©2016 New Delhi Publishers. All rights reserved

BIOTECHNOLOGY

## Molecular characterization of mannose specific lectin gene, ASAL1 from Garlic leaf (*Allium sativum* L.)

Rajveer Singh\*, Navraj Kaur Sarao, Prashant Mohanpuria and Inderjit Singh Yadav

School of Agricultural Biotechnology, Punjab Agricultural University, Ludhiana, Punjab, India

\*Corresponding author: rgrewal372@gmail.com

Paper No. 425

Received: 23-3-2015

Accepted: 6-4-2016

#### Abstract

Plant lectins, especially monocot mannose-binding lectins (MMBLs) play significant roles in plant defense against sap sucking insects. In the present study, lectin gene (*ASAL*1) from garlic leaves (*Allium sativum* L. leaf agglutinin) was isolated, sequenced and characterized using various bioinformatics tools. Full-length cDNA of naturally occurring MMBLs was synthesized from garlic leaf RNA (leaf agglutinin, *ASAL*1) using RT-PCR and was amplified with gene specific primers designed corresponding to the conserved regions of the nucleotide sequences of garlic lectin sequences already available at NCBI. The amplified cDNA was sequenced. Sequence analysis revealed 369bp ORF including C- terminal stop codon, encoding a putative polypeptide of 122 amino acids (13kD) in *ASAL*1. The candidate gene sequence (*ASAL*1) was 30 nucleotides more and showed nine nucleotide changes than the previously reported garlic leaf lectin gene sequence of 339bp with accession number EU252577. *ASAL*1 gene sequence showed maximum (98%) identity with *Allium sativum* lectin mRNA complete cds having accession number DQ525625.1. In *ASAL*1 nine amino acid residues were glycosylated (both N and O linked). A putative conserved domain (4-113) was detected in the deduced amino acid sequence. *ASAL*1 gene is bulb type mannose binding lectin ( $\beta$ - lectin). Phylogenetic analysis revealed that *ASAL*1 falls in close relation with *ACA*.

#### Highlights

- The sequence analysis revealed that candidate gene *ASAL1*, was of 369 bp ORF including C- terminal stop codon, encoding a putative polypeptide of 122 amino acids (13kD).
- *ASAL1* was 30 nucleotides more and showed nine nucleotide changes than the previously reported garlic leaf lectin gene sequence of 339bp with accession number EU252577.
- *ASAL1* gene sequence showed maximum (98%) identity with *Allium sativum* lectin mRNA complete cds having accession number DQ525625.1.
- A putative conserved domain (4-113) was detected in the deduced amino acid sequence. *ASAL*1 gene is bulb type mannose binding lectin (β- lectin).
- Phylogenetic analysis revealed that ASAL1 falls in close relation with ACA.

Keywords: Garlic, cDNA, lectins, RT-PCR, MMBL, phylogenetic analysis

Lectins are proteins of non-immune origin that possess at least one non-catalytic domain that binds stereo-specifically and reversibly to carbohydrates and their derivatives, without initiating their further modification through associated enzymatic activity Van Damme *et al.* (2007). They act as cognate receptors to various cell surface glycoproteins, resulting in several important cell-mediated events ranging from mitogenic processes to plant defense mechanisms Loris (2002) and Sharon (2007). Occurrence of lectins in plants, animals and fungi had been detected much earlier, but plant lectins are



the best studied of them Pusztai (2008). Although, the exact function of these proteins has yet to be elucidated, the high concentrations and their source tissues suggest a role as storage proteins. Furthermore, lectins can also serve as defense molecules against insect herbivores and pathogens. Over the last two decades, abundant constitutively expressed lectins have been identified in many plant species and these show very diverse molecular structures and sugar specificities Van Damme *et al.* (2007). The possible mechanism of lectin toxicity in insects seems to involve the binding of lectin to the gut surface, leading to local lesions in the insect gut Eisemann et al. (1994), thus preventing nutrient absorption. Moreover, MMBLs have been proven to be toxic to hemiptrans both under in vitro as well as in planta conditions and non-toxic to higher animals Bandyopadhyay et al. (2001). Also, it was found to have negligible effects on the development, survival and fecundity of beneficial insects Down et al. (2003).

According to their molecular structure and evolutionary relationship, plant lectins can be subdivided into seven different families, namely legume lectins, chitin-binding lectins, monocot mannose binding lectins (MMBL's), jacalin-related lectins, amaranthine lectin family, cucurbitaceae lectin Van Damme et al. (1998). Mannose-binding lectins are widely distributed in monocotyledonous plants and are believed to play a crucial role in recognition of high-mannose type glycans of foreign micro-organisms or plants predators Barre et al. (2001). Since the first monocot mannose lectin from snowdrop (Galanthus nivalis) bulb was reported Van Damme et al. (1987), more and more monocot MBL's have been found in various tissues of monocot families Alliaceae, Amaryllidaceae, Areaceae, Bromeliaceae, Iridaceae, Liliaceae and Orchidaceae Van Damme et al. (1998). Insect bioassay studies showed that most of the tested mannose-binding lectins purified from Araceae species such as Ariaaema spp. and pinella spp. had more or less insecticidal activities towards cotton aphids (Aphids gossypii Glover) and peach potato aphids (Myzus persicae Sulzer) when incorporated into artificial diets Mao et al. (1999); Li et al. (2000); Yao et al. (2003).

