Annals of Agric. Sci., Moshtohor **ISSN 1110-0419**
Vol. 60(1) (2022), 119 – 126 **https://assjm.journals.ekb.eg**

DNA-barcoding and Identification for Some Common Bean and Teparry Bean Genotypes using *rbcL* Gene

Mayada, Y. E.[1, 2], El-Akkad T. A.[1,2], Agez, A.[3], Refaat M. H.[1,2] and Abd-El-Sabour M. S.[1]
[1] Department of Genetics and Genetic Engineering, Faculty of Agriculture, Benha University, Qalubia, Egypt.
[2]Moshtohor Research Park, Molecular Biology Lab., Benha University, Egypt.
[3] Agricultural Genetic Engineering Research Institute (AGERI), ARC, Giza, Egypt.
Corresponding author**:** **mohamed.refaat@fagr.bu.edu.eg**

Abstract

Among the different biological sources, the common bean, tepary Bean, and the wild type (*Phaseolus spp*.) have many biotechnological applications. The DNA barcode of the use of the *rbcL* gene has proven its usefulness in the study of *Phaseolus spp.* Phylogenetic diversity of genotypes, more than a cryptic introduction, environmental modulation, geographic distribution, and species identification in exclusive bean species. Eleven samples of common bean had been obtained from unique locations in Egypt. The gene of rbcl had proven successful in PCR reaction. Seven samples of complete array with size 4031 base pair have been obtained. The spot ranged from 574 to 575 bp with a common size of 576 bp. The identification of the species published that the samples of the specimen River Garow, Bronco, Cerdon, Goru, Giza3 three belong to *Phaseolus vulgaris*, and two samples of the specimen Teparry113, Teparry116 belong to *Phaseolus acutifolius*, respectively. The *rbcL*-based DNA barcode was once nearly profitable to discover distinctive specimens of *Phaseolus spp* according to species and genus. Some samples of the rbcL gene were no longer sufficient to perceive the stage of the genus and could not differentiate between relatively comparable species.

Key words: Barcoding; *rbcL* gene; phylogenetic diversity*;* Common and Tepary Bean; *Phaseolus spp.*

Introduction

DNA barcoding relies on the discovery of unique conserved areas in different speciemen to collect huge library of standard genomes. DNA bar-coding is a universal way to know the species . Sequences of Nuclear and mitochondrial are lined up in rare speciemen to plan fully DNA-based barcode using to analyze (phylogenetic, genetic diversity, and species distinctions) in single creatures. The model classifies plant germplasm and shows genetic data of a species' ancestral origin inheritance behind it tags new types [***Son, et al. 2003 and Hebert, et al., 2004***].

Chloroplasts are an energetic metabolic team in experienced flowers for converting soft energy into Starches. High sequencing has resulted in the sequencing of nearly 800 distinctive flower genomes of chloroplast [***Kelchner, et al., 2000***]. Two conserved areas of the plastid (chloroplast) genome (*matk* + *rbcl*) had been submitted as barcode regions to characterize a massive collection of angiosperms. The molecular barcode picks out of revealing deception in plants. [***Glossman, 2007 and Sikdar and Dutta 2008***].

For profitable bar-coding experiment, the sequence distinction between 2 species should be more suitable for distinguishing between them . On the other hand , the intra and interspecies difference, the sequence version must be greatly reduced to discriminate the differences 7 chloroplast genomic regions were evaluated via the Plant Biogenic Barcode Working Group Consortium [***Jinbo, et al., 2011***].

The rbcl daily primer has excessive universality but low decision, while matK gives low universality but excessive decision between specific species. A mixture of each *matK* + *rbcl* marker can help discriminate most species. However, to achieve the highest rate of discrimination between two strongly associated species, the China Plant BOL group has already recommended a mixture of internal transcription spacer (*ITS* + *matK* + *rbcl*) [***Li, et al., 2011***].

