Phylogenetic relationships and genetic diversity of the USDA *Vigna* germplasm collection revealed by gene-derived markers and sequencing

MING LI WANG*, NOELLE A. BARKLEY, GRAVES A. GILLASPIE
AND GARY A. PEDERSON
USDA-ARS, Plant Genetic Resources Conservation Unit, 1109 Experiment Street, Griffin, GA 30223, USA

(Received 21 May 2008 and in revised form 19 August 2008)

Summary

Phylogenetic relationships in the USDA *Vigna* germplasm collection are somewhat unclear and their genetic diversity has not been measured empirically. To reveal interspecific phylogenetic relationships and assess their genetic diversity, 48 accessions representing 12 *Vigna* species were selected, and 30 gene-derived markers from legumes were employed. Some high-quality amplicons were sequenced. Indels (insertion/deletions) were discovered from the sequence alignments that were specific identifiers for some *Vigna* species. With regard to revealing polymorphisms, intron-spanning markers were more effective than exon-derived markers. These gene-derived markers were more successful in revealing interspecific polymorphisms than intraspecific polymorphisms at both the DNA fragment and sequence levels. Two different dendrograms were generated from DNA fragment data and sequence data, respectively. The results from these two dendrograms supported each other and showed similar phylogenetic relationships among the *Vigna* species investigated. The accessions clustered into four main groups and 13 subgroups. Each subgroup represents a subgenus or a species. Phylogenetic analysis revealed that an accession might be misclassified in our collection. The putative misclassified accession was further supported by seed morphology. Limited intraspecific genetic diversity was revealed by these gene-derived markers and/or sequences. The USDA *Vigna* germplasm collection currently consists of multiple species with many accessions further classified into specific subspecies, but very few subspecies of the total subspecies available exist within the collection. Based on our results, more attention should be paid to the subspecies, wild forms and/or botanical varieties for future curation in order to expand the genetic diversity of *Vigna* germplasm in the USDA collection.

1. Introduction

Considerable confusion in the synonymy and classification of various *Vigna* species exists in the literature (Verdcourt, 1970; Fery, 1980). The genus *Vigna* comprises seven subgenera and more than 80 species. Some of the species adapt well to a wide range of environmental conditions (such as poor soils and drought) and have been domesticated to cultivated species (Faris, 1965; Verdcourt, 1970; Santalla et al., 1998). Cultivated *Vigna* species are an important protein source in countries where people have limited access to food rich in protein (Singh, 2005). The main cultivated species worldwide include the following: five Asian beans: moth bean (*Vigna aconitifolia* (Jacq.) Marechal), azuki bean (*Vigna angularis* (Willd.) Ohwi and Ohashi), black gram (*Vigna mungo* L.), mung bean (*Vigna radiata* L.), rice bean (*Vigna umbellata* Thunb.); two African beans: bambara groundnut (*Vigna subterranea* L.) and cowpea (*Vigna unguiculata* (L.) Walp); and American *Vigna* beans (Jaaska & Jaaska, 1990; Jaaska, 1999, 2001).

Isoenzymes as biochemical markers have been used to assess genetic diversity and reveal phylogenetic relationships among *Vigna* species (Jaaska & Jaaska, 1990; Pasquet, 1999, 2000). The phylogenetic relationships and genetic diversity among and within *Vigna* species were first assessed by restriction

* Corresponding author. Tel: +1 (770) 229-3342. Fax: +1 (770) 229-3323. e-mail: mingli.wang@ars.usda.gov

https://doi.org/10.1017/S0016672308009889 Published online by Cambridge University Press
Table 1. Selected accessions from Vigna species

