
161 Appendix S2A) and the ITS region were selected for analysis of the complete set of DNAs on the basis
162 of clean amplification and the presence of sequence polymorphisms. For several herbarium DNAs
163 the PCR failed and internal primers were designed to allow amplification of a shorter portion of the
164 locus or to amplify the locus in two portions (Appendix S2A).

165 Sequences were checked manually using Chromas (Technelysium Pty Ltd, 1998-2001) and aligned
166 using Clustal Omega (Sievers and Higgins, 2014). The cpDNA sequences were analyzed individually as
167 well as concatenated. Maximum likelihood phylogenetic analysis was carried out using PHYML
168 (Guindon et al., 2010) with the model of substitution determined by Akaike Information Criterion
169 and with 1000 bootstrap replicates. Bayesian analysis was carried out using MrBayes 3.2.3 (Ronquist
170 et al., 2011); using the www.phylogeny.fr online server. The model of substitution used was GTR+I.
171 The Markov Chain was run for 10,000 generations, sampling every 10, with the first 250 discarded as
172 burn-in.

173 *Microsatellite amplification and analysis*

174 Twenty microsatellite primer pairs were designed from an analysis of three winged bean
175 transcriptome sequences. Reads from two winged bean transcriptomes (Vatanparast et al., 2016)
176 were downloaded from the NCBI short read archive (SRA; accession numbers SRR3039625 and
177 SRR3039626) and assembled together using trinity (Grabherr et al., 2011) and parameters used
178 previously (Chapman, 2015). Microsatellites were identified using misa.pl (http://pgrc.ipk-
179 gatersleben.de/misa/). Reads from a third winged bean transcriptome (Chapman, 2015); SRA
180 accession number SRS826754) were then mapped to the reference to produce bam files which were
181 converted to fasta files using samtools (Li et al., 2009). The sequences were aligned and
182 microsatellite-containing contigs were manually checked for length differences between individuals
183 in the microsatellite region. Twenty contigs with evidence for length variation were used for primer
184 design using primer3 (Rozen and Skaletsky, 2000) (Appendix S2B). A further nine primers from
185 (Wong et al., 2017) were also screened for variation (Appendix S2B).

186 Initially eight winged bean DNAs were amplified to check for reliable and clean amplification and for
187 the presence of polymorphism. PCR used the '3 primer' method (Schuelke, 2000) and were labelled
188 with the fluorophores 6FAM or TET but were otherwise amplified and resolved on agarose as for the
189 cpDNA loci. Post-PCR, loci were combined such that different sized and labelled loci could be
190 resolved together. Samples were diluted 1:30 and run on an ABI 3730xl at the Department of
191 Zoology, University of Oxford with the GeneScan™ 500 size standard. Primer pairs which produced
192 easy to score polymorphic bands were taken forward for amplification on the full set of DNAs.

193 Microsatellite alleles were scored using Genemarker V2.4.0 (http://www.softgenetics.com/) and
194 converted into a matrix for further analyses. Cervus (Kalinowski et al., 2007) was used to provide
195 summary statistics for the data. Genetic distances and Principal Coordinates Analysis (PCO) were
196 computed in GenAlEx V6.5 (Peakall and Smouse, 2006) and a minimum evolution (ME) phylogenetic
197 tree was computed with FastME 2.0 (Lefort et al., 2015; http://www.atgc-montpellier.fr/fastme/)
198 using the same distance matrix. STRUCTURE (Falush et al., 2003) was used to infer genetic clusters in
199 the data; the number of clusters (*K*) was tested from 1 to 8 with five replicate runs (200,000
200 iterations after 50,000 burn-in iterations) per *K*. The most likely number of clusters was determined

201 using the Δ*K* method of Evanno et al. (2005) as calculated in STRUCTURE HARVESTER (Earl and von
202 Holdt, 2012). Genetic variation and differentiation statistics were calculated using GenAlEx ver. 6.5
203 (Peakall and Smouse, 2006).

205 RESULTS

206 DNA was extracted from 53 accessions of winged bean for the microsatellite analysis. Four of these
207 plus a further 18 individuals of other *Psophocarpus* species (one to four individuals per species) and
208 one *Erythrina crista-galli* individual were used in the phylogenetic analyses.

209   *Phylogenetic relationships in* Psophocarpus

210 Three out of twelve cpDNA loci were amplified on the full set of DNAs using internal primers for the
211 majority of samples (Table 1). The three alignments were less than 300bp each due to the inability to
212 amplify larger fragments from the herbarium samples. Nonetheless they provided sufficient
213 numbers of informative sites (Table 1) to create supported, if not fully-resolved, phylogenetic trees.
214 Phylogenetic trees of the individual loci are shown in Appendix S3 and the ML and Bayesian
215 phylogenies of the concatenated data are shown in Fig. 1A and B. We were unable to generate a
216 clean sequence for *psb*C-*trn*S in *Erythrina* or for *rpS*16 in the herbarium sample of *P. grandiflorus*
217 from Ethiopia (gra ETH); hence these are omitted from the relevant individual locus phylogenies. The
218 ITS alignment was 678 nucleotides in length, of which 308 sites were parsimony informative (Table 1;
219 largely because of the high number of fixed differences between winged bean and the other
220 *Psophocarpus* accessions). Unfortunately we were unable to amplify the ITS region from any samples
221 of *P. monophyllus.* ML and Bayesian phylogenetic trees are shown in Fig. 1C, D. Genbank numbers
222 are given in Appendix S1.

