B A
S E

Biotechnol. Agron. Soc. Environ. 1998 **2** (2), 128–134

Use of PCR-RFLP on chloroplast DNA to investigate phylogenetic relationships in the genus *Phaseolus*

Xavier Vekemans [(1)], Olivier Hardy [(1)], Bruno Berken [(1)], Bourlaye Fofana [(2)], Jean-Pierre Baudoin [(2)]

[(1)] Laboratoire de Génétique et d'Écologie végétales. Université Libre de Bruxelles. 1850 Chaussée de Wavre. B-1160 Bruxelles (Belgique).
[(2)] U.E. Phytotechnie des régions intertropicales. Faculté universitaire des Sciences agronomiques de Gembloux. 2, Passage des Déportés. B-5030 Gembloux (Belgique).

Received 9 December 1997, accepted 23 January 1998.

Phylogenetic relationships among 74 accessions belonging to six species of *Phaseolus* are investigated using variation in chloroplast DNA assessed according to a PCR-RFLP protocol. Three fragments of chloroplast DNA are amplified using universal primers, and then digested with 10 restriction enzymes. Thirty-six haplotypes are identified on the basis of the polymorphism in fragment number and size. Three main phylogenetic groups, strongly supported through bootstrap analysis, are identified: (1) accessions from *Phaseolus lunatus* and *Phaseolus xolocotzii*; (2) accessions from *Phaseolus glabellus*; (3) accessions from *Phaseolus vulgaris*, *Phaseolus polyanthus* and *Phaseolus coccineus*. Within the third group, accessions of *Phaseolus coccineus* are scattered along the phylogenetic tree, which provides some evidence that *coccineus* accessions are paraphyletic with respect to *Phaseolus vulgaris* and *Phaseolus polyanthus.* An analysis of molecular variance applied on four species show that they are significantly differentiated with 79% of molecular variance among species and 21% within species. The results agree with previous investigations on chloroplast DNA variation in the genus *Phaseolus*, and suggest that PCR-RFLP methods, which are technically less labour-intensive than previous methods, are of great value for phylogenetic investigations at the generic level.
Keywords. *Phaseolus*, phylogeny, chloroplast DNA, PCR, RFLP, plant genetic resources.

Utilisation de la PCR-RFLPsur de l'ADN chloroplastique pourl'étude des relations phylogénétiques au sein du genre *Phaseolus*. Les relations phylogénétiques entre 74 introductions appartenant à six espèces du genre *Phaseolus* ont été étudiées à partir de la variation d'ADN chloroplastique déterminée par un protocole de PCR-RFLP. Trois fragments d'ADN chloroplastique ont été amplifiés à l'aide d'amorces universelles, et ensuite digérés avec dix enzymes de restriction. Trente-six haplotypes sont identifiés sur la base du polymorphisme du nombre et de la taille des fragments. Trois groupes phylogénétiques, fortement corroborés par l'analyse de "bootstrap", sont identifiés : (1) les introductions de *P. lunatus* et *P. xolocotzii* ; (2) les introductions de *P. glabellus* ; (3) les introductions de *P. vulgaris*, *P. polyanthus* et *P. coccineus*. Au sein du troisième groupe, les introductions de *P. coccineus* sont largement distribuées à travers l'arbre phylogénétique ce qui suggère que *P. coccineus* est paraphylétique par rapport à *P. vulgaris* et *P. polyanthus*. Une analyse de la variance moléculaire, appliquée à quatre espèces, montre qu'elles sont significativement différenciées avec 79 % de la variance moléculaire entre espèces contre 21 % à l'intérieur des espèces. Les résultats sont en accord avec les études précédentes sur la variation chloroplastique au sein du genre *Phaseolus*, et suggèrent que la méthode de PCR-RFLP, qui est techniquement plus aisée que les méthodes précédentes, possède un intérêt considérable pour les études phylogénétiques au niveau générique.
Mots-clés. *Phaseolus*, phylogénie, ADN chloroplastique, PCR, RFLP, ressource génétique végétale.