| No. | PI      | Species                      | Identifier | Seed colour | Seed weight | Origin/collected location |
|-----|---------|------------------------------|------------|-------------|-------------|---------------------------|
| 1   | 633451  | *Phaseolus vulgaris* L.      | BAT 93     | Tan (str)   | 19.2        | USA                       |
| 2   | 164419  | *Vigna aconitifolia* Jacq.  | 8712       | Brown (sol) | 2.59        | India                     |
| 3   | 165479  | *V. aconitifolia* Jacq.     | 8954       | Brown (sol) | 1.87        | India                     |
| 4   | 372355  | *V. aconitifolia* Jacq.     | Kitna      | Brown (m)   | 2.22        | Yemen                     |
| 5   | 312898  | *V. adenantha* G. Mey.      | n/a        | Brown       | 3.28        | Argentina                 |
| 6   | 319448  | *V. adenantha* G. Mey.      | 22156      | Black       | 2.1         | Mexico                    |
| 7   | 93815   | *V. angularis* Willd.       | n/a        | Tan (sol)   | 9.59        | China                     |
| 8   | 157625  | *V. angularis* Willd.       | 26         | Red (sol)   | 8.78        | Korea                     |
| 9   | 360707  | *V. angularis* Willd.       | Treasure small bean | Red (sol) | 8.17        | China                     |
| 10  | 416742  | *V. angularis* Willd.       | 65-078-02570A | Red (sol) | 8.16        | Japan                     |
| 11  | 527686  | *V. angularis* var. *nipponensis* Owhi & H. Ohashi | n/a | Bird bean | 2.37 | Korea                     |
| 12  | 146800  | *Vigna caracalla* L.        | No. 842    | Tan (sp)    | 1.93        | South Africa              |
| 13  | 322588  | *V. caracalla* L.           | IRI 1289   | Brown (m)   | 5.22        | Brazil                    |
| 14  | 310294  | *Vigna longifolia* Verdc.   | 298        | Grey (m)    | 1.38        | Brazil                    |
| 15  | 292866  | *Vigna luteola* Jacq.       | C36-323    | Brown (sp)  | 2.89        | India                     |
| 16  | 406329  | *V. luteola* Jacq.          | TVnu 2     | Brown (sp)  | 2.79        | Tanzania                  |
| 17  | 164316  | *Vigna mungo* L. var. *mungo* | OLANDU     | Black (sp)  | 4           | India                     |
| 18  | 208462  | *V. mungo* L. var. *mungo*  | n/a        | Black (sp)  | 4.32        | Nepal                     |
| 19  | 218104  | *V. mungo* L. var. *mungo*  | MASH       | Black (sp)  | 5.61        | Pakistan                  |
| 20  | 305073  | *V. mungo* L. var. *mungo*  | 5          | Black (sp)  | 4.41        | Thailand                  |
| 21  | 292872  | *Vigna oblongifolia* A. Rich. | C36-307 | Brown (sol) | 3.25        | Nigeria                   |
| 22  | 181585  | *V. oblongifolia* A. Rich. var. *oblongifolia* | Wilmsii  | Black (m)   | 2.57        | South Africa              |
| 23  | 292688  | *V. oblongifolia* A. Rich. var. *parviflora* | C36-328 | Black (m)   | 2.27        | India                     |
| 24  | 406358  | *V. oblongifolia* A. Rich. var. *parviflora* | TVnu 36  | Brown (m)   | 0.79        | Tanzania                  |
| 25  | 164301  | *Vigna radiata* L. var. *radiata* | PACHAPAYARU | Green (sol) | 3.94        | India                     |
| 26  | 171435  | *V. radiata* L. var. *radiata* | 55 | Green (sol) | 4.36        | China                     |
| 27  | 381351  | *V. radiata* L. var. *radiata* | IVau 65001 | Yellow green (sol) | 5.22 | Nigeria                   |
| 28  | 427064  | *V. radiata* L. var. *radiata* | K-970  | Green (sol) | 2.92        | Pakistan                  |
| 29  | 240867  | *Vigna subterranea* L.       | Light reddish brown | Reddish brown (sol) | 36.2 | Uganda                    |
| 30  | 245951  | *V. subterranea* L.          | Bereke     | Reddish brown (sol) | 56.9 | Zimbabwe                  |
| 31  | 378867  | *V. subterranea* L.          | Ditlo      | Red (sol)   | 32          | Nigeria                   |
| 32  | 173933  | *Vigna umbellata* Thunb.     | Oorat      | Red (sol)   | 4.57        | India                     |
| 33  | 208460  | *V. umbellata* Thunb.        | MASH       | Tan (sol)   | 5.34        | Nepal                     |
| 34  | 247689  | *V. umbellata* Thunb.        | Rouge      | Red (sol)   | 5.62        | Zaire                     |
| 35  | 275636  | *V. umbellata* Thunb.        | 1422       | Tan (sol)   | 4.87        | India                     |
| 36  | 291384  | *Vigna unguiculata* L. ssp. *cylindrica* | 276 | Black (spec) | 7.72 | China                     |
| 37  | 292883  | *V. unguiculata* L. ssp. *dekindtiana* | C36-258  | Black (spec) | 1.99 | Nigeria                   |
fragment length polymorphism (RFLP) DNA markers (Fatokun et al., 1993), and then by random amplified polymorphic DNA (RAPD) (Kaga et al., 1996; Santalla et al., 1998; Lakhanpaul et al., 2000; Mimura et al., 2000; Xu et al., 2000; Amadou et al., 2001; Ba et al., 2004; Diouf & Hilu, 2005) and amplified fragment length polymorphism (AFLP) (Tomooka et al., 2002; Zong et al., 2003; Seehalak et al., 2006; Yoon et al., 2007) DNA markers. As more DNA sequence information is now available, internal transcribed spacer (ITS) sequences (Doi et al., 2002; Goel et al., 2002), simple sequence repeat (SSR) markers (Li et al., 2001; Kumar et al., 2002; Wang et al., 2004; Gillaspie et al., 2005) and DNA amplification fingerprinting (DAF) (Simon et al., 2007) were used for assessing the phylogenetic relationships and genetic diversity. Although various DNA markers have been developed from different Vigna species, there are neither a common nor a sufficient set of robust DNA markers available for evaluation of germplasm applicable to all Vigna species. Recently, gene-derived markers were developed across 15 legumes (Choi et al., 2006). Since these markers are derived from putative genes, they may be a good source to reveal phylogenetic relationships and assess genetic diversity among and within Vigna species.

The US germplasm resource for Vigna species is maintained at the USDA-ARS Plant Genetic Resources Conservation Unit (PGRCU) located at Griffin, GA, USA. The phylogenetic relationships in the USDA Vigna germplasm collection are unclear and their genetic diversity is unknown. Revealing the phylogenetic relationships and assessing genetic diversity will help develop strategies for better organization and management of existing as well as further acquisitions of Vigna germplasm. The objectives of the present study were to: (i) reveal phylogenetic relationships and assess genetic diversity of Vigna species in the USDA collection using gene-derived DNA markers, (ii) sequence amplicons generated from gene-derived primers to identify polymorphisms and (iii) evaluate the effectiveness of gene-derived markers in revealing the phylogenetic relationships and assessing genetic diversity among and within Vigna species.

### 2. Materials and methods

(i) **Plant materials and DNA extraction**

Taxonomic classifications of accessions used in the present study are based on the Germplasm Resources Information Network (GRIN; found at http://www.ars-grin.gov/npgs/index.html). Forty-eight accessions from several Vigna species (Table 1) were used in this experiment and all accessions were diploids containing

| Accession | Taxonomic classification | Location | Color      |
|-----------|--------------------------|----------|------------|
| 38        | V. unguiculata L. ssp. pubescens | Tanzania | Yellow & green (spec) |
| 39        | V. unguiculata L. ssp. pubescens | Tanzania | Tan with red eye (sol) |
| 40        | V. unguiculata L. ssp. pubescens | Tanzania | Brown (sol) |
| 41        | V. unguiculata L. ssp. pubescens | Tanzania | Brown with black eye |
| 42        | V. unguiculata L. ssp. pubescens | Tanzania | Brown with white eye |