223 ML phylogenetic relationships using each cpDNA locus individually provided consistent support for
224 four main groups: 1, winged bean; 2, *P. palustris* and *P. scandens* (along with one of the unknown
225 samples); 3, *P. monophyllus* Harms, *P. lecomtei* Tisser. and *P. obovalis* Tisser., and 4, *P. lancifolius*
226 and *P. lukafaensis* (De Wild.) R. Wilczek (plus the second unknown sample)*.* For two of three loci, *P.*
227 *grandiflorus* was allied to the *P. lancifolius - P. lukafaensis* group and for two loci winged bean was
228 sister to the *P. palustris - P. scandens* group (Appendix S3). Concatenating the cpDNA data added
229 support for these groupings under both ML and Bayesian analysis, however the relationships
230 between these groups differs between these analyses (Fig. 1A, B). The main difference is the
231 placement of winged bean, with ML indicating that winged bean's most close African relative is
232 either *P. palustris* or *P. scandens* whereas in the Bayesian analysis winged bean is sister to all other
233 members of the genus.

234 The relationships between taxa in the ITS phylogeny recapitulated those in the cpDNA phylogenies,
235 again indicating four clades, and supporting the sister group relationship of winged bean to the rest
236 of the genus. In both the cpDNA and ITS analyses, the phylogenetic distance between winged bean
237 and the other taxa is relatively large compared to that between most other species (Fig. 1).

238 No sequence diversity was revealed for these three cpDNA fragments or the ITS region amongst the
239 four winged bean samples and divergence within each of the other three groups was absent or
240 minimal. The one exception is the presence of two subgroups in the *P. palustris - P. scandens* group

241 (Fig. 1). This splits *P. palustris* into two clades, one of which contains only two *P. palustris* individuals
242 and the other groups two *P. palustris* individuals with *P. scandens* and one individual of unknown
243 taxonomy (sp1102).

Population genetics of winged bean

245 Five of the 20 primers designed from the combined transcriptomes produced reliable, clean and
246 polymorphic bands. Five further primer pairs from Wong et al. (2017) were also selected based on
247 the same criteria (Appendix S2B). These amplified between two and 11 alleles each (average 5.4)
248 across the panel of winged bean with polymorphism information content (PIC) ranging from 0.370 to
249 0.724 (average 0.546; Table 2) which is as high as, or higher than, several other investigations of
250 legume population structure (Appendix S4). Observed heterozygosity was low (Table 2) which is not
251 surprising for a crop which is largely inbreeding.

252 PCO of the winged bean microsatellite genotypes revealed little clustering of geographically
253 proximal accessions (Fig. 2). For some countries the number of accessions genotyped was low and so
254 it is difficult to determine any fine-scale structure between many countries, however even if we
255 focus on countries where several accessions (≥ 7) were genotyped (Indonesia, Malaysia, Papua New
256 Guinea) there was still no clustering by country. Whilst the partitioning of genetic diversity is not
257 expected to observe political borders (Meirmans, 2015), it might be expected that if genetic
258 variation was related to geography then some patterns would be observed. This was however not
259 the case; even at a continental scale accessions from Africa, Asia and the Americas are not separated.
260 Geographical co-ordinates (or even location descriptions) are not available for many accessions and
261 so a Mantel test (Mantel, 1967) was not possible.

262 In STRUCTURE each individual is assigned proportional membership into each of the *K* clusters.
263 Analysis of the STRUCTURE results suggested that the most likely number of clusters (*K*) was four (Fig.
264 3), however we also present the STRUCTURE results for *K* values of 2 and 3 (Appendix S5) because
265 other levels of clustering might reveal biologically relevant grouping (Meirmans, 2015). If we take 80%
266 membership into a single cluster to represent an individual which is not admixed, then 44 individuals
267 are non-admixed for *K* = 4, with 6, 14, 16, 8 individuals in clusters 1, 2, 3 and 4, respectively. Genetic
268 variation (Table 3), calculated as unbiased expected heterozygosity, was greatest in cluster 3 (0.503),
269 intermediate in clusters 1 and 2 (0.456 and 0.461, respectively) and lowest in cluster 4 (0.307).
270 Genetic differentiation (F_{ST}; Table 3), was lowest between clusters 2 and 3 (0.137) and greatest
271 between clusters 1 and 4 (0.361).

272 Each of the four clusters contains individuals from the three main geographic regions (Asia, Africa,
273 the Americas) with the exception of cluster 1 which only contains Asian (two Indonesian, one Papua
274 New Guinean, two Bangladeshi and one Sri Lankan) winged bean individuals (Fig. 3). A χ^2 analysis
275 revealed that the distribution of accessions into the four clusters (plus a fifth category of unclustered)
276 was essentially random in relation to the geographic origin of the accessions (χ^2 = 9.329, d.f. = 8, *P* =
277 0.315). This indicates that genetic structure in winged bean is not related to geography.

278 However, there was variation in leaflet size when comparing between plants in the STRUCTURE
279 clusters. Leaflets from plants in cluster 2 were significantly smaller ($F_{3,43}$ = 3.74, P = 0.018, 1-way
280 ANOVA) with a smaller perimeter ($F_{3,43}$ = 3.02, P = 0.040, 1-way ANOVA) than those in group 4, with
281 plants in clusters 1 and 3 intermediate (Fig. 4).