The advent of genetic engineering based on recombinant DNA technology has made it

possible to improve specific traits by means of isolating, constructing and transforming useful genes. Therefore, when such lectins are expressed transgenically in plant systems, the plants have been shown to develop the capacity to resist the attack of the sap-sucking insect pests Gatehouse et al. (1996), Rao et al. (1998), Stoger et al. (1999), Foissac et al. (2000), Tang et al. (2001), Sun et al. (2002), Nagadhara et al. (2004) and Dutta et al. (2005). This is considered significant because no other insecticidal agents for transgenic use against sap-sucking insects have proven to be adequately effective. Snowdrop mannose specific lectin, Galanthus nivalis agglutinin (GNA), is the first lectin which has deleterious effects on aphids, leaf hoppers and plant hoppers Rao et al. (1998) and Stoger et al. (1999). Garlic (Allium sativum L.) mannose-binding leaf lectin (ASAL), was very effective against the chickpea aphid, brown plant hopper Majumder et al. (2004) the mustard aphid and the red cotton bug, Bandyopadhyay et al. (2001). The use of phloem specific promoters to express an insecticidal protein in phloem tissue of transgenic plants would be especially useful for conferring resistance to sapsucking insects. An added advantage of using plant insecticidal genes over bacterial genes is that they can be introduced to other plant system without much modification. Thus, the current understanding has been that transgenic plants with lectins offering even partial resistance against any sap-sucking insect would still find acceptance in agriculture Ferry et al. (2004) and could be a useful tool in integrated pest management strategy Banerjee *et al*. (2004). Similar to our study, cloning and molecular characterization of LECASAI lectin gene from garlic (Allum sativum L.) was done by Neha et al. (2014).

Therefore, in the present study, full cDNA sequence of lectin gene from garlic (*Allium sativum*) was cloned and *in-silico* characterized using various bioinformatic tools.

#### Materials and methods

#### Plant material

Plants of garlic variety PG17 were grown in glass house under standard conditions. For the isolation of garlic leaf lectin gene, total RNA was isolated from fresh leaves of four weeks old seedlings. Young garlic leaves were crushed in liquid nitrogen and cellular RNA was isolated using SV Total RNA isolation system (Promega, Madison WI, USA) according to the manufacturer's instructions. RNA pellet was in autoclaved 0.1% DEPC treated double distilled water. The cDNA was synthesized from 1µg of total RNA using oligo-dT primer and reverse transcriptase of Bio-Rad First-Strand cDNA synthesis kit, according to the manufacturer's protocol. In the earlier report, amplification of DNA fragments from genomic DNA by PCR showed no evidence for introns in the coding sequence of the garlic lectin genes Fitches *et al.* (2008). Therefore, in the present study, RT-PCR approach is used for the isolation of lectin gene and full length cDNA of lectin protein was successfully isolated.

# Isolation and Sequencing of garlic leaf lectin encoding gene ASAL1

The PCR (Polymerase chain reaction) based gene amplification approach was carried over to fish out an amplicon of ASAL1 by a set of primer 5'- GGA TTC ATG GGT CCT ACT ACT TCA TCT CCT-3'(Forward) and 5'-GGA TTC TCA AGC AGC ACC GGT GCC ACC CTT -3' (Reverse), using first strand cDNA as the template. A Primer set was designed based on reported-known gene sequence in GenBank database. The designed primers ensured that the coding sequence of the gene corresponding to the mature peptide start with an ATG (methionine) codon. PCR amplification was carried out with 20 µl PCR reaction mixture containing template cDNA (100 ng), primers (each 10 µM), buffers (5X), dNTPs (10 mM), MgCl, (25mM), Taq polymerase (5Units/µl) and distilled deionized water. The PCR profile consisted of initial denaturation at 94°C for 5 min; followed by 35 cycles each with denaturation at 94°C for 45 s, primer annealing at 60°C for 45 s, and primer extension at 72°C for 1 min in Thermal Cycler (Master cycler Gradient-eppendorf<sup>™</sup>). The final extension step was carried out at 72°C for 7 min. Amplified product was separated on 1.0% agarose gel along with Promega 100 bp DNA ladder. The desired amplified product was excised from the gel with sterile blade under UV-transilluminator and eluted from the gel using QIA quick Gel Extraction Kit (Qiagen). The purified product was further verified by sequencing using forward and reverse gene specific primers with ABI Sequencer 3730xl. The retrieved cDNA sequence was analyzed and submitted in GenBank database.

#### **Bioinformatics analysis of ASAL1**

The isolated gene sequence was In-silico analyzed using various online and offline Bioinformatics' tools.

#### Prediction of open reading frame

Open reading frame was Predicted using online tool, ORF finder (www.ncbi.nlm.nih.gov/projects/ gorf). The nucleotide and protein sequences of isolated gene sequence were compared with other plant lectins available at GenBank database using program BLASTn and BLASTp respectively.