Letters in parentheses: m, mottle; sol, solid; spec, speckled; str, streaked.
11 pairs of chromosomes 

\((2n=2x=22)\) (Singh, 2005). Among them, 12 accessions were from \(V.\) unguiculata L. (cowpea, representing six subspecies); five accessions from \(V.\) angularis Willd. (azuki bean); four accessions each from \(V.\) radiata L. (mung bean), \(V.\) mungo L. (black gram), \(V.\) umbellata Thunb. (rice bean) and \(Vigna\) oblongifolia A. Rich.; three accessions each from \(V.\) aconitifolia Jacq. (moth bean) and \(V.\) subterranea L. (bambara groundnut); two accessions each from \(Vigna\) adenantha G. Mey., \(Vigna\) caracalla L. and \(V.\) luteola Jacq.; and one accession each from \(Vigna\) longifolia Verdc. and \(Vigna\) vexillata L. One accession from \(Phaseolus\) vulgaris L. (common bean), which is closely related to the \(Vigna\) genus (Verdcourt, 1970), was also included as an outgroup in the present study. Leaf tissue samples were collected from plants grown in a greenhouse at Griffin, GA, USA. DNA was extracted from leaf tissue using an E.Z.N.A. Plant DNA Miniprep kit from Omega Bio-Tek (Doraville, GA, USA). The DNA was then diluted to 10 ng/\(\mu\)l and later used as a template for PCR.

(ii) **PCR and PCR product separation**

Thirty-two consensus sequences were imported into ClustalX (Thompson et al., 1997) and aligned with published information (Choi et al., 2006) and are listed in Table 2. All PCR reactions, programmes and product separations on agarose gels were performed by following the method described previously by Wang et al. (2006).

(iii) **Allele sequencing**

Before sequencing, PCR products were checked on a 3% agarose gel to verify that only a single band was produced from each sample. The PCR product was treated with 1 \(\mu\)l of exonuclease I (10 units/\(\mu\)l) and 1 \(\mu\)l of shrimp alkaline phosphatase (1 unit/\(\mu\)l) (GE Healthcare, Piscataway, NJ, USA) for every 12 \(\mu\)l of PCR reaction to digest single-stranded DNA and cleave the 5'-phosphate, respectively. The PCR product was also cleaned with a Qiagen PCR cleanup kit (Qiagen, Valencia, CA, USA) to remove excess nucleotides, primers, enzymes and other impurities. Then, 1 \(\mu\)l of the cleaned product was run on an agarose gel with a quantitative marker (Invitrogen, Carlsbad, CA, USA) to determine product concentration and thus prepare the sample for sequencing. Sequencing reactions were prepared following the instructions of the DTCS Quick Start sequencing kit (Beckman Coulter, Fullerton, CA, USA). The sample was sequenced bi-directionally and pUC18 was also sequenced as a positive control. Each sample was sequenced twice to verify the fidelity of the sequenced bases. Samples were injected and sequenced on a Beckman CEQ 8000 by using the LFR-1 method. The sequence module of the software package CEQ 8000 Genetic Analysis System version 8.0.52 from Beckman was used to call the bases after the sequencing was performed. The forward and reverse strands were edited and aligned using AlignIR version 2.0 (LICOR, Lincoln, NE, USA).

(iv) **Morphology comparison**

Seeds were harvested from plants grown in a greenhouse at Griffin, GA, USA. Seeds mainly representing the accession were weighed. Seed-coat colour and size were scanned and recorded with a Hewlett-Packard Scanjet 7400C. In order to confirm the classification, more accessions from the same species or subspecies were requested from the Griffin seed store and compared regarding their seed morphology.

(v) **Data analysis**

Strong clear bands on the gel images were scored as either present (1) or absent (0) for DNA fragment analysis. The data were entered into a binary matrix for analysis. A distance matrix was created between all pairwise combinations by using the proportion of shared allele algorithm in the program MICROSAT v.1.5 (Minch et al., 1997). One hundred bootstrap replicates were generated and a neighbour-joining tree and a consensus tree were constructed using PHYLIP v.3.6 (Felsenstein, 2005). The trees were then viewed and printed using TreeView (Page, 1996).

Thirty-two consensus sequences were imported into ClustalX (Thompson et al., 1997) and aligned for DNA sequencing data analysis. Low gap penalties (gap penalty = 10 and gap extension = 0.1) were applied using the slow and accurate pairwise alignment of ClustalX. The resulting alignment was evaluated for maximum pairwise identity. In general, high gap penalties are suitable for intraspecific data, whereas low gap penalties are suitable for interspecific data. The sequence data were imported into Phylib v.3.6 (Felsenstein, 2005) and maximum likelihood (DNAML) was employed. The program SEQBOOT from the PHYLIP package was used to perform bootstrapping with 100 replicates to test the stability of the clades. CONSENSE was used to create consensus trees from bootstrap replicates.