INTRODUCTION

In the context of conservation of plant genetic resources, investigations on phylogenetic relationships among cultivated species and non-cultivated relatives are useful to highlight priorities in the *ex situ* conservation approaches. Many studies in plants are using chloroplast DNA (cpDNA) variation as a phylogenetic marker (Clegg, Zurawski, 1992), and the most widespread methodology involves restriction digestion followed by electrophoretic separation, Southern transfer to a membrane, and hybridization to a probe (Dowling *et al.*, 1990). An alternative method, using PCR amplification with specific or universal primers, followed by restriction digestion and electrophoretic separation of the fragments, was introduced recently in plant phylogenetic studies (Rieseberg *et al.*, 1992; Tsumura *et al.*, 1995). In a

preliminary study, we showed that this method was useful to estimate cpDNA diversity within several cultivated species of the genus *Phaseolus* (Vekemans *et al.*, 1997).

The genus *Phaseolus* comprises about 55 species (Debouck, 1991), five of which are cultivated species. According to a morphological study by Maréchal *et al.* (1978), the genus is divided into two main groups of species, the *P. vulgaris* L. – *P. coccineus* L. complex, on the one hand, and the group of *P. lunatus* L. and its allies, on the other hand. This pattern has been confirmed by studies on electrophoresis of seed proteins (Manen, Otoul, 1981; Sullivan, Freytag, 1986) and experimental hybridization (Baudoin, Maréchal, 1991). Investigations using cpDNA variation also confirm the existence of these two groups, but the existence of a third group comprising the species *P. glabellus* Piper on its own was suggested, however (Llaca *et al*., 1994; Schmit *et al*., 1993).

This paper is aimed to present results on the use of PCR-RFLP on cpDNA in 74 accessions belonging to the world seedbank collection of the genus *Phaseolus*, to investigate phylogenetic relationships among the six following species: *P. vulgaris*, *P. coccineus*, *P. polyanthus* Greenman, *P. glabellus*, *P. xolocotzii* Delgado, and *P. lunatus*.

MATERIALS AND METHODS

Plant material

Seeds from 74 accessions, sampled throughout the distribution range of the species, were obtained from the *Phaseolus* world collection held at the "Centro Internacional de Agricultura Tropical" CIAT, Cali, Colombia (G accession numbers) and the Belgium National Botanic Garden in Meise (NI and X accession numbers) (See **table 1**). Seeds are allowed to germinate and seedlings grown in a greenhouse for a few weeks. One plant for every accession is used in the molecular analysis.

Molecular methods

DNA isolation is performed on leaf material, according to Fofana *et al*. (1997). Three fragments of cpDNAare amplified by PCR using universal primers described by Demesure *et al*. (1995):
(P1) trnD[tRNA-Asp(GUC)]/trnT[tRNA-Thr(CGU)];

Table 1. Origin of the 74 *Phaseolus* accessions — *Origine des 74 introductions de* Phaseolus.

P. vulgaris	*P. polyanthus*	*P. coccineus*	*P. lunatus*	*P. xolocotzii*	*P. glabellus*
1 G 5686	11 G 35877	36 G 35846	59 G 25551	71 NI 1046	73 NI 304
2 NI 637	12 NI 1023	37 G 35851	60 G 25756	72 X 900	74 NI 820
3 NI 1471	13 NI 1022	38 NI 16	61 G 26144		
4 NI 1429	14 NI 489	39 NI 403	62 NI 529		
5 NI 1396	15 NI 1123	40 NI 606	63 NI 1280		
6 NI 1395	16 NI 1025	41 NI 737	64 NI 1275		
7 NI 1433	17 NI 519	42 NI 762	65 NI 1276		
8 NI 1355	18 NI 1021	43 NI 822	66 NI 1466		
9 NI 1470	19 NI 1208	44 NI 886	67 NI 1421		
10 NI 1469	20 NI 520	45 NI 889	68 NI 1413		
	21 NI 553	46 NI 890	69 NI 1460		
	22 NI 1159	47 NI 912	70 NI 1467		
	23 NI 1015	48 NI 1108			
	24 NI 1340	49 NI 1109			
	25 NI 763	50 NI 1110			
	26 NI 757	51 NI 1124			
	27 NI 758	52 NI 1133			
	28 NI 1166	53 NI 1158			
	29 NI 1294	54 NI 1160			
	30 NI 665	55 NI 1163			
	31 NI 429	56 NI 1265			
	32 NI 1128	57 NI 1330			
	33 NI 1010	58 NI 1428			
	34 NI 1024				
	35 NI 1011				

G: CIAT accession numbers
NI and X: Belgium National Botanic Garden accession numbers

(P2) trnS[tRNA-Ser(GGA)]/trnT[tRNA-Thr(UGU)]; (P3) trnM[tRNA-Met(CAU)]/rbcL[RuBisCO large subunit].