282　The ME tree (Appendix S6) demonstrates that the accessions from the Americas (blue branches in
283　Appendix S6) and from Africa (black branches) are not monophyletic and are instead interspersed
284　amongst the Asian accessions, in line with the PCO and STRUCTURE results. Because the tree is
285　unrooted it is impossible to confirm the order of the branching events. The accessions grouped into
286　the four clusters by STRUCTURE correspond well, but not perfectly, with the ME tree groups (see
287　shading of the accession names in the ME tree; Appendix S6). For example, five of six accessions
288　assigned to cluster 1 form a monophyletic group in the ME tree and 13 of the 14 accessions assigned
289　to cluster 2 form a monophyletic group (albeit with one cluster 3 and three admixed individuals also
290　found in this group). Admixed individuals (based on STRUCTURE) are found scattered throughout the
291　ME tree.

292

293　DISCUSSION

294　　　*Phylogenetic relationships in* Psophocarpus

295　The four clades (Fig. 1) identified in the molecular phylogenetic analyses largely group species by
296　geography. Winged bean appears genetically distinct and is the only species found in Asia (with a
297　few reports of *P. scandens* although this is introduced and cultivated). The clade comprising
298　*Psophocarpus grandiflorus, P. lancifolius* and *P. lukafaensis* represents the East African species (*P.
299　grandiflorus* and *P. lancifolius* have overlapping distributions in Uganda/Kenya and *P. lancifolius* and
300　*P. lukafaensis* overlap in Zambia/Botswana; (Verdcourt and Halliday, 1978). The clade comprising *P.
301　monophyllus*, *P. lecomtei* and *P. obovalis* represents Western (*P. monophyllus*) and Central (Central
302　African Republic, Democratic Republic of Congo; *P. lecomtei* and *P. obovalis*) African species, all with
303　restricted distributions (Verdcourt and Halliday, 1978). The final clade comprises the widespread
304　species *P. palustris* (West to Central Africa) and *P. scandens* (Central and South Eastern Africa),
305　which overlap in their distributions in Benin/Nigeria/Cameroon (Verdcourt and Halliday, 1978).
306　Although precise identification of the two unknown samples from the Democratic Republic of Congo
307　is not possible yet, they appear to be allied to the *P. lancifolius* group (sp1091) and the *P. palustris -
308　P. scandens* group (sp1102). The lack of resolution between some of the species suggests recent
309　divergence and evolution within each group, but relatively more ancient divergence between the
310　groups.

311　Several aspects of the molecular-based groupings are consistent with previous morphological
312　analyses. For example (Verdcourt and Halliday, 1978; Maxted, 1990; Nur Fatihah et al., 2012) all
313　proposed close relationships between *P. lancifolius* and *P. lukafaensis*, between *P. monophyllus*, *P.
314　lecomtei* and *P. obovalis*, and between *P. palustris* and *P. scandens* (Table 4). The differences
315　between these three studies can be attributed to the placement of *P. grandiflorus* and winged bean,
316　and in fact our study is not completely in line with any of these scenarios.

317　Verdcourt and Halliday (1978) suggested that *P. grandiflorus* was the progenitor of winged bean, and
318　Maxted (1990) also suggested a close relationship between these taxa, as well as with *P. palustris*
319　and *P. scandens*. The analysis by Nur Fatihah et al. (2012), however, placed *P. grandiflorus* sister to
320　the remainder of the genus (Table 4). Our findings show *P. grandiflorus* is most closely related to,
321　but distinct from, *P. lancifolius* and *P. lukafaensis* (Fig. 1), a different pattern to that suggested by
322　any of the morphological analyses (Table 4).

323 In addition to the *P. grandiflorus* vs. *P. lancifolius - P. lukafaensis* split, only one other intra-clade
324 divergence is supported; genetic differences were found within *P. palustris*, with the four individuals
325 split into two clades in both the cpDNA and ITS analyses (Fig. 1). One pair of individuals are on a
326 branch of their own, whereas the other pair are genetically identical to two individuals of *P.*
327 *scandens* and one of the samples of uncertain taxonomic status (sp1102). The first pair of *P. palustris*
328 individuals are from Western Africa (Burkina Faso and The Gambia) where *P. scandens* is absent,
329 whereas the other pair are from Democratic Republic of the Congo and Nigeria where the species
330 distributions overlap, and intermediates have been reported (Verdcourt and Halliday, 1978; Maxted,
331 1990). This could mean that hybridization has taken place in this region of sympatry causing the
332 introgression of cpDNA and the ITS region from *P. scandens* into *P. palustris*. Alternatively, the
333 morphological similarity (Maxted, 1989) and our finding of a lack of consistent genetic
334 differentiation between *P. palustris* and *P. scandens* could instead mean that the taxonomy of these
335 two species needs revision.

336 We did not manage to resolve all interspecific relationships, hence we can only comment on the
337 sectional and subgenus classifications. Based on our results however we propose that the
338 morphological analyses carried out previously (Verdcourt and Halliday, 1978; Maxted, 1990; Nur
339 Fatihah et al., 2012) do not accurately define genetic groupings and therefore may need to be
340 updated. We would like to add more markers before making any commitments to a new
341 classification; however we tentatively suggest four sectional delimitations (Table 4). In particular we
342 highlight that *P. tetragonolobus* is genetically distinct from other members of the genus, in contrast
343 to all morphological analyses which place *P. tetragonolobus* in the same section as *P. palustris* and *P.*
344 *scandens*. The morphological characters which group these taxa (petiole < 5 cm and pods with a
345 lignified exocarp; Nur Fatihah et al., 2012) may therefore represent the ancestral condition in the
346 group, having been lost in the other taxa, or are products of convergent evolution.