3. **Results and discussion**

(i) **Gene-derived markers, genetic diversity and phylogenetic relationships**

All 30 gene-derived primers amplified DNA fragment(s) from the accessions tested. Employed as DNA markers, 26 (86.7%) were polymorphic and
| Name      | Origin | Gene Type | F-primer | R-primer | Size (bp) |
|-----------|--------|-----------|----------|----------|-----------|
| BV164524  | V. radiata | 6DSC IS  | CTCACGAAAGCCATTGGAGGCACAGCAACTTG | TGGCAATGAGCATCAGCATTCTTTTGA | 263–311   |
| BV164858  | V. radiata | AAT IS   | TGCTTTACGATGCACCAAAGGCTCCTTAG | TCCGACATTAGATCATAAGTGG | 432–596  |
| BV164971  | V. radiata | ACL IS   | GTAAGGTCACGACTGTATTAATGACCAAC | TCTCAAACTCTCTCAATTGGGACAG | 447–509  |
| BV164954  | V. radiata | BIPA IS  | GAGGAGTCTCACAAGAGGATTTTC | GCTTTTCTATGTTGACATAGTTTCA | 363–405  |
| BV165019  | V. radiata | CPOX2 IS | GATAATGCTGCCAATATTGCACCTTCA | GCTCAGAGCCCACCTCTTTAG | 253–286  |
| BV165028  | V. radiata | CTP ED   | GTTTGACCCGGACACCAACTCTATAGAGTA | TGGCAATGAGCATCAGCATTCTTTTGA | 303–316  |
| BV165051  | V. radiata | CYSK IS  | GGAATGGCACTGTTGACATAGTTTCA | AATGGAGGACACTCTGTTGACATAGTTTCA | 310–423  |
| BV165086  | V. radiata | DK326 IS | GAAACCTGTCGACTCTTCA | TTAGCATGATTAAACCCATATG | 539–631  |
| BV165128  | V. radiata | ENOL IS  | CATCAAGGCCGAGGACACTTCA | GCTCAGAGCCCACCTCTTTAG | 1087–1104|
| BV165137  | V. radiata | EST763 STS | CACTCTAAAAAGGGCCAGAAGGTTTAGC | TTATGACCAATATCCTTTTCA | 348–369  |
| BV165163  | V. radiata | FENR IS  | ATGCTATGCCCCAAAGATCAAATGC | TCTCAGCAACAGATGGCTTAGAAGT | 479–508  |
| BV165196  | V. radiata | GLNA IS  | GAAATTGCGTGTGCTTCACATG | TGGTAGTTGTCTGCTACATGGAAG | 498–502  |
| BV165202  | V. radiata | GLP ED   | GACTCAACAGCAGTGCATATGCTTTCTCCT | TTATGACCAATATCCTTTTCA | 278–356  |
| BV165337  | V. radiata | MU107 STS | ATATTGACGATAATGGCCATTTAGCT | AGACTCTGCTGCTCAGTGTTTCAAT | 491–518  |
| BV165346  | V. radiata | MU141 STS | TTGATCAAGGCAAGAATTTAAT | GCCCTCCACAAAGTAAAGTTC | 636–652  |
| BV165392  | V. radiata | MSU72 IS  | ATGAGGTCGAGCGTGGGAGATTTTTG | AGCATGCTGCTGCTCAGTGTTTCAAT | 392–415  |
| BV165438  | V. radiata | Ntlim1 IS | GAAATGGTGTCAGGAGCAGTGGGAGA | GTGCTCCTGGTCCCAACCCACCACTTTCA | 382–450  |
| BV165476  | V. radiata | PDC ED   | GCTGTGAGGCCCCATCAGTAGCAGC | CCAATGCAAGGCTCCAGAGCCTCACACACA | 341–365  |
| BV165482  | V. radiata | PEPCASE ED | TGGCAAGCATATATCACTTCGAG | GAATACTACACAAAGCCGGTTCCTTTCG | 323–350  |
| BV165526  | V. radiata | PNDKN IS  | GGCGGAAACAGACTCTCACATGATC | CCAAGCTGCGATGAGCAGGTGTTTGT | 553–694  |
| BV165531  | V. radiata | PP IS    | GTGCACAGCAGTACGGAACCAACCA | CTGCTTCCTGCTGCTGCTTTCGAG | 450–636  |
| BV165543  | V. radiata | PPH ED   | ATGGTCTTTGGCCACTGGAAGTGCAGAAGAAG | GCCAGGAAATGACGATGATAAGAAA | 287–299  |
| BV165627  | V. radiata | RBPCO ED | CCCATCTGATGACTGGTGAG | TTAGAAGAATCTGACCGGACATGATG | 273–280  |
| BV165651  | V. radiata | RNAH IS  | GCTTGCAACCTGACCAAGTATCAAG | TTAGCCTCAGAAAGATCAGTAC | 670–727  |
| BV165660  | V. radiata | RNAR IS  | GTTTGCAATGATGATGAGTGGTGAAGA | GGTAGGGCAATGATGGGGACACCA | 312–465  |
| BV165710  | V. radiata | SHMT IS  | ACCCAACCTCGATGACTGCACTTAC | TTAGCCTCGATGAGCAGCACCTC | 570–575  |
| BV165732  | V. radiata | SUSY IS  | TCGCGAATGACACACACAACTTTCCA | GTGACGCTGCTGCTGCTGCTTTCG | 481–488  |
| BV165766  | V. radiata | TGFRIP ED | ATTCTCTGGAGGAGGAGGACGAAGC | CAAGCTTATGACCAACACAGAATCAC | 323–353  |
| BV165781  | V. radiata | tRNLS IS | GCTGTCGAGCTTGTTGAGGAGAAG | GCAATCCCTCCACCTCAGTAAGT | 293–324  |
| BV165813  | V. radiata | TUP IS   | GAATGGGATGCTATGGGAAGT | TGGATCAGTGCCACACCATTTTAT | 745–754  |

a Markers were monomorphic for *Vigna* germplasm.

IS, intron-spanning marker; ED, exon-derived marker; STS, sequence-tagged site-derived marker.
4 (13.3%) were monomorphic. There were no polymorphic markers identified within an accession. All of the markers derived from the intron-spanning (IS) regions were polymorphic, whereas only four of the seven markers from the exon-derived (ED) regions were polymorphic (Table 2). For example, marker BV165019 derived from the intron-spanning region of the gene cationic peroxidase 2 (COPX2) revealed interspecific polymorphisms but not intraspecific polymorphisms, whereas the marker BV165028 from the exon-derived region of carboxyl-terminal peptidase (CTP) was monomorphic for all the accessions examined (Fig. 1). The results from gene-derived markers demonstrated that at the DNA fragment level, (i) intron-spanning markers were more effective in revealing polymorphisms than exon-derived markers and (ii) gene-derived markers were more effective in revealing polymorphisms among species than within species (Fig. 1 and Table 2). However, future studies should include more accessions from a single species to evaluate intraspecific polymorphism.