PCR reactions are performed in 100 µl of the following PCR-mix: 100 µM dNTP, 1.5 mM $MgCl_2$, PCR-buffer [from Eurogentec, 10X buffer as 750 mM Tris-HCl pH 9.0, 200 mM $(NH_4)_2SO_4$, 0.1% (w/v) Tween 20], 1 U *Taq*-DNA polymerase (Goldstar, from Eurogentec), 25 pM of each primer, 150 ng total DNA. The thermocycler is programmed, using a touch-down protocol, for one cycle of denaturation at 94°C for 4 min, then 30 cycles of 1 min denaturation at 92°C/ 1 min annealing starting from 65°C to 50°C with a decrease of one degree every two cycles/2 min elongation at 72°C, then a final elongation for 15 min at 72°C. Five microliters of the PCR products are used for restriction digestion in a final volume of 15 µl for each of the following restriction enzymes (Eurogentec): *Alu* I, *Ava* I, *Hae* II, *Hha* I, *Hin*f I, *Msp* I, *Nde* II, *Rsa* I, *Scr*F I, and *Taq* I. The digested fragments are then visualised by ethidium bromide after gel electrophoresis either on 2% agarose gels (NuSieve 3:1) or on 5% polyacrylamide gels, depending on the size and number of fragments obtained. Data are collected as presence or absence of each scorable fragment (see **figure 1**).

Data analysis

The proportion of shared fragments between each pair of haplotypes is determined as

$$F = \frac{2m_{xy}}{m_x + m_y}$$

where m_x and m_y are the total number of fragments scored in haplotypes x and y, respectively, and m_{xy} is the number of shared fragments between x and y. The matrix of pairwise (1-*F*) values, i.e. the proportion of fragments which are not shared, is directly used to reconstruct phylogenetic relationships among haplotypes by the neighbour-joining method using procedure NEIGHBOR from the PHYLIP software package (Felsenstein, 1993). One thousand sets of bootstraps are performed by

(1) sampling, with replacement, individual fragments from the original data set,
(2) computing a new (1-*F*) matrix for each bootstrap,
(3) reconstructing trees as described above, and
(4) summarizing the results using procedure CONSENSE from PHYLIP.

An analysis of molecular variance (AMOVA, according to Excoffier *et al.*, 1992) is performed on

