

RESEARCH ARTICLE

BLAST analysis of Phaseolin gene from common bean (*Phaseolus vulgaris*)

■ ROSY RANI¹, PANKAJ KUMAR², JITENDER SINGH², AMIT KUMAR SINGH¹, ANCHAL RANI¹, PRADEEP SHUKLA¹ AND PRAGATI MISRA^{1*}

¹Sam Higginbottom Institute of Agriculture, Technology and Sciences, ALLAHABAD (U.P.) INDIA

²Sardar Vallabhbhai Patel University of Agriculture and Technology, MEERUT (U.P.) INDIA

ARTICLE INFO

Received : 17.02.2014

Revised : 04.03.2014

Accepted : 14.03.2014

Key Words :

Phaseolus vulgaris, Phaseolin gene, Nucleotide Database, BLAST

*Corresponding author:

Email: pragatimisra3@rediffmail.com

ABSTRACT

The common bean (*Phaseolus vulgaris*) contains major seed storage protein which is known as the phaseolin. The phaseolin is the salt-soluble glycoprotein that accounts of some 50 per cent of the total protein in mature bean seeds. The computational packages and online servers are the current tools used in the nucleotide or protein sequence analysis and characterization. The query full length nucleotide sequence of phaseolin gene was subjected to BLAST analysis with the homologous nucleotide sequences. The sequences homologues to the phaseolin gene query sequence were retrieved on the bases of query coverage, maximum identity and e-value.

How to view point the article : Rani, Rosy, Kumar, Pankaj, Singh, Jitender, Singh, Amit Kumar, Rani, Anchal, Shukla, Pradeep and Misra, Pragati (2014). BLAST analysis of Phaseolin gene from common bean (*Phaseolus vulgaris*). *Internat. J. Plant Protec.*, 7(1) : 151-153.

INTRODUCTION

The common bean (*Phaseolus vulgaris*) is a herbaceous annual plant domesticated independently in ancient Mesoamerica and the Andes, and now grown worldwide for its edible bean, popular both dry and as a green bean. The leaf is occasionally used as a leaf vegetable, and the straw is used for fodder. It is a staple food for many people due to its energy, protein, dietary fibre and minerals content (Haytowitz *et al.*, 1981; Norton *et al.*, 1985). Botanically, the common bean is classified as a dicotyledon. The seeds of common bean have valuable nutritional properties due to the fact that they are an important source of fibre, minerals and vitamins, as well as to their low content of fat and sodium (Sgarbieri and Whitaker, 1982). A diet including beans provides substantial health benefits, decreasing the risk of heart and renal diseases (Kabagambe *et al.*, 2005), protecting against several cancer types (Kolonel *et al.*, 2000) and helping in the control of overweight and obesity (Celleno *et al.*, 2007).

The common bean containing major seed storage protein is known as the phaseolin (Bollini and Vitale, 1981; Gepts and

Bliss, 1986). Phaseolin is the salt-soluble glycoprotein that accounts for some 50 per cent of the total protein in mature bean seeds. It was one of the first plant proteins to be translated *in vitro* from mRNA and one of the first plant genes isolated. It was also the first developmentally regulated plant gene to be expressed in a heterologous plant species through *Agrobacterium*-mediated transformation. Studies on Phaseolin have provided insight to many aspects of plant protein synthesis, from fundamental molecular mechanisms to practical goals such as the improvement of the bean's nutritional quality. The computational packages and online servers are the current tools used in the nucleotide or protein sequence analysis and characterization. There has been an ever increasing growth in the size of nucleotide and protein databases due to advances in sequencing technology. There is a high demand in the area of computational biology to extract useful information from these massive databases. This has led to the usage and development of high performance computing power to help researchers sift through the voluminous biological data. Supercomputers, clusters, and custom designed bioinformatics hardware solutions attempt to tackle the various problems stemming from

analysis of biological data.