DNA molecule is a safer than RNA and decided in all plant tissues. Therefore, thoroughly DNA-based markers are favored for the identification of legume plants.

Objectives of the present topic were 4 in number: (1) amplification and sequencing of rbcl regions in 11 common bean cultivars of *Phaseolus vulgaris* and two wild accessions of *Phaseolus acutifolius* (2) to study the homology model of the rbcl gene between two species using the Basic Local Alignment Research Tool (BLAST), (3) comparative genomic analysis, to study these sequences in elements between different genotypes of common bean and tepary bean (4) to explore the relationship between *Phaseolus ssp.* The development of the study would also be useful for the identification, authentication, and detection of alteration of the genetic material of legumes in national and international trade.

Materials and Methods

1. Plant Material:

A set of 11 common bean cultivars of Phaseolus vulgaris and two wild accessions of *Phaseolus acutifolius* (2n=22) were evaluated under Egyptian condition. The 11 *Phaseolus vulgaris* genotypes included belonged to the gene pools of common bean (River grow, GORU, Cerdon, Bronco, Giza3, Scala, Giza 6, Matilda, Paulista, Nambale and Samantha) and two wild accession Teparry113(G40083) and Teparry116(G40084),Seeds of these cultivars were obtained from the National Germplasm Resources Laboratory, Beltsville, USA and Germplasm Preservation Laboratory, Faculty of Agriculture at Moshtohor, Benha University and Horticulture Research Institute, Agricultural Research Center, Dokki, Egypt.

2. DNA Barcoding and Sequencing Analysis:

For most of the seven samples of wild and common beans (*Phaseolus spp.*), DNA was once extracted from (100 mg) of leaves the usage of DNeasy Plant Mini Kit (Qiagen, Santa Clarita, CA) and DNA template with specific bands in For Polymerase chain reaction amplification and rbcL sequencing regions, with specific primers were respectively *rbcLF* (5'ATGTCACCACAAACAGAGACTAAAGC3') and *rbcLR*(5'TCGCATGTACCTGCAGTAGC3') The product size predicted by PCR was 600 bp.

DNA amplification was performed by a thermal cycler(PCR) TProfessional (Biometra, Germany) using specific protocol and PCR products were posted on the UV Gel Documentation System (Fire Reader XSD5520Mand) with DNA Marker ladder (100 bp); Amplified products for all PCR had been purified the use of Qiagen® PCR Purification Kit(Qiagen, Santa Clarita, CA).

The purified PCR product and incubated at room temperature for two minutes and hold DNA at (-20°C). The rbcL PCR merchandise DNA sequencing by using Sanger DNA sequencing method (Macrogen® Inc., Seoul, Rep. of Korea).

The similarity search of the PCR product was performed with a Basic local alignment searching tool (BLAST), MEGA and Neighbor-Joining(NJ) (***Thompson, et al. 1994; Altschul, et al.1997 and Kumar et al. 2001***), using all the regions of the two different species of *Phaseolus* were used as seven request sequences (LC578835.1, LC578837.1, LC578838.1, LC578839.1, LC57840.1, LC57840.1 and LC578845.1).

Results

primers of PCR rbcL were successful in amplifying and producing a PCR product with expected band size (≠ 600 bp) (**Fig. 1**).

2. Species identification:

All species sequences had been efficiently recognized and most of the sequences had an identity level higher than 96%. However , the coverage of queries representing the size of comparable nucleotides between the query (our PCR product sequences) and the challenge (rbcL genes in the NCBI database) domain from 99.65% to 100%, indicating a similarity to all common bean queries. , could be the low variety of sequences of this species in the database or the high effectivity of the rbcL gene to pick out these genotypes.

The species identification revealed that the specimen's samples River Garow(LC578835.1), Bronco(LC578838.1), Cerdon (LC578837.1), GORU(LC578836.1), Giza 3(LC578839.1) belongs to *Phaseolus vulgaris*, and two specimen's samples Teparry113 (LC578840.1) , Teparry116 (LC578845.1) belongs to *Phaseolus acutifolius* , respectively.