347

348 *The origin of winged bean*

349 Whether the closest African relative of winged bean is a member of the *P. palustris - P. scandens*
350 group (as seen in the cpDNA ML analysis and most morphological analyses) or not (our cpDNA
351 Bayesian analysis and the ITS analysis) is unclear, but given the large genetic distance (Fig. 1) it
352 seems that winged bean should probably be considered in a different section. This suggests that
353 identification the direct progenitor of winged bean still eludes us and could imply that Verdcourt and
354 Halliday (1978)'s conclusion, that the progenitor of winged bean is a now-extinct Asian species, is
355 true. Alternatively winged bean as a crop could be morphologically indistinct from a progenitor and
356 so individuals are all identified as the cultivated taxon.

357 No sequence differences in the four winged bean samples were observed and it is not unusual for a
358 crop to exhibit very low sequence variation relative to wild relatives (Dane and Lang, 2004; Wills and
359 Burke, 2006; Besnard et al., 2011) due to the genetic bottleneck typically associated with a
360 domestication event. In this study, this is not because the four sequenced individuals are especially
361 closely related to each other; instead these are present in different regions of the microsatellite-
362 based PCO.

363 *Population genetics of winged bean*

364   Using ten microsatellites we genotyped 53 accessions of winged bean from throughout the
365   distribution in an effort to elucidate genetic structure, should it be present.

366   Our population genetic analyses suggest that genetic structure in winged bean is roughly partitioned
367   into four groups (Fig. 3); however this is unrelated to geography. Instead the four groups typically
368   comprise accessions from throughout the distribution. This seems to be relatively rare in studies of
369   the partitioning of genetic variation within crop species, however other studies of winged bean have
370   come to similar conclusions (Mohanty et al., 2013; Chen et al., 2015). This can be explained by one
371   or more of a diffuse or multiple origins, significant cross-breeding and/or the transport of seed
372   between locations. Nine individuals showed admixture (i.e., no more than 80% ancestry in one of
373   the four clusters) which could indicate recent cross-breeding or an ancestral admixed 'stock' of
374   winged bean from which the other populations are derived. Yet there was no geographic pattern to
375   these nine individuals, coming from all three continents sampled and six different countries.
376   Although we were unable to measure pod and flower characteristics on the winged bean accessions
377   we grew in the greenhouse, we found that leaf size differed between some of the groups found in
378   the STRUCTURE analysis (Fig. 4). If we assume that morphological similarity correlates with genetic
379   similarity then this backs up our finding that the global distribution of winged bean germplasm
380   represents the result of large-scale movement of accessions and/or extensive hybridization with
381   morphologically similar accessions found in disparate geographic locations.

382   The reasons behind the lack of geographic structure in the genetic data are not clear; however we
383   can speculate some reasons for the successful international trade of winged bean which would
384   erode geographic groupings. First, few legume crops are well-adapted to humid tropics therefore
385   winged bean fits a specific niche, encouraging its international trade to countries with specific
386   environments (Khan, 1976). Second, it has been reported that winged bean needs no specific
387   nodulating bacteria (NAS, 1975b) and so this may have aided its expansion into other regions, where
388   other legumes may be slowed or prevented from establishing. A final possibility is that since the
389   renewed interest in the winged bean in the 1970s there has been much seed-sharing between
390   institutes. We found that the passport data for some accessions was either missing or contradictory
391   (these accessions were not used), but this raises the possibility that some of the geographic data we
392   used is incorrect.

393   CONCLUSIONS

394   Using a combination of marker types (from slow-evolving cpDNA loci to fast-evolving nuclear
395   microsatellites) we have been able to investigate the relationship of winged bean to other members
396   of the genus *Psophocarpus* as well as search for population structure in the crop. We show that the
397   wild progenitor of winged bean remains elusive, or that cultivated winged bean is insufficiently
398   diverged from a wild form to be identified as separate taxon. Within the crop there is subtle
399   population structure which is largely unrelated to geography suggesting that winged bean might
400   have rapidly been traded worldwide after a single or multiple origins.

401

402

Table 1 - Locus information for the three cpDNA regions studied. Primer sequences and their sources are given in Appendix S2.

Locus	NI [a]	Locus length	#PI sites [b]	Model [c]
*psb*C-*trn*S	22	290	14	HKY85
*psb*M-*trn*D	23	279	19	GTR
*rp*S16	22	232	33	GTR
Combined cp	n/a	801	68	GTR+I
ITS	20	678	308	TN93+G

Notes:

[a] Number of individuals (out of 23) successfully amplified and sequenced; [b] Number of parsimony informative sites; [c] Model of sequence evolution determined by PhyML (HKY, (Hasegawa et al., 1985)]; GTR, (Lanave et al., 1984)]; TN93, (Tamura and Nei, 1993)]); +I, with the proportion of invariable sites calculated; + G, modelled by a discrete Gamma distribution).