A total of 134 polymorphic bands (DNA fragments) were observed with an average of approximately five bands per marker, which were subsequently used for phylogenetic analysis. A dendrogram generated from gene-derived DNA marker data is shown in Fig. 2 and the accessions examined clustered into four main groups (American bean group, African bean group, Asian bean group and intermediate group). One accession (Plant Introduction number (PI) 633451) from *P. vulgaris*, which was used as the outgroup, was different from most of the accessions, but most similar (genetic distance of 0.027) to one accession (PI 146800) from *V. caracalla*. This was not surprising because *P. vulgaris* was taxonomically separated from the *Vigna* genus only since 1970 (Verdcourt, 1970).

Group I included three accessions: two accessions (PI 312898 and PI 319448) from *V. adenantha*, and one accession (PI 322588) from *V. caracalla*. Species of *V. adenantha* and *V. caracalla* belong to the same subgenus *Sigmoidotropis* (Piper) and were also classified as American beans using biochemical markers (Jaaska, 2001). Interestingly, PI 312898 was closely related (distance = 0.013) to PI 322588, which was from a different species (*V. caracalla*, supported by a 59% bootstrap value). Species within group I was closely related to *P. vulgaris*. This result was similar to previous phylogenetic analyses of the Phaseolus–Vigna complex (Jaaska, 2001; Goel et al., 2002).

Group II contained only one accession (PI 310294), which belongs to *V. longifolia*. This group is designated as an intermediate group.

Group III contained 20 accessions (from PI 164419 to PI 171435 in Fig. 2) covering five species or five Asian beans, representing five subgroups. Accessions from each species formed a cluster. Three accessions classified as *V. aconitifolia* (moth bean) were examined. Two accessions (PI 165479 and PI 372355) clustered together with little genetic variation (genetic distance = 0.001) and were supported by a 66% bootstrap value. These accessions were different
(genetic distance = 0.027) from accession PI 164419. Therefore, some genetic diversity was detected within moth bean. Four accessions of *V. mungo* var. *mungo* (black gram, from PI 208462 to PI 305073) were examined and they formed a distinct cluster. Five accessions of *V. angularis* (azuki bean, from PI 527686 to PI 360707) were examined and very little genetic difference was detected among these accessions as demonstrated by the short branches and small genetic distance values among these accessions ranging from 0.0008 to 0.0018. Furthermore, strong support for monophyly was apparent with the support of a 73% bootstrap value. The accession, PI 527686 was previously determined to be the wild form (or progenitor) of cultivated azuki bean (Mimura et al., 2000) and the phylogeny supports this notion. Four accessions were examined within *V. umbellata* (rice beans). Accessions (PI 247689 and PI 275636) and (PI 173933 and PI 208460) formed two small clusters that were supported by 71 and 62% bootstrap values, respectively.

Fig. 2. Neighbour-joining tree of *Vigna* species. Bootstrapping was performed with 100 replicates and values greater than 40% were placed on the branches.
Four accessions of *V. radiata* (mung bean, PI 171435, PI 164301, PI 381351 and PI 427064) were examined and formed a small cluster. Very limited genetic diversity was detected in mung bean (genetic distance ranging from 0.0009 to 0.003). This result was consistent with the genetic diversity revealed by SSR markers in mung bean (Gillaspie et al., unpublished results) and EcoTILLING (Barkley et al., 2008), which may suggest a narrow genetic base for *V. radiata*. Low levels of genetic diversity were also revealed among 32 Indian mung bean cultivars (Lakhanpal et al., 2000). Based on branch nodes from our results, mung bean (*V. radiata*) was genetically very different from the other four beans.

**Group IV** contained 22 accessions covering five species (including two African beans: bambara groundnut and cowpea), representing five subgroups. Accessions from each species formed a cluster. The species within group IV originated in Africa and therefore this group was named the African bean group. Two accessions (PI 292866 and PI 406329) from *V. luteola* clustered with a genetic distance of 0.008. Four accessions were examined within *V. oblongifolia* and they formed a small cluster with the support of a 72% bootstrap value. However, PI 406358 was genetically similar to three other accessions with genetic distances ranging from 0.011 to 0.014 (PI 292868, PI 181585 and PI 292872; Fig. 2). Based on phylogenetic analysis, it seems that the *V. luteola* species was closely related (distance = 0.22–0.31) to *V. oblongifolia* with the support of a 60% bootstrap value. One accession (PI 406390) from *V. vexillata* clustered closely to *V. unguiculata* and *V. subterranea*, which was consistent with the early RFLP analysis (Fatokun et al., 1993) and morphological observations (Bisht et al., 2005). Three accessions (PI 378867, PI 240867 and PI 245951) were examined within *V. subterranea* (bambara groundnut) and they formed a small cluster with the support of a 67% bootstrap value. However, PI 240867 and PI 245951 were closely related genetically (distance = 0.005) with the support of a 69% bootstrap value and were distinct from PI 378867. In a previous study, genetic classification of *V. subterranea* accessions was related to geographic origin and accessions collected from Nigeria were very different from accessions collected from Zimbabwe (Amadou et al., 2001). In the present study, PI 378867 collected from Nigeria was also very different from PI 245951 (genetic distance of 0.01) collected from Zimbabwe. Twelve accessions (PI 632904–PI 582578) from cowpea (*V. unguiculata*) clustered together, forming a subgroup. There were six subspecies examined within *V. unguiculata* (Table 1). The accession (PI 632910) from subspecies *p. sesquipedalis* may be closely related (distance = 0.005–0.007) to the accessions (PI 632903 and PI 632904) from subspecies *stenophylla* and formed a small cluster.