3. Phylogenetic analysis:

The *rbcL* regions sequences together with immoderate similar GenBank genes were used to generate the data base of nucleotide for phylogenetic evaluation (Fig. 3). The phylogenetic tree advanced for LC578835.1 (River Garow) divided specimens from into four groups. The common branch duration changed into 0.0016, with a version of (0.0), domain from (-0.011 to 0.0009). The question (River Garow) with shaped a monophyletic clade together with exceptional species belongs to Phaseolus genus. While the Species (*Phaseolus lunatus* and *Phaseolus coccineus*) formed a particular clade (Fig. 3A).

In LC578838.1 (Bronco) the phylogenetic tree endorse size is (0.00148) with a variance of (61.44). Where the tree dimension domains from (0.0011 to 0.0032). The phylogenetic tree consists of 6 clades, the area Bronco genotype shaped one clade with any different *Phaseolus lunatus*. The *Phaseolus* genus shaped separate clades, which encouraged that rbcL gene can effectively differentiate between these genera (Fig. 3B).In query LC578837.1 (Cerdon), the phylogenetic tree suggests that the dimension is 0.00154 with a variance of 61.54. Where the tree dimension ranged from (0.0011 to 0.0027).

The phylogenetic tree is composed of 5 clades, the location Cerdon genotype formed one clade with each other *Phaseolus lunatus and Phaseolus coccineus*. The Phaseolus genus formed separate clades, which endorsed that rbcL gene can successfully differentiate amongst those genera (Fig. 3C).

The query's LC578839.1 and LC578838.1 (Giza3 and Bronco) genotypes fashioned with sequences belongs to the identical genus (*Phaseolus vulgaris*) a clade, whilst sample 6 shaped with rbcL sequence belongs to the identical species (Fig. 3D-E).Clearly clustered all rbcL sequences of Phaseolus lunatus in one clade and *Phaseolus coccineus* in another, indicating that even though sequences more than a few inside the species, variations are minimal pattern to cluster the same species in one clade indicating the DNA barcode conceivable of the rbcL gene.

Moreover, the phylogenetic tree built with LC578840.1 (Teparry-113) and different GenBank sequences divided into 6 clades, every with a recommend dimension and variance of (0.0069 and 98.55), respectively.

The branch sizes on the phylogenetic tree ranged from 0.0016 to 0.024. Different *Phaseolus ssp. rbcL* sequences formed a clade with the LC578840.1 (*Phaseolus acutifolius*) (Fig. 3F).

For the LC578845.1 (Teparry-113) tree, the suggested department size and variance were (0.0069 and 98.55), respectively. The branch size ranged from 0.013 to 0.003 mm. *Phaseolus lunatus, Phaseolus coccineus, Phaseolus carteri,* and *Orbexilum pedunculatum* sequences were all found in the LC578845.1 (*Phaseolus acutifolius*) clade (Fig. 3G).

Discussion

Information on plastid genome sequences is of brilliant price importance for barcoding plants. Sequence data performs a key position in figuring out and featured species, thus allowing the certificates of industrial and particularly products.

We used the chloroplast trnL (***Taberlet, et al., 2007***) and the rpoC1 zones and the ITS2 nuclear vicinity to barcode the necessary Mediterranean legume species. Using DNA barcoding sequences, (***Gao and Chen, 2009***) evaluated the opportunity of 4 coding chloroplast areas (rpoB, rpoC1, rbcL, and matK) and two noncoding nuclear areas (ITS, ITS2) serving as barcodes for the plants in the Fabaceae family.