Table 2 - Polymorphism statistics for the ten SSR markers analyzed. Primer sequences are given in Appendix S2.

Locus [a]	NA [b]	NI [c]	Hobs [d]	Hexp [e]	PIC [f]	Min [g]	Max [h]
MAC4	2	53	0.132	0.498	0.372	310	312
MAC7	11	52	0.404	0.715	0.670	221	240
MAC14	4	51	0.196	0.516	0.470	321	330
MAC15	4	46	0.065	0.623	0.538	235	246
MAC17	6	46	0.087	0.402	0.370	240	267
Pt7.2	6	49	0.265	0.770	0.724	424	433
Pt53	6	47	0.170	0.594	0.553	308	324
Pt68.1	4	49	0.245	0.630	0.567	225	237
Pt76.1	5	49	0.306	0.687	0.634	200	209
Pt93.1	6	51	0.255	0.597	0.562	267	279
Average	5.4	49.3	0.213	0.603	0.546		

Notes:

[a] MAC primers were designed here, Pt primers are from Wong et al. (2017); [b] Number of alleles detected; [c] Number of individuals (out of 53) successfully amplified and scored; [d] Observed heterozygosity; [e] Expected heterozygosity; [f] Polymorphic Information Content; [g] Minimum allele size; [h] Maximum allele size

Table 3 – Genetic variation within and differentiation between genetic clusters of winged bean identified using STRUCTURE

| STRUCTURE cluster | uHe | F_{ST} | | | |
		C1	C2	C3	C4
C1	0.456	X	X	X	X
C2	0.461	0.194	X	X	X
C3	0.503	0.210	0.137	X	X
C4	0.307	0.361	0.229	0.230	X

[1] uHe, Unbiased heterozygosity; F_{ST}, Fixation index

Table 4 – Summary of the morphological phylogenetic analyses of the genus *Psophocarpus* alongside tentative taxonomic rearrangements drawn from this study.

Species	(Verdcourt and Halliday, 1978)	(Maxted, 1990)	(Nur Fatihah et al., 2012)	This study
P. lukafaensis	subgen. *Vignopsis*	sect. *Vignopsis*	sect. *Vignopsis*	sect. *Vignopsis*
P. lancifolius	subgen. *Vignopsis*	sect. *Vignopsis*	sect. *Vignopsis*	sect. *Vignopsis*
P. grandiflorus	sect. *Psophocarpus*	sect. *Psophocarpus*	subgen. *Longipedunculares*	sect. *Vignopsis*?
P. tetragonolobus	sect. *Psophocarpus*	sect. *Psophocarpus*	sect. *Psophocarpus*	sect. *tetragonolobus?*
P. scandens	sect. *Psophocarpus*	sect. *Psophocarpus*	sect. *Psophocarpus*	sect. *Psophocarpus*
P. palustris	sect. *Psophocarpus*	sect. *Psophocarpus*	sect. *Psophocarpus*	sect. *Psophocarpus*
P. obovalis	sect. *Psophocarpus*	subgen. *Lophostigma*	subgen. *Lophostigma*	sect. *Lophostigma*
P. lecomtei	sect. *Unifoliolatae*	subgen. *Lophostigma*	subgen. *Lophostigma*	sect. *Lophostigma*
P. monophyllus	sect. *Unifoliolatae*	subgen. *Lophostigma*	subgen. *Lophostigma*	sect. *Lophostigma*

FIGURE LEGENDS

Figure 1 – Phylogenetic trees for winged bean and other *Psophocarpus* species (rooted with *Erythrina*). Species names are followed by the country of origin (ISO 3166-1 alpha-3 code) (see Appendix S1 for further details). (A, B) ML and Bayesian phylogenetic trees for combined cpDNA regions *psbC-trnS*, *psbM-trnD*, and *rpS16*; (C, D) ML and Bayesian phylogenetic trees for the ITS data.

Figure 2 – Principal Coordinates Analysis of the microsatellite data for cultivated accessions of winged bean.

Figure 3 – Results of the STRUCTURE analysis of cultivated winged bean based on ten microsatellites. The most highly supported number of clusters was 4 (inset). Each accession is represented by a horizontal bar and the shading of each bar corresponds to percentage membership to each of the four clusters.

Figure 4 – Apical leaf area (A) and perimeter (B) for winged bean accessions found in each of the four STRUCTURE clusters. Shared letters above the bars indicate no significant difference (1-way ANOVA).

Appendix S1 – List of accessions of winged bean accessions used in this study. (A) Those analyzed for microsatellite variation, and (B) those used in the cpDNA and ITS sequencing.

Appendix S2 – Primer sequences. (A) For cpDNA and ITS sequencing, and (B) for microsatellite analysis**.**

Appendix S3 – Phylogenetic trees for the analyses of individual cpDNA loci. (A) *psbC-trnS*, (B) *psbM-trnD*, and (C) *rpS16*. * indicates support > 70%, ** indicates support > 90%.

Appendix S4 – Comparison of Polymorphic Information Content (PIC) of microsatellite from investigations of cultivated legumes.

Appendix S5 – Results of the STRUCTURE analysis of cultivated winged bean based on ten microsatellites. Results for 2, 3, and 4 clusters are presented. Each accession is represented by a horizontal bar and the colors of each bar correspond to percentage membership to each of the clusters.

Appendix S6 – Minimum Evolution phylogenetic tree for 53 accessions of winged bean.