Four accessions (PI 582578, PI 582469, PI 582470 and PI 612607) from subspecies *unguiculata*, three accessions (PI 610582, PI 419163 and PI 215659) from subspecies *sesquipedalis*, one accession (PI 292883) from subspecies *dekindiana* and one accession (PI 291384) from subspecies *cylindrica* formed a small cluster with the support of a 68% bootstrap value. Obviously, some genetic diversity was detected among these subspecies but compared with other species examined within group IV; the genetic diversity detected was minimal, as demonstrated by short branch lengths and low genetic distance among accessions ranging from 0.005 to 0.013 within *V. unguiculata*. This result was consistent with results from other studies (Li et al., 2001; Diouf & Hilu, 2005) in which a narrow genetic base was also found in cowpea breeding lines and local varieties from Senegal. A single domestication event between wild and cultivated cowpea may be the explanation for the narrow genetic base within cowpea (Coulibaly et al., 2002; Ba et al., 2004). Another possible explanation is that the type of DNA markers employed may also affect the level of polymorphism revealed. To reveal genetic diversity of cowpea (*V. unguiculata* var. *unguiculata*), SSR, RAPD and gene-derived markers were used in two previous studies as well as in the present study, respectively. When 26 DAF primers were employed, 54 cowpea accessions (*V. unguiculata* var. *unguiculata*) were classified into separate groups (Simon et al., 2007). DAF may be a highly efficient system for the generation of polymorphic DNA markers for revealing cowpea genetic diversity.

(ii) DNA sequencing, genetic diversity and phylogenetic relationships

To detect polymorphism at the DNA sequence level, 32 amplicons generated from the marker BV165019 were sequenced and the sequence alignment is shown in Fig. 3. Possible sequence errors were identified at a primer site when the consensus sequences were constructed from forward and reverse reads. The first few base pairs of the sequences with possible errors were removed when the phylogenetic tree was generated from sequence data. The phylogenetic tree generated from sequence data is shown in Fig. 4. Comparing the sequence alignment of the common bean (*P. vulgaris*, PI 633451) with the remaining sequences derived from species within the genus *Vigna*, several small deletions or insertions (indels) were identified. This implies that indels may play an important role in differentiation and speciation (Fig. 3). According to the phylogenetic analysis from DNA bands (size of the DNA fragments), PI 322588 from *V. caracalla* formed a cluster with accessions (PI 312898 and PI 319448) from *V. adenantha*. Our sequence data further confirmed the above phylogenetic relationship. Within the
Fig. 3. Sequence alignment of gene-derived marker BV165019 alleles generated from selected *Vigna* accessions produced by using AlignIR version 2.0.
species of *V. adenantha*, there was a base pair deletion at position 340 observed between PI 319448 and PI 312898 (Fig. 3).

Although accessions from different subspecies within cowpea were sequenced, there were no sequence polymorphisms identified (Fig. 3). The accessions from cowpea formed a cluster with a bootstrap value of 57% and no variation among accessions. The low genetic diversity revealed within cowpea may be explained by a single domestication event (Pasquet, 1999). In comparison with cowpea accessions, *V. vexillata* (PI 406390) had a four-base-pair insertion (154–157), a large deletion (182–194) and one-base-pair deletion (259) (Fig. 3). Within the Asian bean group, two accessions from each species were sequenced and all sequenced Asian bean accessions formed a group (Fig. 4) with little diversity among the species within this group. Sequence variation (point
mutation, insertion or deletion) was identified among these five species. However, only one point mutation (A/G) at position 196 was identified within the species *V. mungo* (Fig. 3). The species *V. longifolia* (PI 310294) within the African bean group was distinguished from other African bean accesses by several smaller deletions (Fig. 3). The species *V. subterranea* (PI 240867 and PI 245951) were separated from the species *V. luteola and V. oblongifolia*. There was a point mutation identified (from T to C) at position 258 within the species *V. luteola*. Overall, sequencing gene-derived amplicons detected more variation among species than within species, which was consistent with the results from detecting size difference of DNA fragments on an agarose gel. The topologies of the two dendrograms generated from DNA fragment data and DNA sequence data were very similar (Figs 2 and 4). The method of sequencing DNA amplicons for characterization of germplasm may be expensive, but the method for separating DNA fragments on agarose gels may require processing a large number of DNA markers to be effectively utilized in some genera.

(iii) **Seed morphological observation and phylogenetic relationships**

Seed morphology from 48 accessions was observed and recorded (shown in Fig. 5). The phylogenetic analysis showed that two accessions (PI 146800 and PI 322588) from *V. caracalla* were placed into different groups. Accession PI 146800 had small speckled tan colour seeds (1.93 g per 100 seeds), whereas PI 322588 had large solid brown seeds (5.22 g per 100 seeds) (Fig. 5 and Table 1). Furthermore, within 313 base pairs of the sequenced amplicons, eight point mutations were identified between these two accesses (Fig. 3) and these two accesses were also clustered into the same subgroup from the phylogenetic analysis of markers and sequence data (Figs 2 and 4). It is suspected that these two accesses may belong to different species. Actually, the seed morphologies of PI 322588 (*V. caracalla*) and PI 312898 (*V. adenantha*) were very similar (Fig. 5) and these two accesses were also clustered into the same subgroup from the phylogenetic analysis of markers and sequence data (Figs 2 and 4). It is suspected that these two accesses may belong to the same species (*V. adenantha*). To confirm this speculation, more accesses were classified as *V. adenantha* were requested from the Griffin seed store and their seed morphology was compared. The seed morphology of these two accesses and other accesses from the same species was very similar. Further confirmation of the possible misidentified accesses will include collecting more observational data (for example, seedling morphology, flowering characteristics and other traits) and experiments need to be conducted. Future work will include growing these accesses and collecting descriptor data to determine whether this accession has been misclassified or mislabelled during curation of this crop.

There were five accesses investigated from *V. angularis* for phylogenetic analysis. One of them (PI 527686) was the wild form, which was different from the other four accesses (Fig. 2). The seed size of PI 527686 (2.37 g per 100 seeds) was almost four times smaller than the cultivated form (from 8.16 to 9.59 g per 100 seeds, Table 1). From phylogenetic analysis, three *V. subterranea* accesses clustered into two distinct groups. The accession (PI 378867) was distinct from the remaining accesses (PI 240867 and PI 245951) (Fig. 2). The seed morphology was consistent with the phylogenetic analysis. The former accession had a red seed coat, whereas the latter two accesses had reddish brown seed coats.