BLAST stands for Basic Local Alignment Search Tool. It was developed by the National Centre for Biotechnology Information (NCBI) in USA. Even though the name of BLAST suggests only local alignment, but in reality it can carry out both local and global alignment. The main idea behind BLAST was to feed the computer with an input sequence and produce an output when it is run against a selected database. The output is then interpreted to derive information about the input sequence. As mentioned earlier, the main purpose of using BLAST is sequence alignment. In carrying out a local alignment, BLAST breaks down an input sequence into smaller parts and compares them with the database. No gaps are introduced in local alignment in order to force the input sequence to match with the database. A match or a mismatch is indicated by the presence or absence of a vertical line between the alphabets of the input sequence and the database sequence. Unlike local alignment, sequences are guaranteed to match with the input sequence. Here, gaps are introduced whenever a mismatch is found and the input sequence is shifted by some places in order to match it with the database. One other important aspect of global alignment is that the whole sequence is taken at once and matched with the database, unlike in local alignment where an input sequence is broken down into smaller components.

MATERIAL AND METHODS

The BLAST (Basic Local Alignment Search Tool) is an algorithm for comparing primary biological sequence information, such as the amino-acid sequences of different proteins or the nucleotides of DNA sequences. A BLAST search enables a researcher to compare a query sequence with a library or database of sequences, and identify

library sequences that resemble the query sequence above a certain threshold. The obtained full length nucleotide sequence of *Phaseolin* gene through sequencing from YVR life sciences, Ghaziabad was subjected to BLAST analysis (BLASTX, National Center for Biotechnology Information - NCBI) with the homologous nucleotide sequences. The sequences homologous to the query sequence were retrieved on the bases of query coverage, maximum identity and e-value.

BLAST analysis for the alignment :

BLAST (Basic Local Alignment Search Tool) is an effective tool in bioinformatics. Inside the blast homepage (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) BLAST option was clicked. In the next page, the fasta format of this query nucleotide sequence was uploaded and selects the nucleotide database and finally clicked on BLAST option. After a waiting time the BLAST result appeared.

RESULTS AND DISCUSSION

The full length nucleotide sequence of *Phaseolin* gene was subjected to BLAST analysis (Table 1) with earlier reported phaseolin genes, which were retrieved from the NCBI database. The ten sequences were showed high range of query coverage with low e-value and maximum identity were used for BLAST analysis. These sequences showed maximum identity *i.e.* from 79 per cent to 100 per cent and wide range of query coverage *i.e.* from 27 per cent to 98 per cent with low e-value. The sequences *Phaseolus vulgaris* mRNA for beta-type phaseolin precursor (accession no. X03004.1), *Phaseolus vulgaris* mRNA for alpha-type phaseolin precursor (accession no. X02980.1), *Phaseolus vulgaris* Sanilac clone 2-13 phaseolin (Phs) mRNA, complete cds (accession no. U01132.1) and *Phaseolus vulgaris* Sanilac clone 1-12 phaseolin (Phs)

Table 1: Blast results of query nucleotide sequence from NCBI database

Accession	Description	Max score	Total score	Query coverage	E value	Max ident
X52626.1	<i>Phaseolus vulgaris</i> gene for alpha-phaseolin	3253	3253	98%	0.0	97%
J01263.1	<i>Phaseolus vulgaris</i> beta-type phaseolin storage protein gene, complete cds	3127	3451	98%	0.0	99%
V01163.1	Part of the gene for phaseolin in <i>Phaseolus vulgaris</i> (bean)	1450	1450	46%	0.0	96%
X03004.1	<i>Phaseolus vulgaris</i> mRNA for beta-type phaseolin precursor	584	2558	72%	9e-163	100%
X02980.1	<i>Phaseolus vulgaris</i> mRNA for alpha-type phaseolin precursor	564	2296	70%	1e-156	100%
U01132.1	<i>Phaseolus vulgaris</i> Sanilac clone 2-13 phaseolin (Phs) mRNA, complete cds	564	2363	73%	1e-156	100%
U01131.1	<i>Phaseolus vulgaris</i> Sanilac clone 1-12 phaseolin (Phs) mRNA, complete cds	562	2525	73%	4e-151	100%
U01121.1	<i>Phaseolus lunatus</i> phaseolin (Phs) mRNA, complete cds	492	1201	42%	6e-135	95%
AM905848.1	<i>Vigna unguiculata</i> partial vicilin gene, exons 1-6	545	545	42%	4e-151	79%
DQ538335.1	<i>Vigna radiata</i> 8S globulin beta isoform precursor mRNA, complete CDS	462	839	27%	4e-126	95%