#### Analysis of protein

The protein family, conserved domain as well as the carbohydrate specificity to which the identified protein bound was predicted from the protein sequence by the online tool known as pfam (pfam. Sanger.ac.uk/). The structure of the protein was modelled using homology modelling based approach with the help of Modeller 9v5 software. The various five different models generated, best model satisfying constraints of protein structure was selected on the basis of Ramchandran plot. For quality assessment of modeled proteins, these models were further analyzed through SAVES server (http://services.mbi.ucla.edu/SAVES/). The amino acid profile viz., molecular weight, percentage of each amino acid, positively and negatively charged residue, instability index, aliphatic index, theoratical PI and grand average of hydrophobicity (GRAVY) was predicted by Expasy Prot-Param tool (expasy. org/protparam/). Sub cellular localization of the protein was identified using Cello V 2.5 server (http//cello life.ncbi.edu.tw).

#### Sequence alignments and Phylogenetic analysis

Multiple sequence clustal alignment of the deduced amino acid sequence with other monocot mannose binding lectins of other families retrieved from GenBank was performed using ClustalW2-Phylogeny (http://www.ebi.ac.uk/Tools/phylogeny/ clustalw2\_phylogeny/). A phylogenetic tree was generated using neighbor-joining method Thompson *et al.* (1994) to study the evolutionary relationships between the candidate gene and other sequences of different mannose binding plant lectins.



#### Singh et al.

#### **Results and discussion**

#### Isolation and Sequence analysis of ASAL1 gene

Full length cDNA of garlic leaf lectin gene was amplified by RT-PCR. Very sharp amplicon of expected size of 540+bp (Figure 1) was obtained as earlier reported by Yarrasi et al. (2008). The gene has full-length cDNA of 540bp with 369bp ORF (including C terminal stop codon). The candidate gene sequence (ASAL1) was 30 nucleotides more and showed nine nucleotide changes than the previously reported garlic leaf lectin gene sequence of 339bp with accession number EU252577 (Figure 2). Five substitutions were observed in amino acid positions, V<sup>5</sup>, R<sup>83, 85</sup> and L<sup>97</sup> and E<sup>103</sup> (Figure 3). These changes were due to single base pair mutation. Most mutation in the DNA sequence during evolution does not result in a change in protein function. These amino acids differences are not likely to affect the binding ability of the mannose. Therefore, the mannosebinding motifs which are important for Allium sativum lectin are well protected Kai et al (2004). The gene sequence was deposited in the GenBank with accession number KJ767739 and was named as Allium sativum agglutinin lectin 1 (ASAL1). It has 50% A+T content and 50% G+C content. Sequence comparison by performing BLASTn search, revealed that ASAL1 had high homology (>90%) with many other mannose binding lectins, suggesting that ASAL1 belonged to the monocot mannose-binding lectin superfamily. ASAL1 gene sequence showed maximum (98%) identity with Allium sativum lectin mRNA complete cds having accession number DQ525625.1.

#### Amino acid sequence analysis

A BLASTp comparison of deduced amino acid sequence of ASAL1 gene with other mannosebinding lectins revealed high sequence similarity (50-93%) at amino acid level. ASAL1 showed maximum identity (98%) with Allium sativum lectin having accession number ABE70332 (nucleotide id is DQ525625.1). The ASAL1 gene codes for a polypeptide of 122 amino acids (Figure 4) with molecular weight (13kDa) consistent with the previously reported lectins from Alliaceae family Hossain et al. (2006) and Yarasi et al. (2008).

The peptide starts with Methionine (M) and ends with Tyrosine (Y). The richest amino acid deduced

from ASAL1 protein was Val (10.7%), followed by Gly (9.8%), Asn (9%), Arg (8.2%), Tyr (7.4%), Leu and Ser (6.6%), Asp, Ala and Thr (4.9%), Gln, Cys and Ile (4.1%), Glu (3.3%), Trp, Met and Lys (2.5%), Phe and Pro (1.6%) and His (0.8%). The number of buried and exposed amino acids were predicted through NetSurfP (http://www.cbs.dtu.dk/services/ NetSurfP/). In ASAL1, out of 122 amino acids, 53 (43.44%) were buried and 69 (56.55%) were exposed. Moreover nine amino acid residues (T<sup>7</sup>, N<sup>32, 71</sup>, S<sup>45, 78,</sup> <sup>80,102,106,113</sup>) were glycosylated in ASAL1 and both N linked and O linked glycosylation sites were found.

#### Analysis of ASAL1 protein

#### Protein family and stability

The ASAL1 protein belonged to the family B-Lectin, showed specificity towards D-mannose. A putative conserved domain (from amino acid 4-113) was detected in the deduced amino acid sequence using reserved position specific blast (RPS-BLAST), which predicted that ASAL1 gene is bulb type mannose binding lectin ( $\beta$ - lectin). The domains contained a three fold internal repeats or mannose binding sites (Figure 4) in which the consensus sequence motif QXDXNXVXY is involved in  $\alpha$ -D-mannose recognition hence, mannose-binding motifs are well protected in the ASAL1 gene sequence. The motif I was highly conserved and in motif II, similar to all monocot mannose binding lectins second residue was variable. In motif III, third residue valine was replaced by Lucine. The three mannose-binding motifs identified in candidate gene also agrees with the lectin genes, detected in garlic Afolabi-Balogun et al. (2012) and Van Damme et al. (1995), onion Hossain et al. (2006) and Van Damme et al. (1993) and Taxus media Kai et al. (2004). The conserved motifs were often conservative during evolution, while some variations on un-conserved domain can form the molecular foundation for the diversity of the structures and functions Kai et al. (2004).