LITERATURE CITED

BESNARD, G., P. HERNÁNDEZ, B. KHADARI, G. DORADO, and V. SAVOLAINEN. 2011. Genomic profiling of plastid DNA variation in the Mediterranean olive tree. *BMC Plant Biology* 11: 80-80.

BITOCCHI, E., L. NANNI, E. BELLUCCI, M. ROSSI, A. GIARDINI, P. S. ZEULI, G. LOGOZZO, et al. 2012. Mesoamerican origin of the common bean (Phaseolus vulgaris L.) is revealed by sequence data. *Proceedings of the National Academy of Sciences of the United States of America* 109: E788-E796.

CAMPBELL, B. M., S. J. VERMEULEN, P. K. AGGARWAL, C. CORNER-DOLLOFF, E. GIRVETZ, A. M. LOBOGUERRERO, J. RAMIREZ-VILLEGAS, et al. 2016. Reducing risks to food security from climate change. *Global Food Security* 11: 34-43.

CHAPMAN, M. A. 2015. Transcriptome sequencing and marker development for four underutilized legumes. *Applications in Plant Sciences* 3: 1400111.

CHAPMAN, M. A., and J. M. BURKE. 2012. Evidence of selection on fatty acid biosynthetic genes during the evolution of cultivated sunflower. *Theoretical and Applied Genetics* 125: 897-907.

CHAPMAN, M. A., C. H. PASHLEY, J. WENZLER, J. HVALA, S. TANG, S. J. KNAPP, and J. M. BURKE. 2008. A genomic scan for selection reveals candidates for genes involved in the evolution of cultivated sunflower (*Helianthus annuus*). *The Plant Cell* 20: 2931-2945.

CHEN, D., X. YI, H. YANG, H. ZHOU, Y. YU, Y. TIAN, and X. LU. 2015. Genetic diversity evaluation of winged bean (*Psophocarpus tetragonolobus* (L.) DC.) using inter-simple sequence repeat (ISSR). *Genetic Resources and Crop Evolution* 62: 823-828.

CULLIS, C., and K. J. KUNERT. 2017. Unlocking the potential of orphan legumes. *Journal of Experimental Botany* 68: 1895-1903.

DANE, F., and P. LANG. 2004. Sequence variation at cpDNA regions of watermelon and related wild species: implications for the evolution of *Citrullus* haplotypes. *American Journal of Botany* 91: 1922-1929.

USDA FOOD COMPOSITION DATABASES 2015. http://ndb.nal.usda.gov/ [accessed 6th December 2015].

DEMPEWOLF, H., G. BAUTE, J. ANDERSON, B. KILIAN, C. SMITH, and L. GUARINO. 2017. Past and Future Use of Wild Relatives in Crop Breeding. *Crop Science* 57: 1070-1082.

DOYLE, J. J., and J. L. DOYLE. 1990. Isolation of plant DNA from fresh tissue. *Focus* 12: 13-15.

EARL, D. A., and B. M. VON HOLDT. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4: 359-361.

EVANNO, G., S. REGNAUT, and J. GOUDET. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14: 2611-2620.

FALUSH, D., M. STEPHENS, and J. K. PRITCHARD. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164: 1567-1587.

FOYER, C. H., H. M. LAM, H. T. NGUYEN, K. H. M. SIDDIQUE, R. K. VARSHNEY, T. D. COLMER, W. COWLING, et al. 2016. Neglecting legumes has compromised human health and sustainable food production. *Nature Plants* 2.

GRABHERR, M. G., B. J. HAAS, M. YASSOUR, J. Z. LEVIN, D. A. THOMPSON, I. AMIT, X. ADICONIS, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29: 644-U130.

GUINDON, S., J. F. DUFAYARD, V. LEFORT, M. ANISIMOVA, W. HORDIJK, and O. GASCUEL. 2010. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Systematic Biology* 59: 307-321.

HARDER, D. K., and J. SMARTT. 1992. Further evidence on the origin of the cultivated winged bean, *Psophocarpus tetragonolobus* (L) DC (Fabaceae) - chromosome numbers and the presence of a host-specific fungus. *Economic Botany* 46: 187-191.

HASEGAWA, M., H. KISHINO, and T. YANO. 1985. Dating of the human-ape split by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* 21: 160-174.

HAWKES, J. G. 1991. The importance of genetic resources in plant breeding. *Biological Journal of the Linnean Society* 43: 3-10.

KAJITA, T., H. OHASHI, Y. TATEISHI, C. BAILEY, and J. J. DOYLE. 2001. rbcL and Legume Phylogeny, with Particular Reference to Phaseoleae, Millettieae, and Allies. *Systematic Botany* 26: 515-536.

KALINOWSKI, S. T., M. L. TAPER, and T. C. MARSHALL. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology* 16: 1099-1106.

KHAN, T. N. 1976. Papua New Guinea: A centre of genetic diversity in winged bean (*Psophocarpus tetragonologus* (L.) DC.). *Euphytica* 25: 693-706.

LANAVE, C., G. PREPARATA, C. SACCONE, and G. SERIO. 1984. A new method for calculating evolutionary substitution rates. *Journal of Molecular Evolution* 20: 86-93.

LEFORT, V., R. DESPER, and O. GASCUEL. 2015. FastME 2.0: a comprehensive, accurate and fast distance-based phylogeny inference program. *Molecular Biology and Evolution* 32: 2798-2800.