Phylogenetic analysis from DNA fragment data classified 12 investigated cowpea accesses into two small clusters. The first small cluster contained three accesses (PI 632904, PI 632910 and PI 632903) and the second small cluster contained nine accesses (from PI 291384 to PI 582578) (Fig. 2). All accesses within the first cluster had small seeds (1.23, 1.6 and 1.78 g, respectively) and this cluster may be called the ‘wild group’. All accesses except one within the second small cluster had large seeds (from 7.72 to 20.89 g) (Table 1 and Fig. 5) and this small cluster may be called the ‘domesticated group’. The result from seed morphological observation was consistent with those from phylogenetic analysis. Our results demonstrated that phylogenetic analysis with morphological re-examination may provide a more complete approach to classify accesses or to examine misidentified accesses in a plant germplasm collection.

Very limited genetic variation (especially diversity within a species) was detected within the USDA *Vigna* germplasm collection based on the present and previous studies. Currently, some genetic gaps exist in the USDA *Vigna* collection. For example, on the botanical variety level, there are at least three varieties (var. *radiata*, var. *sublobata* and var. *setulosa*) available within *V. radiata* (Bisht et al., 2005). However, most accesses preserved in the USDA collection are from *V. radiata* var. *radiata*. Only one accession is maintained from *V. radiata* var. *sublobata*, while there are no accesses classified as *V. radiata* var. *setulosa* in the USDA collection. There are at least two botanical varieties (*V. mungo* var. *mungo* and *V. mungo* var. *silvestris*) within *V. mungo* (Seethalak et al., 2006), but only *V. mungo* var. *mungo* was collected and maintained as part of our collection. At the species level, some newly described ones (for example, *V. aridicola*, *V. exilis*, *V. nepalensis*, *V. tenuicaulis* and other species) are available (Tomooka et al., 2002). Tomooka et al. found that the species *V. aridicola* was closely related to *V. aconitifolia* (moth bean), *V. exilis* was

https://doi.org/10.1017/S0016672308009889 Published online by Cambridge University Press
closely related to *V. umbellata* (rice bean), and both *V. nepalensis* and *V. tenicaulis* were closely related to *V. angularis* (azuki bean). Although these newly described species could be potentially important for improving the cultivated species, none of them have yet been added to the USDA *Vigna* germplasm collection.

In conclusion, gene-derived markers are efficient to reveal phylogenetic relationships. Forty-seven *Vigna* accessions have been classified into four notable groups. Gene-derived markers are more effective at revealing polymorphism among species than within species. A few polymorphisms were identified within species by sequencing amplicons generated from

![Image](https://example.com/image.png)
gene-derived primers. The classification from DNA fragment analysis was consistent with the classification from DNA sequence analysis. Moreover, the genetic classification was supported by seed morphological observation. There was limited genetic diversity within the current USDA Vigna germplasm collection. In order to expand the genetic base of the USDA Vigna germplasm, new botanical varieties, subspecies and species need to be added to the USDA collection by germplasm curation and exchanges.

We are grateful to Drs Brad Morris, Richard Fery and Paul Raymer for comments and suggestions to improve the quality of the manuscript, Mrs Lee-Ann Chalkey for providing the seeds and Mr James Chalkey for help with scanning the seeds for morphological observation. We thank Mr Brandon Tonnis for his excellent technical assistance.

References

Amadou, H. I., Bebeli, P. J. & Kaltsikes, P. J. (2001). Genetic diversity in Bambara groundnut (Vigna subterranean L.) germplasm revealed by RAPD markers. Genome 44, 995–999.

Ba, F. S., Pasquet, R. S. & Gepts, P. (2004). Genetic diversity in cowpea (Vigna unguiculata (L.)) as revealed by RAPD markers. Genetic Resources and Crop Evolution 51, 539–550.

Barkley, N. A., Wang, M. L., Gillaspe, A. G., Dean, R. E., Pederson, G. A. & Jenkins, T. M. (2008). Discovering and verifying DNA polymorphism in a mung bean [V. radiate (L.) R. Wilczek] collection by EcoTILLING and sequencing. BMC Research Notes 1, 28.

Bisht, I. S., Bhat, K. V., Lakanpal, S., Latha, M., Jayan, P. K., Biswas, B. K. & Singh, A. K. (2005). Diversity and genetic resources of wild Vigna species in India. Genetic Resources and Crop Evolution 52, 53–68.

Choi, H. K., Luckow, M. A., Doyle, J. & Cook, D. R. (2006). Development of nuclear gene-derived molecular markers linked to legume genetic maps. Molecular Genetics and Genomics 276, 56–70.

Coulibaly, S., Pasquet, R. S., Papa, R. & Gepts, P. (2002). AFLP analysis of the phenotypic organization and genetic diversity of Vigna unguiculata L. Walp. reveals extensive gene flow between wild and domesticated types. Theoretical and Applied Genetics 104, 358–366.

Diof, D. & Hilu, K. W. (2005). Microsatellites and RAPD markers to study genetic relationships among cowpea breeding lines and local varieties in Senegal. Genetic Resources and Crop Evolution 52, 1057–1067.

Doi, K., Kaga, A., Tomooka, N. & Vaughan, D. A. (2002). Molecular phylogeny of genus Vigna subgenus Ceratotropis based on rDNA ITS and atpB-trnL intergenic space of cpDNA sequences. Genetics 114, 129–145.

Faris, D. G. (1965). The origin and evolution of the cultivated forms of Vigna sinensis. Canadian Journal of Genetics and Cytology 7, 433–452.

Fatokun, C. A., Danesh, D., Young, N. D. & Stewart, E. L. (1993). Molecular taxonomic relationships in the genus Vigna based on RFLP. Theoretical and Applied Genetics 86, 97–104.