The number of positive charged (Asp + Glu) and negatively charged (Arg + Lys) residues were 10 and 13 respectively. The ASAL1 protein is a stable protein as the instability index was 34.47. As the values of Grand average of hydrophobicity (GRAVY) were negative (-0.343) therefore, the protein is hydrophilic in nature. The isoelectric point value of the protein was 8.64. The lectin proteins are generally secretary in nature therefore, sub cellular localization of the



Fig. 1: Isolation of *Allium sativum* agglutinin leaf lectin cDNA [M = Promega 100 bp+ DNA ladder, NC = Negative control, Lane 1 = No amplification on 40ng/µl of garlic cDNA, Lane 2= Amplification on 80ng/µl of garlic cDNA, Lane 3= Amplification on 100ng/µl of garlic cDNA of expected size 550 bp, Lane 4= Negative control for Actin primers, Lane 5-7= Amplification of actin primers on garlic cDNA with 180 bp as amplicon size.]

EU252577	ATGGCCAGGAACCTACTGACGAACGGTGAAGGACTCTACGCAGGCCAATCACTGGATGTA	
KJ767739	ATGGCCAGGAACGTACTGACTAACGGTGAAGGACTCTACGCAGGCCAATCCCTGGATGTA	
	*********** ******* *******************	
EU252577	GAACAATACAAGTTTATAATGCAGGATGACTGCAACCTCGTGCTGTACGAATACAGCACC	
KJ767739	GAACAATACAAGTTTATAATGCAGGATGACTGCAACCTCGTGCTGTACGAATACAGCACC	ĺ
	***************************************	
511252577		
EU252577		
KJ/6//39		
	***************************************	
FI1252577		
K1767739		
	*****	
EU252577	GTAAGAGGGAACGGGAACTATATCCTGGTGCTTCAGAAGGACAGGAACGTTGTCATTTAC	
KJ767739	GTAAGAAGGAACCGGAACTATATCCTGGTGCTTCAAAAGGACAGGAACCTTGTCATTTAC	
	***** ***** ***************************	
EU252577	GGATCTGATATTTGGTCTACTGGTACCTACAGAAGATAG	
KJ767739	GGATCTGAAATTTGGTCCACTGGTACCTACAGAAGAAGTGTGGGGTGGAGCTTGTTGTCAT	
	******* ******* *****	
EU252577		
KJ767739	TGTTATTAA	

Fig. 2: Alignment of isolated complete cDNA sequence of garlic leaf lectin gene (KJ767739) with previously reported sequence (EU252577). The Asterisk mark (\*) represents bases showing identity and gapped space shows base pair changes.

EU252577 KJ767739	MARNLLTNGEGLYAGQSLDVEQYKFIMQDDCNLVLYEYSTPIWASNTGVTGKNGCRAVMQ MARNVLTNGEGLYAGQSLDVEQYKFIMQDDCNLVLYEYSTPIWASNTGVTGKNGCRAVMQ ****:	60 60
EU252577 KJ767739	RDGNFVVYDVNGRPVWASNSVRGNGNYILVLQKDRNVVIYGSDIWSTGTYRR RDGNFVVYDVNGRPVWASNSVRRNRNYILVLQKDRNLVIYGSEIWSTGTYRRSVGGACCH ***********************************	112 120
EU252577 KJ767739	 CY 122	

Fig. 3: Alignment of amino acid sequence of isolated complete cDNA sequence of garlic leaf lectin gene (KJ767739) with previously reported sequence (EU252577). The Asterisk mark (\*) represents bases showing identity and gapped space shows base pair changes. The alignment was performed with the online services of ClustalW2. The sign (:), shows amino acid changes.

66	at	<mark>qa</mark> g	aaa	cgt	att	ggt	gaa	caa	cga	agg	act	gta	cgc	agg	ccaa
	M	R	N	V	L	V	N	N	E	G	L	Y	A	G	Q
111	tc	cct	agt	cgt	aga	aca	igta	cac	ctt	tat	aat	gca	gga	tga	ctgc
	S	L	V	V	E	Q	Y	T	F	I	M	Q	D	D	С
156	aa	cct	tgt	act	gta	cga	ata	cag	cac	ССС	cat	ctg	ggc	ctc	aaac
	N	L	V	L	Y	E	Y	S	T	P	Ι	Ŵ	A	S	N
201	ac	ggg	cgt	cac	cgg	caa	aaa	cgg	gtg	cag	ggc	cgt	gat	gca	ggtt
	T	G	V	T	G	K	N	G	С	R	A	V	M	Q	V
246	ga	tgg	caa	ctt	tgt	ggt	cta	cga	tgt	taa	ggg	gcg	tgc	cgt	ctgg
	D	G	N	F	V	V	Y	D	V	K	G	R	A	V	W
291	gc	cag	taa	cag	cag	aag	agg	gaa	cgg	aaa	cta	cat	cct	ggt	gctt
	A	S	N	S	R	R	G	N	G	N	Y	Ι	ľ	V	L
336	Ca	gga	gga	cag	gaa	cgt	tgt	tat	tta	tgg	atc	tga	tat	ttg	gtct
	Q	E	D	R	N	V	V	Ι	Y	G	S	D	Ι	W	S
381	ac	tgg	tac	gta	cag	gaa	ia <u>tq</u>	<u>a</u> 4	01						
	T	G	T	Y	R	K	*								