LI, H., B. HANDSAKER, A. WYSOKER, T. FENNELL, J. RUAN, N. HOMER, G. MARTH, et al. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25: 2078-2079.

MANTEL, N. 1967. The detection of disease clustering - a generalised regression approach. *Cancer Research* 27: 209-220.

MAXTED, N. 1989. A Phenetic Investigation of the *Psophocarpus palustris*: *P. scandens* Complex (Leguminosae, Phaseoleae, Phaseolinae). *Kew Bulletin* 44: 731-742.

MAXTED, N. 1990. A phenetic investigation of *Psophocarpus* Neck. ex DC. (Leguminosae - Phaseoleae). *Botanical Journal of the Linnean Society* 102: 103-122.

MEIRMANS, P. G. 2015. Seven common mistakes in population genetics and how to avoid them. *Molecular Ecology* 24: 3223-3231.

MICKELBART, M. V., P. M. HASEGAWA, and J. BAILEY-SERRES. 2015. Genetic mechanisms of abiotic stress tolerance that translate to crop yield stability. *Nature Reviews: Genetics* 16: 237-251.

MOHANTY, C., S. VERMA, V. SINGH, S. KHAN, P. GAUR, P. GUPTA, M. NIZAR, et al. 2013. Characterization of winged bean (*Psophocarpus tetragonolobus* (L.) DC.) based on molecular, chemical and physiological parameters. *American Journal of Molecular Biology* 3: 187-197.

NAS. 1975a. The winged bean. A high protein crop for the tropics. National Academy of Sciences, Washington, D.C., USA. .

NAS. 1975b. Underexploited tropical plants with promising economic value. National Academy of Sciences, Washington, D.C., USA. .

NUR FATIHAH, H. N., N. MAXTED, and L. RICO ARCE. 2012. Cladistic analysis of *Psophocarpus* Neck. ex DC. (Leguminosae, Papilionoideae) based on morphological characters. *South African Journal of Botany* 83: 78-88.

PEAKALL, R., and P. E. SMOUSE. 2006. GENALEX 6: Genetic Analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6: 288-295.

PEOPLES, M. B., D. F. HERRIDGE, and J. K. LADHA. 1995. Biological nitrogen fixation: An efficient source of nitrogen for sustainable agricultural production? *Plant and Soil* 174: 3-28

PRAKASH, D., P. N. MISRA, and P. S. MISRA. 1987. Amino acid profile of winged bean (*Psophocarpus tetragonolobus* (L.) DC): a rich source of vegetable protein. *Plant Foods for Human Nutrition* 37: 261-264.

RAHMAN, M. M., A. M. ISLAM, S. M. AZIRUN, and A. N. BOYCE. 2014. Tropical Legume Crop Rotation and Nitrogen Fertilizer Effects on Agronomic and Nitrogen Efficiency of Rice. *The Scientific World Journal* 2014: 490841.

RASBAND, W. S. 1997-2013. ImageJ (imagej.nih.gov/ij/).

RONQUIST, F., M. TESLENKO, P. VAN DER MARK, D. AYRES, A. DARLING, S. HOHNA, B. LARGET, et al. 2011. MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* 61: 539-542.

ROZEN, S., and H. J. SKALETSKY. 2000. Primer3 on the WWW for general users and for biologist programmers. *In* S. Krawetz AND S. Misener [eds.], Bioinformatics Methods and Protocols: Methods in Molecular Biology., 365-386. Humana Press, Totowa, NJ.

SCHMIDHUBER, J., and F. N. TUBIELLO. 2007. Global food security under climate change. *Proceedings of the National Academy of Sciences of the United States of America* 104: 19703-19708.

SCHMUTZ, J., P. E. MCCLEAN, S. MAMIDI, G. A. WU, S. B. CANNON, J. GRIMWOOD, J. JENKINS, et al. 2014. A reference genome for common bean and genome-wide analysis of dual domestications. *Nature Genetics* 46: 707-713.

SCHUELKE, M. 2000. An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology* 18: 233-234.

SIDDIQUE, K. H. M., C. JOHANSEN, N. C. TURNER, M. H. JEUFFROY, A. HASHEM, D. SAKAR, Y. T. GAN, et al. 2012. Innovations in agronomy for food legumes. A review. *Agronomy for Sustainable Development* 32: 45-64.

SIEVERS, F., and D. G. HIGGINS. 2014. Clustal Omega, Accurate Alignment of Very Large Numbers of Sequences. *In* D. J. Russell [ed.], Multiple Sequence Alignment Methods, vol. 1079, Methods in Molecular Biology, 105-116.

SMARTT, J. 1980. Some observations on the origin and evolution of the winged bean (*Psophocarpus tetragonolobus*). *Euphytica* 29: 121-123.

TAMURA, K., and M. NEI. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution* 10: 512-526.

TECHNELYSIUM PTY LTD. 1998-2001. Chromas 2.12 www.technelysium.com.au.

THARANATHAN, R. N., and S. MAHADEVAMMA. 2003. Grain legumes - a boon to human nutrition. *Trends in Food Science & Technology* 14: 507-518.

VARSHNEY, R. K., R. K. SAXENA, H. D. UPADHYAYA, A. W. KHAN, Y. YU, C. KIM, A. RATHORE, et al. 2017. Whole-genome resequencing of 292 pigeonpea accessions identifies genomic regions associated with domestication and agronomic traits. *Nature Genetics* 49: 1082-1088.