The sequencing information for rbcL turned into applied to create a dendrography depiction of the variations through the species studied. Although it isn't always a real phylogenetic tree, it could be used to differentiate among *Phaseolus coccineus* and *Phaseolus lunatus*, in addition to *Vigna unguiculata* and *Vigna radiata*. Because of the technique hired to assemble the tree, the position of the various species covered on this examine withinside the phylogenetic tree can't be immediately as compared to placements mounted in different studies. We received 100% identity with *Phaseolus lunatus, Phaseolus coccineus, and Phaseolus carteri* from the National Center for Biotechnology Information database with *Phaseolus vulgaris*. DNA barcodes were approach in local species which could be a step to installation DNA totally definitely tracking protocols for economic plant adulteration below national and global trading. However, there are lot of legumes genomes lacking series information.

A genus totally based identity tool could be a technique of preference in species identity in indigenous plants. Resequencing of additional loci for target-primarily based totally upgrades could as an alternative be beneficial to study greater conserved genomic areas among notable plant species. The findings of contemporary-day find out about propose that using frequent primers (rbcL) for DNA barcoding is worthwhile for amplification, identity, and discrimination of above-cited indigenous plant species. The amplification fulfillment fees have been 100% for all the 3 species. Both the primer pairs resulted sharp bands that have been required for reliable DNA sequencing. Thus, using DNA barcodes primers systems is reliable, fast, and inexpensive tool identity of legumes flora at genus and species degree of indigenous land species.

Generating a large DNA database while focusing additional conserved areas would be effective in identifying legume flowers [***Padial, et al., 2010; Schindel and Miller, 2005***]. These data would also be useful in discovering the taxonomy, ecology, phylogeny, and morphology of distinctive species [***Hajibabaei, et al., 2007***]. However, improving new amplification protocols and techniques with novel primer cocktails would dramatically revolutionize the discipline of DNA barcodes through the representation of additional information about the genome of unique species.

Conclusion

Finally , Chloroplast genomes' sequences had been used for phylogenetic evaluation via ML based totally on unique Phaseolus species. *Phaseolus vulgaris* and *Phaseolus lunatus* are sister species, *Phaseolus lunatus* is extra carefully associated to *Phaseolus vulgaris, Vigna unguiculata* and *Vigna radiata*. Consistent with the gene order results, they're all of subtribe phaseolinae.

According to species and genus, *rbcL* based DNA barcoding was almost effective in identifying distinct seven common and wild bean genotypes specimens. Some *rbcL* specimens were unable to distinguish between extremely similar species and were unable to identify genus level. We propose that, in addition to *rbcL*, further DNA barcoding approaches be used to improve species identification resolution.

These findings are extremely helpful in the development of molecular markers for barcoding. Teparry bean has a unique genome when compared to other legume species (*Phaseolus acutifolius*).

References

Altschul SF., Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, 1997; **25**(17): 3389–3402. https://doi.org/10.1093/nar/25.17.3389

(Gao and Chen, 2009)

Glossman, M.D. CHIH-DFT determination of the molecular structure and infrared and ultraviolet spectra of-solanine. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, (2007); 66(1): 208-211.

Hajibabaei M, Singer GA, Hebert PD, Hickey DA. DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. Trends in Genetics, (2007); 23(4): 167-172.

Jinbo U, Kato T, Ito M. Current progress in DNA barcoding and future implications for entomology. Entomological Science, (2011); 14(2): 107-124.

Kelchner SA. The evolution of non-coding chloroplast DNA and its application in plant systematics. Annals of the Missouri Botanical Garden, (2000); 87(4): 482-498.

Kumar S, Tamura K, Jakobsen IB and Nei M (2001) MEGA2: molecular evolutionary genetics analysis software. Bioinformatics 17(12):1244–1245

Li FW, Kuo LY, Rothfels CJ, Ebihara A, Chiou WL, et al. rbcL and matK earn two thumbs up as the core DNA barcode for ferns. PLoS One, (2011); 6(10): e26597.

Padial JM, Miralles A, De la Riva I, Vences M. The integrative future of taxonomy. Frontiers in Zoology, (2010); 7(1): 1.