Fig. 4: The full-length cDNA sequence (366 bp) and deduced amino acid sequence (122aa) of *Allium sativum* agglutinin leaf lectin (*ASAL1*). The start codon (atg) and stop codon (tga) underlined. The conserved mannosebinding sites (QDNVY) are showed in blocks. Asterisk mark (\*) represents that stop codon can not be translated into any amino acid.



Fig. 5: Three-dimensional structure of ASAL1.  $\beta$  sheets are indicated as arrows, turns and loops are indicated as lines and the amino acids constituting glycosylation are shown with respective residue number.

GNA4 ACA ASALI AAA APC ASA ASA ASL AAP DFA ACALI DPA ACL	HAKASLLILATIFLGVITPSCLSEHILYSGETLPTGGFLTSGSFVFIQQDCH HARNILLTHGEGLYAGGSLVEQYTFIQQDCH HARNILLTHGEGLYAGGSLDVEQYTFIQQDCH NGRTTPSPKLDYSITTYAAILITLASTOHARNILLTHGEGLYAGGSLDVEQYTFIQQDCH 
	·····
GHA3 GHA4 ACA ASAL1 AAA ASA ASA ASA2 ASA ASA2 ASL AAP DFA ACAL1 DPA ACL	LVL-NIVDKPINATIITGGLS-SDCSLSMQTDGILVV/TPS0KPINAS0TDGQNBIVVCIA LVL-NIUDKPINATIITGGLS-SDCSLSMQTDGILVV/TPS0KPINAS0TDGQNBIVVCIA LVL-NYSTPINASITGVTGKNGCRAMP.QbQBiFVV/TWNKBPVNASISKRBNBIVTLVUQ LVL-NYSTPINASITGVTGKNGCRAMP.QbQBiFVV/TWNKBPVNASISKRBNBIVTLVUQ LVL-NYSTPINASITGVTGKNGCRAMP.QbQBiFVV/TWNKBPVNASISKBNBINTLVUQ LVL-NYSTPINASITGVTGKNGCRAMP.QbQBiFVV/TWNKBPVNASISKBNBINTLVUQ LVL-NYSTPINASITGVTGKNGCRAMP.QbQBiFVV/TWNKBPVNASISKBNBINTLVUQ LVL-NYSTPINASITGVTGKNGCRAMP.QbQBiFVV/TWNKBPVNASISKBNBINTLVUQ LVL-NYSTPINASITGVTGKNGCRAMP.QbQBiFVV/TWNKBPVNASISKBNBINTLVUQ LVL-NYSTPINASITGVTGKNGCRAMP.QbQBiFVV/TWNKBPVNASISKBNBINTLVUQ LVL-NYSTPINASITGVTGKNGCRAMP.QbQBiFVV/TWNKBPVNASISKBNBINTLVUQ LVL-NYSTPINASITGVTGKNGCRAMP.QbQBiFVV/TWNKBPVNASISKBNBINTLVUQ LVL-NYSTPINASITGVTGKNGCRAMP.QbQBiFVV/TWNKBPVNASISKBNBINTLVUQ LVL-NYSTPINASITGVTGKNGCRAMP.QbQBiFVV/TWNKBPVNASISKBNBINTLVUQ LVL-NYSTPINASITGVTGKNGCRAMP.QbQBiFVV/TWNKBPVNASISKBNBINTLVUQ LVL-NYSTPINASITGVTGKNGCRAMP.QbQBIFVV/TWNKBPVNASISKBNBINTLVUQ LVL-NYSTPINASITGVTGKNGCRAMP.QbQBIFVV/TSNKBPVNASISKBNBINTLVUQ LVL-NYSTPINASITGVTGKNGCRAMP.QbQBIFVV/TSNKBPVNASISKBNBINTLVUQ LVL-NYSTPINASITGVGKNGCRAMP.QbQBIFVV/TSNKBPVNASISKBNBINTLVUQ LVL-NYSTPINASITGVGKNGCRAMP.QbQBIFVV/TSNKBPVNASISKBNBINTLVUQ LVL-NYSTPINASITGVGKNGCRAMP.QbQBIFVV/TSNKBPVNASISKBNBINTLVUQ LVL-NYSTPINASITGVGKNGCRAMP.QbQBIFVV/TSNKBPVNASISKBNBINTLVUQ LVL-NYSTPINASITGVGKNGCRAMP.QbQBIFVV/TSNKBPVNASISKBNBINTLVUQ LVL-NYSTPINASITGVGKNGCRAMP.QbQBIFVV/TSNKBPVNASISKBNBINTLVUQ LVL-NYSTPINASITGVGKNGCRAMP.QbQBIFVV/TSNKBPVNASISKBNBINTLVUQ LVL-NYSTPINASITGVGKNGCRAMP.QbQBIFVV7DVGKKANASISKBNBINTLVUQ LVL-NYSTPINASITGVGKNGCRAMP.QbQBIFVV7DVGKKANPASISKBNBINTLVUQ LVL-NYSTPINASITGVGKNGCRAMP.QbQBIFVV7DVGKKANPASISKBNBINTLVUQ LVL-NYSTPINASITGVGKNGCRAMP.QbQBIFVV7DVGKKANPASISKBNBINTLVUQ LVL-NYSTPINASITGVGKNGCRAMP.QbQBIFVV7DVGKKANPASISKBNBINTLVUQ LVL-NYSTPINASITGVGKNGCRAMP.QbQBIFVV7DVGKKANPASISKBNBINTLVQD LVL-NYSTPINASITGVGKNGCRAMP.QbQBIFVV7DVGKKANPASISKBNBINTLVQD LVL-NYSTPINASITGVGKNGCANPADQBIFNVCDVINGKANPASISKBNBINTLVQD LVL-NYSTPINASITGVGKNGCANPADQBIFNCZD
GHA3 GHA4 ACA ASAL1 AAA ASA ASA2 ASA2 ASA ASA2 ASL AAP DFA ACAL1 DPA ACL	KDRI/NVZYTTIRNATGTHTGAVGIPASPPSE
GHA3 GHA4 ACA ASAL1 AAA APC ASA ASA2 ASL ASA2 ASL ASA2 ASL DFA ACAL1 DPA ACL	- - - - - - - - - - -