Felsenstein, J. (2005). PHYLIP (Phylogeny Inference Package), Version 3.6 [computer program]. Distributed by the author. Seattle, WA: Department of Genomics, University of Washington.

Fery, R. L. (1980). Genetics of Vigna. In Horticultural Reviews, vol. 2 (ed. J. Janick), pp. 311–394. Westport, CT: AVI Publishing Company.

Gillaspie, A. G., Hopkins, M. S. & Dean, R. E. (2005). Determining genetic diversity between lines of Vigna unguiculata subspecies by AFLP and SSR markers. Genetic Resources and Crop Evolution 52, 245–247.

Goel, S., Raina, S. N. & Oghara, Y. (2002) Molecular evolution and phylogenetic implications of internal transcribed spacer sequences of nuclear ribosomal DNA in the Phaseolus–Vigna complex. Molecular Phylogenetics and Evolution 22, 1–19.

Jaaska, V. (1999). Isoenzyme diversity and phylogenetic affinities among the African beans of the genus Vigna Savi (Fabaceae). Biochemical Systematics and Ecology 27, 569–589.

Jaaska, V. (2001). Isoenzyme diversity and phylogenetic relationships among the American beans of the genus Vigna Savi (Fabaceae). Biochemical Systematics and Ecology 29, 1153–1173.

Jaaska, V. & Jaaska, V. (1990). Isoenzyme variation in Asian beans. Botanical Acta 103, 281–290.

Kaga, A., Tomooka, N., Egawa, Y., Kosaka, K. & Kamijima, O. (1996). Species relationships in the subgenus Ceratotropis (genus Vigna) as revealed by RAPD analysis. Euphytica 88, 17–24.

Kumar, S. V., Tan, S. G., Quah, S. C. & Yusoff, K. (2002). Isolation of microsatellite markers in mung bean, Vigna radiata. Molecular Ecology Notes 2, 96–98.

Lakhanpal, S., Chadha, S. & Bhat, K. V. (2000). Random amplified polymorphic DNA (RAPD) analysis in India mung bean (Vigna radiata (L.)) cultivars. Genetica 109, 227–234.

Li, C. D., Fatokun, C. A., Ubi, B., Singh, B. B. & Scoles, G. J. (2001). Determining genetic similarities and relationships among cowpea lines and cultivars by microsatellite markers. Crop Science 41, 189–197.

Mimura, M., Yasuda, K. & Yamaguchi, K. (2000). RAPD variation in wild, weedy and cultivated azuki beans in Asia. Genetic Resources and Crop Evolution 47, 603–610.

Minch, E., Ruiz-Linares, A., Goldstein, D., Feldman, M., Kidd, J. R. & Cavalli-Sforza, L. L. (1997). Microsat 1.5: A Computer Program for Calculating Various Statistics on Microsatellite Allele Data. Available at http://hpglstanfordedu/projects/microsat/programs/

Page, R. D. M. (1996). TREEVIEW. Version 1.6.6. An application to display phylogenetic trees on personal computers. Computer Applied Bioscience 12, 357–358.

Pasquet, R. S. (1999). Genetic relationships among subspecies of Vigna unguiculata (L.) Walp. based on allozyme variation. Theoretical and Applied Genetics 98, 1104–1119.

Pasquet, R. S. (2000). Allozyme diversity of cultivated cowpea Vigna unguiculata (L.) Walp. Theoretical and Applied Genetics 101, 211–219.

Santalla, M., Power, J. B. & Davey, M. R. (1998). Genetic diversity in mung bean germplasm revealed by RAPD markers. Plant Breeding 117, 473–478.

Seelahal, W., Tomooka, N., Waranyuwat, A., Thipyapong, P., Laosuwan, P., Kaga, A. & Vaughan, D. A. (2006). Genetic diversity of the Vigna germplasm from Thailand and neighboring regions revealed by AFLP. Genetic Resources and Crop Evolution 53, 1043–1059.

Simon, M. V., Benko-Iseppon, A. M., Resende, L. V., Winter, P. & Kahl, G. (2007). Genetic diversity and phylogenetic relationships in Vigna Savi germplasm revealed by DNA amplification fingerprinting. Genome 50, 538–547.
Singh, R. J. (2005). Landmark research in grain legumes. In Genetic Resources, Chromosome Engineering, and Crop Improvement (ed. R. J. Singh & P. P. Jauhar), pp. 1–9. Boca Raton, FL: Taylor and Francis Group.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acid Research 25, 4876–4882.

Tomooka, N., Yoon, M. S., Doi, K., Kaga, A. & Vaughan, D. (2002). AFLP analysis of diploid species in the genus Vigna subgenus Ceratotropis. Genetic Resources and Crop Evolution 49, 521–530.

Verdcourt, B. (1970). Studies of the Leguminosae–Papilionoideae for the flora of tropical East Africa. IV. Kew Bulletin 24, 507–570.

Wang, M. L., Mosjidis, J. A., Morris, J. B., Dean, R. E., Jenkins, T. M. & Pederson, G. A. (2006). Genetic diversity of Crotalaria germplasm assessed through phylogenetic analysis of EST-SSR markers. Genome 49, 707–715.

Xu, R. Q., Tomooka, N., Vaughan, D. A. & Doi, K. (2000). The Vigna angularis complex: genetic variation and relationships revealed by RAPD analysis, and their implications for in situ conservation and domestication. Genetic Resources and Crop Evolution 47, 123–134.

Yoon, M. S., Lee, J., Kim, C. Y. & Baek, H. J. (2007). Genetic relationships among cultivated and wild Vigna angularis (Willd.) Ohwi et Ohashi and relatives from Korea based on AFLP markers. Genetic Resources and Crop Evolution 54, 875–883.

Zong, X. X., Kaga, A., Tomooka, N., Wang, X. W., Han, O. K. & Vaughan, D. (2003). The genetic diversity of the Vigna angularis complex in Asia. Genome 46, 647–658.