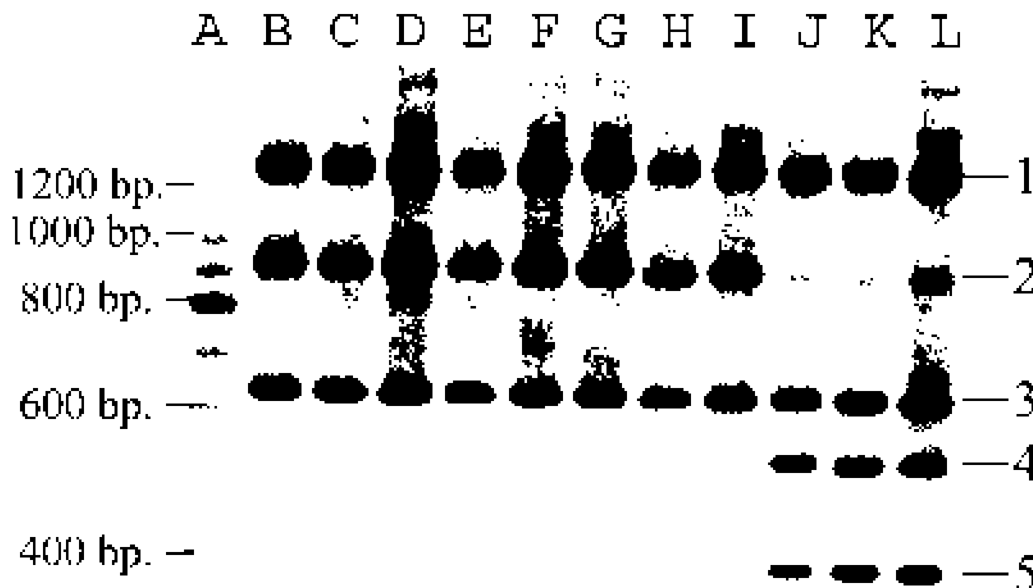

Figure 1. Example of restriction fragment patterns obtained after electrophoresis on a 2% agarose gel of 11 samples of DNA amplified using primers P3 and digested with *Scr*F I. Lane A corresponds to a 100 bp DNA-ladder marker (Pharmacia). Lanes B to E correspond to accessions of *P. polyanthus*, lanes F to I to accessions of *P. coccineus*, lanes J and K to *P. glabellus*, and lane L to *P. lunatus*. Fragments 1 and 3 are shared among all accessions. Fragment 2 is shared only among accessions of *P. polyanthus* and *P. coccineus*. Fragments 4 and 5 are shared only among accessions of *P. glabellus* and *P. lunatus* — *Exemple de patrons de restrictions obtenus après électrophorèse sur gel d'agarose à 2 % de 11 échantillons d'ADN amplifiés à l'aide des amorces P3 et digérés avec Scr*F I. *La ligne A correspond à une échelle de poids moléculaire par pas de 100 pb (Pharmacia). Les lignes B à E correspondent aux introductions de* P. polyanthus*, les lignes F à I aux introductions de* P. coccineus*, les lignes J et K à* P. glabellus*, et la ligne L à* P. lunatus. *Les fragments 1 et 3 sont présents parmi toutes les introductions. Le fragment 2 n'est présent que chez les introductions de* P. polyanthus *et* P. coccineus. *Les fragments 4 et 5 s'observent uniquement chez* P. glabellus *et* P. lunatus.

four species (*P. vulgaris*, *P. polyanthus*, *P. coccineus*, and *P. lunatus*) to estimate the proportion of cpDNA variation at intraspecific and interspecific levels, respectively. The AMOVA is performed using the program ARLEQUIN (Schneider *et al.*, 1997) with the matrix of presence or absence of each scorable fragment as input. This program tests the significance of the genetic differentiation among species using a numerical resampling method (permutation test).

RESULTS AND DISCUSSION

Number of fragments and haplotypes

During PCR, primer pairs P2 and P3 amplified fragments of sizes 1,180 bp (base pairs) and 2,650 bp, respectively, whereas fragments amplified with P1 had variable sizes ranging from 1,420 to 1,530 bp (**Table 2**). The seven restriction enzymes used on P1 fragments generated a total of 132 distinct DNA fragments. For P2, six restriction enzymes were used, and 30 fragments could be scored. For P3, 121 fragments were generated using eight restriction enzymes. A total of 283 fragments were scored, with 244 (80%) of them showing polymorphism among the 74 accessions studied.

Overall, 36 distinct haplotypes are identified. All 23 accessions of *P. coccineus* constitute distinct haplotypes, whereas only one haplotype is found among all 25 accessions of *P. polyanthus*. For *P. vulgaris*, three haplotypes are distinguished among the 10 accessions, and for *P. lunatus*, five haplotypes out of 12 accessions. Distinct haplotypes are also found for each of the two accessions of *P. glabellus* and of *P. xolocotzii*.

Table 2. Number of distinct fragments observed after digestion of cpDNAamplified with primer pairs P1, P2 and P3, from 74 accessions of *Phaseolus* species, using ten different restriction enzymes — *Nombre de fragments distincts observés après digestion d'ADN chloroplastique amplifié à l'aide des amorces P1, P2 et P3, pour 74 introductions d'espèces du genre* Phaseolus*, utilisant dix enzymes de restriction différents.*

Enzymes	Primer pairs P1	P2	P3
Size of undigested fragment (bp)	1,420–1,530	1,180	2,650
Alu I	7	n.a.	15
Ava I	n.a.	3	3
Hae III	n.a.	3	n.a.
Hha I	n.a.	5	n.a.
*Hin*f I	30	n.a.	19
Msp I	11	n.a.	11
Nde II	18	7	31
Rsa I	21	3	16
*Scr*F I	16	n.a.	5
Taq I	29	9	21
Total(1)	132(127)	30(24)	121(89)

n.a. = not analysed.
(1) the number of polymorphic fragments among the 74 accessions is given under parentheses.

Phylogenetic analysis

The phylogenetic tree obtained using neighbour-joining on the 36 haplotypes is shown on **figure 2** with bootstrap values higher than 75% indicated on individual branches. Three major groups of haplotypes are separated by long branches with bootstrap values higher than 95%:

(1) a group comprising haplotypes from *P. lunatus* and *P. xolocotzii*;

(2) a group containing the two haplotypes from *P. glabellus*;

(3) a large group comprising haplotypes from *P. coccineus*, *P. vulgaris* and *P. polyanthus*.