Fig. 6: Multiple alignment of the amino acid sequence of *ASAL*1 with those of mannose binding lectins from other families. The alignment was performed with the online services of ClustalW2. Gaps were introduced for optimal alignment and maximal similarities between all compared sequences. The highly conserved amino acids and motifs were shown in Asterisk mark (\*).



Fig. 7: Phytograph of *ASAL*1 with other mannosebinding lectin gene sequences.

IJAEB

*ASAL1* protein was predicted with 2.376 cello score which showed that the protein is extra cellular in nature.

#### Homology modelling of proteins

Three dimensional structure of protein play a key role in protein function. Homology modelling requires a template structure to build the model for new proteins. The structure of ASAL1 contained 84% residues in core region, 11.3 % in allowed region and 4.7 residues in generously allowed region, no residue was in disallowed region (Table 1). ASAL1 showed a maximum similarity with the chain A of Mannose-Specific Agglutinin (Lectin) from Garlic (PDB id: 1KJ1). It has been observed from the tertiary structure of the ASAL1 protein that it possess, three fold internal repeat i.e.,  $\beta$ prism type-2 structure, as shown by most of the mannose binding lectins from the Alliaceae family Afolabi-Balogun et al. (2012) and Van Damme et al. (1995, 1993). The structure of ASAL1 protein (Figure 5) had 13 random coils and 12  $\beta$  strands but no helix. Molecular characterization of ASAL1 such as mannose-binding site analysis, secondary and three-dimensional structures analysis revealed that ASAL1 had many typical characters possessed by monocot mannose-binding lectins, implying that ASAL1 might have similar functions with many other mannose-binding lectins such as inhibition of fungal growth or insect feeding.

#### Sequence alignments and Phylogenetic analysis

The availability of numerous amino acid sequences encoding Mannose-binding lectins (MMBLs) from Alliaceae, Bromeliceae, Dioscoreaceae and Poaceae species enabled us to analyze the homology and molecular evolution of the different members of this superfamily of monocot Man-binding lectins. The online tool ClustalW2 was used for multiple sequence alignments (Figure 6) of the deduced amino acid sequence of ASAL1 with monocot mannose binding lectins of other families. The sequence alignment assured that conserved motifs containing internal repeats remain unchanged in all lectin families. A phylogenetic tree was constructed to study the evolutionary relationships between the candidate gene and other sequences of different mannose binding plant lectins. All the monocot mannose binding lectins were grouped in to three clusters (Figure 7). The first cluster comprised of ACA (Allium cepa; ABB54692), GNA3 (Galanthus nivalis; AAL07476), GNA4 (Galanthus nivalis; AAL07477), DPA (Dioscorea polystachya; BAD67184), ACL (Ananas comosus; AAM28277), DFA (Diplachne fusca; AD205812), AAP (Allium altaicum preproprotein; ADN26577), ACAL1 (Allium cepa N53; KJ767740) and ASAL1 (Allium sativum PG17; KJ767739) lectin gene sequences, which are from some taxonomically unrelated families such as Alliaceae, Dioscoreaceae, Bromeliaceae and Poaceae. The candidate gene ASAL1 was grouped in first cluster but surprisingly other lectin gene sequences from garlic viz. ASA2 and ASL used for phylogenetic relationship were fall in second cluster with unrelated families. Kai et al. (2004) also reported that the distribution of lectin genes is independent to the family to which these belong. The possible reason for this evolutionary behavior of MMBLs is supposed to be amino acid changes in Mannose binding sites Barre et al. (1996).