Schindel DE, Miller SE. DNA barcoding a useful tool for taxonomists. Nature, (2005); 435(7038): 17-17.

Sikdar M, Dutta U. Traditional phytotherapy among the Nath people of Assam. Ethno-med, (2008); 2(1): 39-45.

Son YO, Kim J, Lim JC, Chung Y, Chung GH. Ripe fruits of Solanum nigrum L. inhibits cell growth and induces apoptosis in MCF-7 cells. Food and Chemical Toxicology, (2003); 41(10): 1421-1428.

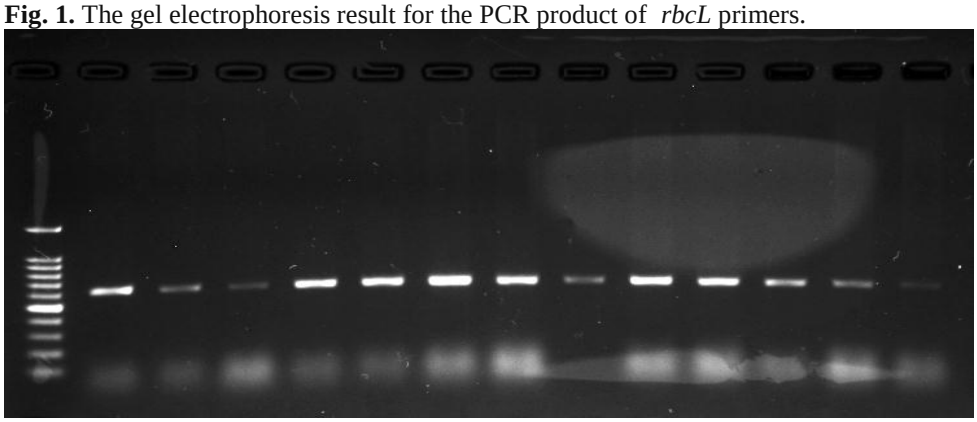
Taberlet P, Coissac E, Pompanon F and Gielly L (2007). Power and limitations of the chloroplast *trnL* (UAA) intron for plant DNA barcoding. *Nucleic Acids Res.* 35: e14

Thompson JD, Higgins DG, Gibson TJ and cluster W. Improving the sensitivity of progressive multiple
sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, 1994; **22**(22): 4673–4680. https://doi.org/10.1093/nar/22.22.4673.

Table 1 The NCBI blast results details for the Seven *Phaseolus spp*. genotypes *rbcL* Region sequences.

Genotypes	Taxonomic Level		Sequence Length	Max Score	Total Score	Sequence Cover	E - value Blast	Blast Similarity %	Similar Sequences No.
	Genus	Species							
River garow (LC578835.1)	*Phaseolus*	*vulgaris*	576	1061	1061	100%	0.0	100	
GORU (LC578836.1)	*Phaseolus*	*vulgaris*	577	1061	1061	100%	0.0	100	
Cerdon (LC578837.1)	*Phaseolus*	*vulgaris*	576	1061	1061	100%	0.0	100	LT576851.1 LT576853.1 MK348626.1 MK348625.1 LT576850.1 KJ773734.1
Bronco (LC578838.1)	*Phaseolus*	*vulgaris*	574	1061	1061	100%	0.0	100	
Giza3 (LC578839.1)	*Phaseolus*	*vulgaris*	576	1050	1050	99.65%	0.0	100	
Teparry-113 (LC578840.1)	*Phaseolus*	*acutifolius*	576	1050	1050	99.65%	0.0	100	
Teparry-116 (LC578845.1)	*Phaseolus*	*acutifolius*	576	1050	1050	99.65%	0.0	100	

Fig. 1. The gel electrophoresis result for the PCR product of *rbcL* primers.