Groups (1) and (3) correspond to the two main groups of species of *Phaseolus*, according to the taxonomic study of Maréchal *et al*. (1978). The observation of a third main group, comprising only *P. glabellus*, is in accordance to the study of Schmit *et al*. (1993) which is also based on cpDNA variation but with a technique involving restriction digestion of the whole chloroplast DNA. Group (1) is further divided, with a bootstrap value of 100%, into haplotypes from *P. lunatus*, on the one hand, and haplotypes from *P. xolocotzii*, on the other hand. This is the first study reporting phylogenetic information on *P. xolocotzii* from cpDNA variation. It appears clearly that this species is closely related but distinct from *P. lunatus*. Haplotypes belonging to *P. coccineus* are found in almost all lineages within group (3) and show strong differentiation within the species. This confirms previous observations that cpDNA in *P. coccineus* is highly variable as compared to other species of *Phaseolus* (Llaca *et al*., 1994; Vekemans *et al*., 1997).

Haplotypes from *P. vulgaris* and *P. polyanthus* form a group supported only at a bootstrap value of 70.5% (data not shown), which also contains accession NI 1109 from *P. coccineus.* This group is embedded within the overall phylogeny of *P. coccineus* which is thus paraphyletic to *P. vulgaris* and *P. polyanthus*. According to previous investigations with cpDNA variation, *P. polyanthus* appeared more closely related to *P. vulgaris* than to *P. coccineus* (Schmit *et al*., 1993; Llaca *et al*., 1994), whereas based on nuclear markers the reverse was found to be true (Pinero, Eguiarte, 1988). From our results, it appears that *P. vulgaris* and *P. polyanthus* are closely related, but that they both lie within the range of variation of *P. coccineus*. This could explain the above discrepancy.