### Conclusion

All evidence suggests that the presently isolated lectin gene, ASAL1 is a distinct lectin gene from garlic. Lectins are evolving continuously in nature irrespective of their family and show independent distribution in phylogenetic system. Since like all other mannose binding lectins, the presently isolated lectin gene (ASAL1) contains three fold internal repeat (QXDNXVXY) in mannose binding domain, which play a key role in mannose binding and agglutination. Since the agglutinating properties of lectin are responsible for insecticidal properties of lectin, hence ASAL1 would have the ability to induce insecticidal effect if successfully transformed into other plants. Hence cloning, characterization and expression of ASAL1 in transgenic plants would enable us to study its potential insect resistance function in the future.

#### Acknowledgements

The authors are highly acknowledged to Dr. Kuldeep Singh, Director, School of Agricultural Biotechnology for providing facilities for sequencing, Dr. A.S. Dhatt for providing the seed of garlic variety PG17 and my Advisor Dr. N.K. Sarao (Assistant plant Breeder) for her judicious guidance and support. Funding support by School of Agricultural Biotechnology is also gratefully acknowledged.



#### References

- Afolabi-Balogun, N.B., Inuwa, H.M., Ishiyaku, M.F., Bakare-odunoola, M.T., Nok, A.J. 2012. Isolation and characterization of a mannose-binding insecticidal lectin gene from *Allium sativum* (Garlic) and its putative role in insect resistance using bioinformatics tools. *Infection, Genetics and Evolution* **12**: 1508-1512.
- Bandyopadhyay, S., Roy, A. and Das, S. 2001. Binding of garlic (*Allium sativum*) leaf lectin to the gut receptors of Homopteran pests is correlated to its insecticidal activity. *Plant Science* **161**: 1025-1033.
- Banerjee, S., Hess, D., Majumder, P., Roy, D. and Das, S. 2004. The interactions of Allium sativum leaf agglutinin with a chaperonin group of unique receptor protein isolated from a bacterial endosymbiont of the mustard aphid. *Journal of Biological Chemistry* **279**: 23782- 23789.
- Barre, A., Van Damme, E.J.M., Peumans, W.J. and Rouge, P. 1996. Structure-function relationship of monocot mannose-binding lectins. *Plant Physiology* **112**: 1531-1540.
- Barre, A., Bourne, Y., Van Damme, E.J.M., Peumans, W.J., Rouge, P. 2001. Mannose binding plant lectins: different structural scaffolds for a common suger-recognition process. *Biochimie* 83: 645-651.
- Down, R.E., Ford, L., Woodhouse, S.D., Majerus, M.E.N., Gatehouse, J.A. and Gatehouse, A.M.R. 2003. Tritrophic interactions between transgenic Potato expressing snowdrop lectin. *Transgenic Research* **12**: 229-234.
- Dutta, I., Saha, P., Majumder, P., Sarkar, A., Chakraborti, D., Banerjee, S. and Das, S. 2005. The efficacy of a novel insecticidal protein Allium sativum leaf lectin (ASAL) against Homopteran insects monitored in transgenic tobacco. *Plant Biotechnology Journal* **3**: 601-611.
- Eisemann, C.H., Donaldson, R.A., Pearson, R.D., Cadagon, L.C., Vuocolo, T., Pellam, R.L. 1994. Larvicidal action of lectins on Luciliacuprina; mechanism of action. *Entomologia Experimentalis et Applicata* **72**: 1-11.
- Ferry, N., Edwards, M.G., Mulligan, E.A., Emami, K., Petrova, A.S., Frantescu, M., Davison, G.M. and Gatehouse, A.M.R. 2004. Engineering resistance to insect pests. In: P Christou and H Klee (Eds) Hand Book of Plant Biotechnology. *John Wiley & Sons Ltd. West Sussex UK* 1: 373-394
- Fitches, E., Wiles, D., Douglas, A.E., Hinchliffe, G., Audsley, N. and Gatehouse, J.A. 2008. The insecticidal activity of recombinant garlic lectins towards aphids. *Insect Biochemistry and Molecular Biology* **38**: 905-915.
- Foissac, X.N.T., Loc, P., Christou, A., Gatehouse, A.M.R. and Gatehouse, J.A. 2000. Resistance to green leafhopper (Nephotettix virecens) and brown plant hopper (*Nilapravata lugens*) in transgenic rice expressing snowdrop lectin (*Galanthus nivalis* agglutinin; GNA). *Journal of Insect Physiology* **46**: 573-583.
- Gatehouse, A.M.R., Down, R.E., Powell, K.S., Sauvion, N., Rahbe, Y., Newell, C.A., Merry weather, A., Hamilton, W.D.O. and Gatehouse, J.A. 1996. Transgenic potato plants with enhanced resistance to the peach-potato aphid, Mizuspersicae. *Entomologia Experimentalis et Applicata* **79**: 295-307.