VATANPARAST, M., P. SHETTY, R. CHOPRA, J. J. DOYLE, N. SATHYANARAYANA, and A. N. EGAN. 2016. Transcriptome sequencing and marker development in winged bean (*Psophocarpus tetragonolobus*; Leguminosae) 6: 29070.

VERDCOURT, B., and P. HALLIDAY. 1978. A revision of *Psophocarpus* (Leguminosae: Papilionoideae: Phaseoleae). *Kew Bulletin* 33: 191-227.

WHITE, T., T. BRUNS, S. LEE, and J. TAYLOR. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *In* I. MA, G. DH, S. JJ, AND W. TJ [eds.], PCR protocols: a guide to methods and applications., 315–322. United States: Academic Press., United States: Academic Press.

WILLS, D. M., and J. M. BURKE. 2006. Chloroplast DNA variation confirms a single origin of domesticated sunflower (*Helianthus annuus* L.). *Journal of Heredity* 97: 403-408.

WONG, Q. N., A. S. TANZI, W. K. HO, S. MALLA, M. BLYTHE, A. KARUNARATNE, F. MASSAWE, et al. 2017. Development of Gene-Based SSR Markers in Winged Bean (*Psophocarpus tetragonolobus* (L.) DC.) for Diversity Assessment. *Genes* 8.

ZEVEN, A. C., and J. M. J. DE WET. 1982. Dictionary of cultivated plants and their regions of diversity. 2nd ed. Centre for Agricultural Publishing and Documentation, Wageningen, Netherlands.

DATA ACCESSIBILITY

cpDNA and ITS sequences have been deposited in GenBank (http://www.ncbi.nlm.nih.gov/genbank) with accessions numbers MF980721 - MF980787 and MG969796 - MG969815, respectively.
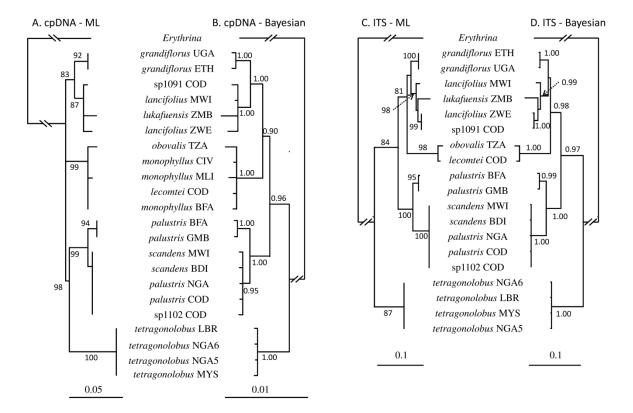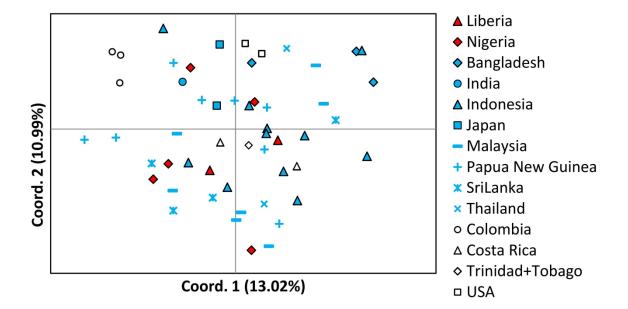
Figure 1

Figure 2

Legend:
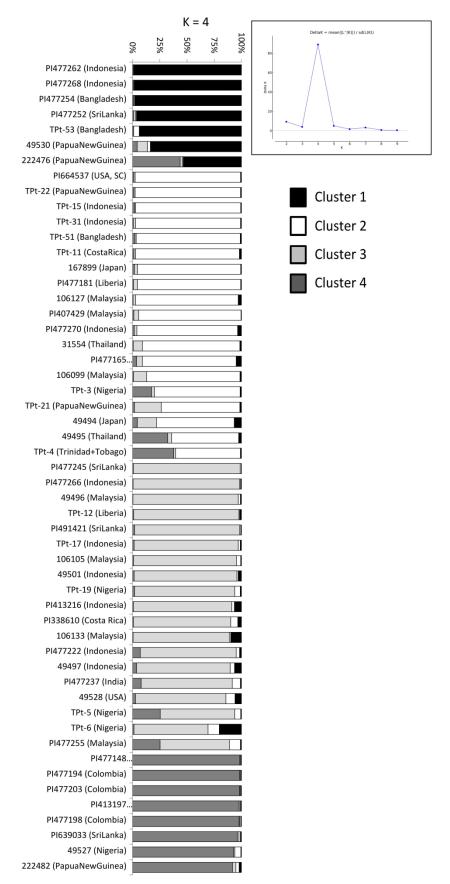- ▲ Liberia
- ◆ Nigeria
- ◆ Bangladesh
- ● India
- ▲ Indonesia
- ■ Japan
- ▬ Malaysia
- ✚ Papua New Guinea
- ✖ SriLanka
- ✕ Thailand
- ○ Colombia
- △ Costa Rica
- ◇ Trinidad+Tobago
- □ USA

Coord. 1 (13.02%)

Coord. 2 (10.99%)

Figure 3

Figure 4