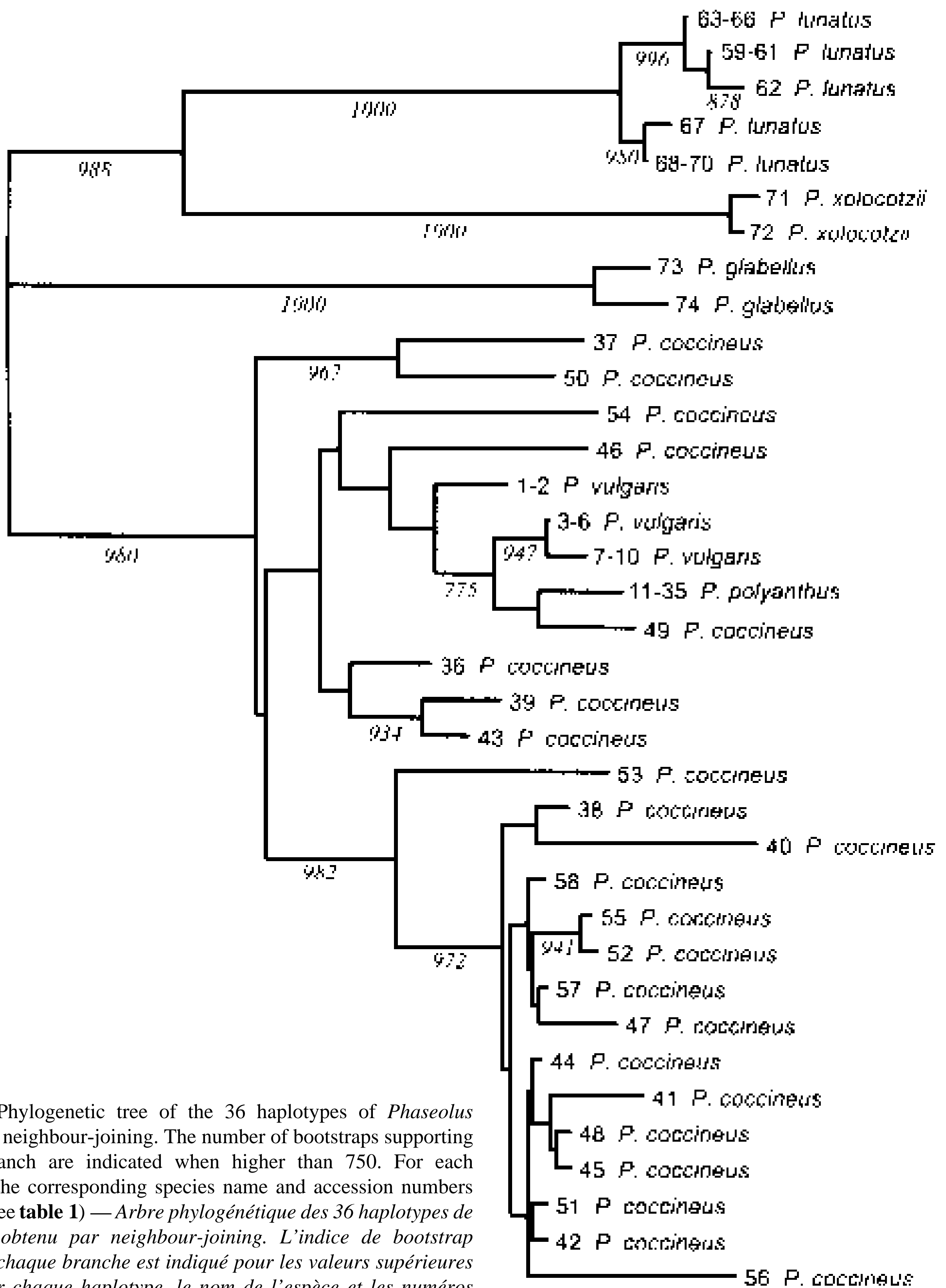

Figure 2. Phylogenetic tree of the 36 haplotypes of *Phaseolus* obtained by neighbour-joining. The number of bootstraps supporting a given branch are indicated when higher than 750. For each haplotype, the corresponding species name and accession numbers are given (see **table 1**) — *Arbre phylogénétique des 36 haplotypes de* Phaseolus, *obtenu par neighbour-joining. L'indice de bootstrap supportant chaque branche est indiqué pour les valeurs supérieures à 750. Pour chaque haplotype, le nom de l'espèce et les numéros d'introduction correspondants sont donnés (voir* **tableau 1***).*

Analysis of molecular variance (AMOVA)

An AMOVA was performed to estimate the proportion of cpDNA variation observed within or between species (**Table 3**). The analysis is restricted to the four species with appropriate sample size (*P. vulgaris*, *P. coccineus*, *P. polyanthus*, *P. lunatus*). Results show that there is significant variation in cpDNA among species (P < .001), which amounts to 79.2% of the total variation observed, but also that substantial variation at the intraspecific level is occurring (20.8% of the overall molecular variance). It appears thus that the PCR-RFLP method used in this study is useful to investigate phylogenetic relationships within the genus *Phaseolus*, but also that it may be of interest in phylogeographic studies of cultivated *Phaseolus* species. Its main advantage lies in the higher number of accessions which can be analysed concurrently, as opposed to other methods: in the present study we could survey 74 accessions of *Phaseolus* whereas Schmit *et al.* (1993) and Llaca *et al.* (1994) studied, respectively, 33 and 30 accessions.

Table 3. Analysis of molecular variance on cpDNA variation in four species of *Phaseolus* based on a matrix of presence or absence of restriction fragment — *Analyse de la variance moléculaire appliquée à la variation de l'ADN chloroplastique pour quatre espèces de* Phaseolus*, basée sur une matrice présence – absence de fragments de restriction.*

Source	Degrees of freedom	Sum of squares	Variance components	Percentage of variation
Among species	3	1129.9	22.23	79.2 (P < .001)
Within species	66	386.3	5.85	20.8
Total	69	1516.2	28.08	100.0

CONCLUSIONS AND PERSPECTIVES

In this study, we show that the genus *Phaseolus* can be divided in at least three major monophyletic groups of species, based on chloroplast DNA variation. One group includes *P. lunatus* and *P. xolocotzii*, another group is restricted to *P. glabellus*, and a third group comprises *P. vulgaris*, *P. polyanthus* and *P. coccineus*. Only six species were investigated here, and further studies are needed to include others. More specifically, species of undetermined phylogenetic status, such as the Mesoamerican species *Phaseolus ritensis* Jones, *Phaseolus polystachyus* B.S.P., and *Phaseolus maculatus* Scheele, should be included to test whether they actually belong to the phylogenetic group of *P. lunatus*. Moreover, analysis of a larger number of accessions from *P. vulgaris*, *P. polyanthus*, and *P. coccineus* would be useful to clarify the relationships among these three taxa, and to further test the paraphyletic relationship of *P. coccineus* to the other two species.