- Hossain, M.A., Maiti, M.A., Basu, A., Sen, S., Gosh, A.K. and Sen, S.K. 2006. Transgenic expression of onion leaf lectin gene in Indian mustard offers protection against aphid colonization. *Crop Science* **46**: 2022-2032.
- Kai, G., Zhao, L., Zheng, J., Zhang, L., Miao, Z., Sun, X. and Tang, K. 2004. Isolation and characterization of a new mannose-binding lectin gene from *Taxus media*. *Journal* of *Biosciences* 29 (4): 399-407.
- Li, R., Gao, W., Ji, D. and Li, C. 2000. Isolation of agglutinin from Arisaemacons arguieum and analysis on its resistance to *Aphis gossypii*. *Acta* **12**: 54-56.
- Loris, R. 2002. Principles of structures of animal and plant lectins. Biochim Biophys *Acta*. **1572**:198–208.
- Majumder, P., Banerjee, S. and Das, S. 2004. Identification of receptors responsible for binding of the mannose specific lectin to the gut epithelial membrane of the target insects. *Glycoconjugate Journal* **20**: 525-530.
- Mao, X., Gao, L., Li, C. and Li, R. 1999. Antibiological effects of several phyto hemagglutinins on *Aphis gossypii*. *Journal of Shanxi Agricultural University* **19**: 122-124.
- Nagadhara, D., Ramesh, S., Pasalu, I.C., Kondala, Y., Rao, N.P., Sarma, V.D. and Reddy, K.V. 2004. Transgenic rice plants expressing the snowdrop lectin gene (gna) exhibit high- level resistance to the white backed planthopper (*Sogatella furcifera*). *Theoretical and Applied Genetics* **109**: 1399-1405.
- Neha, G., Pankaj, K., Jitender, S., Anchal, R., Anil, S. and Prasann, K. 2014. Cloning and molecular characterization of LECASAI lectin gene from garlic (Allum sativum L.). *International Journal of Agriculture, Environment and Biotechnology* 7: 1-10.
- Pusztai, A. 2008. Plant Lectins. Cambridge: Cambridge University Press.
- Rao, K.V., Keerti, S., Rathore, I., Thomas, K., Hodges, X., Stoger, F.E., Sudhakar, D., Williams, S., Christou, P., Bharathi, M., David, P.B., Kevin, S.P., Jacqui, S., Angharad. M.R.G. and John, A.G. 1998. Expression of snowdrop lectin (gna) in transgenic rice plants confers resistance to rice brown plant hopper. *Plant Journal* 15: 469-477.
- Sharon, N. 2007. Lectins: Carbohydrate-specific reagents and biological recognition molecules. *Journal of Biological Chemistry* 282: 2753–2764.
- Stoger, E., William, S., Christou, P., Down, R.E. and Gatehouse, J.A. 1999. Expression of the insecticidal lectin from snowdrop (*Galanthus nivalis* agglutinin; GNA) in transgenic wheat plants: effects on predation by the grain aphid Sitobion avenae. *Molecular Breeding* 5: 65-73.
- Sun, X., Wu, A., Tang, K. 2002. Transgenic rice lines with enhanced resistance to the small brown planthopper. *Crop Protection* **21**: 511-514.
- Tang, K., Zhao, E., Sun, X., Wan, B., Qi, H. and Lu, X. 2001. Production of transgenic rice homozygous lines with enhanced resistance to the rice brown planthopper. *Acta Biotechnology* 21: 117-128.

IJAEB

- Thompson, J.D., Higgins, D.G., Gibson, T.J. and Wolfinger, R.D. 1994. Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**: 4673-4680.
- Van Damme, E.J.M., Anthony, K.A., Peumans, W.J. 1987. Isolation and characterization of lectins from bulbs. *FEBS Letters* 215: 140-144.
- Van Damme, E.J.M., Smeets, K., Torrekens, S., Van-Leuven, F., Peumans and W.J. 1993. The mannose-specific lectins from ramsons (*Allium ursinum* L.) are encoded by three sets of genes. *European Journal of Biochemistry* 217: 123-129.
- Van Damme, E.J.M., Goossens, K., Smeets, K., Van-Leuven, F., Verhaert, P. and Peumans, W.J. 1995. Characterization and molecular cloning of the lectin from *Arum maculatum*. *Plant Physiology* **107**: 1147-1158.
- Van Damme, E.J.M., Peumans, W.J., Pusztai, A. and Bardocz, S. 1998. Handbook of plant lectins: properties and biomedical applications. Bognor Regis, UK: John Wiley and Sons.

- Van Damme, E.J., Nakamura-Tsuruta, S., Smith, D.F., Ongenaert, M., Winter, H.C., Rouge, P., Goldstein, I., Hanqing, M.O., Kominami, J. and Culerrier, R. 2007. Phylogenetic and specificity studies of two-domain GNA-related lectins: generation of multi-specificity through domain duplication and divergent evolution. *Biochemistry Journal* 404:51–61.
- Yao, J., Zhao, X., Zhang, H., Liao, Z., Chen, F., Song, J., Sun, X. and Tang, K. 2003. Molecular cloning and characterization of a novel lectin gene from Pinellia ternata. *Cell Research* 13: 301-308
- Yarasi, B., Sadumpati, V., China, P.I., Dasavantha, R.V. and Venkateswara, R.K. 2008. Transgenic rice expressing *Allium sativum* leaf agglutinin (ASAL) exhibits highlevel resistance against major sap-sucking pests. *BMC Plant Biology* **8**: 102-107.