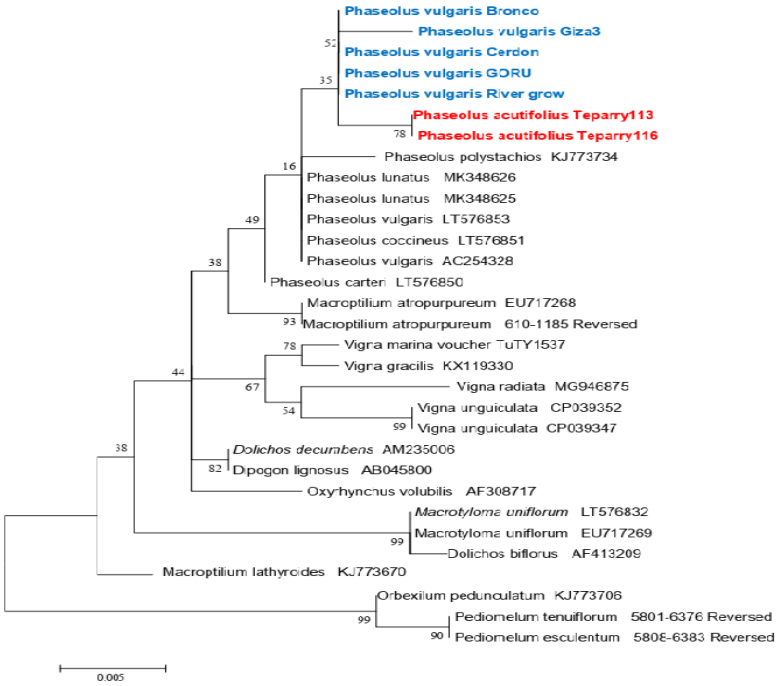


Fig. 2. The phylogenetic trees depicting the relationship between different *Phaseolus spp.* Genotypes specimens and GenBank sequences.

Fig. 3. The phylogenetic trees depicting the relationship between different *Phaseolus spp.* Genotypes specimens and GenBank sequences.

تشفير الحمض النووي DNA وتحديد بعض الانماط الوراثية الشائعة

للفاصوليا والفاصوليا أكوتيفولي أو تيباري (Teparry) باستخدام جين rbcL

ميادة يوسف السيد أ.م د تامر احمد العقاد ت. ، د عمرو محمد عجيز أ.د/ محمد حسن رفعت أ.د/ محمد سراج الدين عبد الصبور

1 قسم الوراثة والهندسة الوراثية ، كلية الزراعة ، جامعة بنها ، القليوبية ، مصر.

2 مجمع بحوث مشتهر ، معمل البيولوجيا الجزيئية ، جامعة بنها ، مصر.

3 معهد بحوث الهندسة الوراثية الزراعية (AGERI) ، مركز البحوث الزراعية ، الجيزة ، مصر

تم دراسة العديد من العينات التى تنتمى الى *(Phaseolus spp)*للتعرف على الرمز الجينى للحمض النووى DNA وذلك باستخدام الجين rbcL . تم الحصول على احدى عشر عينة من الفاصوليا من مواقع مختلفة من المحافظات المصرية تم اجراء هذا الاختبار عليها.

أثبت جين rbcl نجاحه في تفاعل البوليميريز المتسلسل. تم الحصول على سبع عينات من مجموعة كاملة بحجم 4031 زوج أساسي. وتراوحت البقعة من 574 إلى 575 نقطة أساس بحجم مشترك قدره 576 bp. تم تحديد الأنواع المنشورة بأن عينات River Garow، Bronco، Cerdon ، Goru، جيزة 3 تنتمي إلى Phaseolus vulgaris ، وعينتين من عينات Teparry113، Teparry116 تنتمي إلى Phaseolus acutifolius ، على التوالي. كان الرمز الشريطي للحمض النووي المستند إلى rbcL مناسبا لاكتشاف عينات مميزة من Phaseolus spp وفقًا للأنواع والجنس.

الكلمات الأساسية: تشفير ؛ جين rbcL التنوع الوراثي؛ Common Bean; Phaseolus