Acknowledgements

We thank C. Lefèbvre, P. Meerts, and P. du Jardin for discussions. This work was financed by grant n° 2.4515.92 from the Belgian National Fund for Scientific Research.

Bibliography

Baudoin JP., Maréchal R. (1991). Wide cross and taxonomy of pulse crops, with special emphasis on Phaseolus and Vigna. *In* Ng NQ., Perrino P., Attere A., Zedan H. *Crop genetic resources of Africa*. Ibadan, Nigeria: IITA-IBPGR-UNEP/CNR, p. 287–302.

Clegg MT., Zurawski G. (1992). Chloroplast DNA and the study of plant phylogeny: present status and future prospects. *In* Soltis PS., Soltis DE. *Molecular systematics of plants*. New York : Chapman & Hall, p. 1–13

Debouck DG. (1991). Systematics and morphology. *In* Van Schoonhoven A., Voysest O. *Common beans: research for crop improvement.* Wallingford, UK: C.A.B. International, p. 55–118.

Demesure B., Sodzi N., Petit RJ. (1995). A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNAin plants. *Mol. Ecol.* **4**, p. 129–131.

Dowling TE., Moritz C., Palmer JD. (1990). Nucleic acids II: restriction site analysis. *In* Hillis DM., Moritz C. *Molecular systematics.* Sunderland, USA: Sinauer associates, p. 250–317.

Excoffier L., Smouse PE., Quattro JM. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**, p. 479–491.

Felsenstein J. (1993). PHYLIP (Phylogeny Inference Package). Seattle, U.S.A. (Distributed by the author, Department of Genetics, University of Washington).

Fofana B., Vekemans X., du Jardin P., Baudoin JP. (1997). Genetic diversity in Lima bean (*Phaseolus lunatus* L.) as revealed by RAPD markers. *Euphytica* **95**, p.157–165.

Llaca V., Delgado SA., Gepts P. (1994). Chloroplast DNA as an evolutionary marker in the *Phaseolus vulgaris* complex. *Theor. Appl. Genet*. **88**, p. 646–652.

Manen JF., Otoul E. (1981). Etudes électrophorétiques et détermination des fractions protéiques principales chez quelques cultivars élites de *Phaseolus lunatus* L. et de *Phaseolus vulgaris* L. *Bull. Rech. Agron. Gembloux* **16**, p. 309–326.

Maréchal R., Mascherpa JM., Stainier F. (1978). *Étude taxonomique d'un groupe complexe d'espèces des genres* Phaseolus *et* Vigna (Papilionaceae) *sur la base de données morphologiques et polliniques, traitées par l'analyse informatique*. Genève, Suisse : Conservatoire et Jardin Botaniques de Genève.

Pinero D., Eguiarte L. (1988). The origin and biosystematic status of *Phaseolus coccineus* ssp. *polyanthus*: electrophoretic evidence. *Euphytica* **37**, p. 199–203.

Rieseberg LH., Hanson MA., Philbrick CT. (1992). Androdioecy is derived from dioecy in *Datiscaceae*: evidence from restriction site mapping of PCR-amplified chloroplast DNA fragments. *Syst. Bot*. **17**, p. 324–336.

Schmit V., du Jardin P., Baudoin JP., Debouck DG. (1993). Use of chloroplast DNA polymorphisms for the phylogenetic study of seven *Phaseolus* taxa including *P. vulgaris* and *P. coccineus*. *Theor. Appl. Genet.* **87**, p. 506–516.

Schneider S., Kueffer JM., Roessli D., Excoffier L. (1997). *ARLEQUIN, software for population genetic data analysis*. Geneva, Switzerland: Biometry and Genetics Laboratory, University of Geneva.

Sullivan JG., Freytag G. (1986). Predicting interspecific compatibilities in beans (*Phaseolus*) by seed protein electrophoresis. *Euphytica* **35**, p. 201–209.

Tsumura Y., Yoshimura K., Tomaru N., Ohba K. (1995). Molecular phylogeny of conifers using RFLP analysis of PCR-amplified specific chloroplast genes. *Theor. Appl. Genet*. **91**, p. 1222–1236.

Vekemans X., Berken B., Fofana B., Baudoin JP. (1997). Preliminary results on chloroplast DNA diversity estimation within *Phaseolus* species using PCR-RFLP. *Annu. Rep. Bean Improv. Coop.* **40**, p. 11–12.

(18 ref.)