mRNA, complete cds (accession no. U01131.1) showed 100 per cent identity with the query sequence of *Phaseolin* gene. These sequences showed 70 per cent to 73 per cent query coverage and low e-value. These data indicated that these sequences rely between the query sequence of *Phaseolin* gene. The *Phaseolus vulgaris* beta-type phaseolin storage protein gene, complete cds (accession no J01263.1) showed 99 per cent identity with the query coverage of 98 per cent with zero e-value. The *Phaseolus vulgaris* gene for alpha-phaseolin (accession no. X52626.1) has 97 per cent identity with query sequence and 98 per cent query coverage with zero e-value. The part of the gene for Phaseolin in *Phaseolus vulgaris* (bean) (accession no. V01163.1) showed 96 per cent identity with 46 per cent query coverage and zero e-value whereas *Phaseolus lunatus* phaseolin (Phs) mRNA, complete cds (accession no. U01121.1) and *Vigna radiata* 8S globulin beta isoform precursor mRNA, complete CDS (accession no. DQ538335.1) showed 95 per cent identity with 42 per cent and 27 per cent query coverage, respectively and low e-value. The *Vigna unguiculata* partial vicilin gene, exons 1-6 (accession no. AM905848.1) showed 79 per cent identity with low e-value and query coverage of 42 per cent on BLAST analysis with query sequence of *Phaseolin* gene (Table 1).

BLAST (Basic Local Alignment Search Tool) (Altschul *et al.*, 1990, Altschul *et al.*, 1997) is a widely used bioinformatics application for rapidly querying nucleotide (DNA) and protein sequence databases. Given a query sequence, the goal is to find the most similar sequences from a large database. This has applications in the identification of functions and structures of unknown sequences or understanding the evolutionary origin of DNA or protein query sequences. Due to its high importance, BLAST also finds a place in almost all the bioinformatics solutions.

REFERENCES

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990). Basic local alignment search tool. *J. Mol. Biol.*, **215**(3): 403-410.
- Altschul, S.F., Madden, L.T., Schffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997). Gapped blast and psi-blast: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**(17) : 3389-3402.
- Bollini, R. and Vitale, A. (1981). Genetic variability in charge microheterogeneity and polypeptide composition of phaseolin, the major storage protein of *Phaseolus vulgaris*; and peptide maps of its three major subunits. *Physiol. Plant*, **52** : 96–100.
- Celleno, L., Tolaini, M.V., D'Amore, A., Perricone, N.V. and Preuss, H.G. (2007). A dietary supplement containing standardized *Phaseolus vulgaris* extract influences body composition of overweight men and women. *Internat. J. Med. Sci.*, **4** (1) : 45–52.
- Gepts, P. and Bliss, F.A. (1986). Phaseolin variability among wild and cultivated common beans (*Phaseolus vulgaris*) from Colombia. *Econ. Bot.*, **40**(4) : 469-478.
- Haytowitz, D.B., Marsh, A.C. and Matthews, R.H. (1981). Content of selected nutrients in raw, cooked and processed legumes. *Food Tech.*, **35** (3) : 73-74.
- Kabagambe, E.K., Baylin, A., Ruiz-Narvarez, E., Siles, X. and Campos, H. (2005). Decreased consumption of dried mature beans is positively associated with urbanization and nonfatal acute myocardial infarction. *J. Nutr.*, **135** (7) : 1770–1775.
- Kolonel, L.N., Hankin, J.H., Whittemore, A.S., Wu, A.H., Gallagher, R.P. and Wilkens, L.R. (2000). Vegetables, fruits, legumes and prostate cancer: A multiethnic case-control study. *Cancer Epidem. Biomark. Prev.*, **9** (8): 795–804.
- Norton, G., Bliss, F.A. and Brezan, R. (1985). Biochemical and nutritional attributes of grain legumes. In : R. J. Summerfield, & E. H. Roberts (Eds.), *Grain legume crops* (pp. 73-114). London (UK).
- Sgarbieri, V.C. and Whitaker J.R. (1982). Physical, chemical and nutritional properties of common bean (*Phaseolus*) proteins. In: *Advances in food research*, Vol. 28, C.O. Chichester, E.M. Mrak, B.S. Schweigert (Eds.), Academic Press, New York, NY, USA pp. 93–166.

7th
Year
★★★★★ of Excellence ★★★★★