The Andhra Agric. J 64 (2): 391-397, 2017

Cloning of Defense Related Gene ARID/ BRIGHT Against Fusarium Wilt in Chickpea

G NagaRaju and Subhra Chakraborty
National Institute of Plant Genome Research, J.N.U. Campus, NewDelhi-110 067, India

ABSTRACT

The regulation of gene expression in response to internal and environmental cues occurs at many levels in the plant and animal cell. Myriad transcription factors, particularly those which function *in trans*, control the efficiency with which the transcription apparatus is assembled and the rate at which transcripts are produced from a genetic locus. Transcription factors (TFs) are conventionally defined by their ability to bind specific DNA sequences and regulate transcription. The ARID (AT-rich interaction domain) is a billion year old DNA-binding domain that has been identified in all sequenced higher eukaryotic genomes. The ARID consensus sequence spans about 100 amino acid residues, and structural studies identify the major groove contact site as a modified helix-turn-helix motif. In green plants 187 ARID genes are identified out of which 13 are present in Arabidopsis and 21 in Rice, one each in Lotus and Medicago etc. ARID containing genes are also present in vitis, ricinus, glycine, maize, sorghum, populous and barley. In the present study, the transcript profiling during *Fusarium* wilt in chickpea led to the identification of an ARID transcription factor. The gene was found to be differentially expressed and it was of great interest to clone the gene and study its role in immune response. Here we present the expression study of the *CaAB* transcription factor in response to *Fusarium.* We also demonstrate its tissue specific expression, copy number and subcellular localization.

Key words: *ARID/BRIGHT, Fusarium wilt and Chickpea*

Transcription factors (TFs) are conventionally defined by their ability to bind specific DNA sequences and regulate transcription (Verk *et. al.,* 2009). They can perform these functions in many ways. The repertory of known DNA-binding domains, interactions with other TFs and chromatin proteins, and means of influencing transcription continue to grow. TFs have often identified as key metabolic or developmental regulators, and at least initially provided an easily understood scheme for the orchestration of gene expression, cell differentiation, and homeostasis. Genome sequencing and genomic analyses, however, have abundantly confirmed early suspicions that cells interpret genomic information by mechanisms that are both complicated and varied. The ever-increasing amount of data presents new challenges and opens new horizons in the analysis of TFs and their functions. Eukaryotic genomes display remarkable diversity in their transcription factor (TF) repertoires, in terms of both presence and prevalence of different TF families in different lineages. It is estimated that TFs constitute between 0.5 and 8% of the gene content of eukaryotic genomes, with both the absolute number and proportion of TFs in a genome roughly scaling with the complexity of the organism. Most eukaryotic TFs tend to recognize short, degenerate DNA sequence motifs, in contrast to the larger motifs preferred by prokaryotic TFs. Arabidopsis TFs are characterized by a large number of genes and by the variety of gene families when compared with those of *D. melanogaster* or *C. elegans* . Around half of Arabidopsis TFs are plant specific and possess DBDs found only in plants. AP2-ERF, NAC, Dof, YABBY, WRKY, GARP, TCP, SBP, ABI3- VP1 (B3), EIL and LFY are plant- specific TFs. TFs act as transcriptional activators or repressors. In common with other eukaryotes, TFs containing domains rich in the acidic amino acids glutamine or proline, such as TOC1, DREBs, ARFs and GBF1, are transcriptional activators. The following Table 1.1 summarizes the transcription factor types involved in defense against pathogens.

ARID- newly discovered Transcription factor family in plants

The ARID (AT-rich interaction domain) is a billion year old DNA-binding domain that has been identified in all sequenced higher eukaryotic genomes (Wilsker et al, 2002). The ARID consensus sequence spans about 100 amino acid residues, and structural studies identify the major groove contact site as a modified helix-turn-helix motif (Iwahara and Clubb, 1999). The ARID consensus was first identified in the mouse B-cell-specific transcription factor bright and is the product of the dead ringer (dri, also known as retained, retn) gene of Drosophila melanogaster (Kortschak *et. al.,* 2000) DRI and Bright were each isolated in searches designed specifically to identify proteins binding to AT-rich sequences, but neither turned out to contain a previously known DNA-binding domain (Webb, 2001) . Identification of DNA binding sequences conserved between Bright and DRI defined the parameters of a new DNA-binding domain, whose name was inspired by the interaction of these proteins with AT-rich DNA elements. Since the discovery of the ARID, many additional proteins containing this domain have been identified. The cellular functions of ARID proteins include participation in the regulation of cell growth, differentiation, and development. The ARID domain is both ancient and widespread, occurring in (some) protozoa, green algae, higher plants, fungi, and metazoans. In green plants 187 ARID genes are identified out of which 13 are present in Arabidopsis and 21 in Rice, one each in Lotus and Medicago (Zhu et. al., 2008). ARID containing genes are also present in vitis, ricinus, glycine, maize, sorghum, populous and barley. Here for the first time we report cloning of ARID gene from Chick plant against Fusarium wilt.

MATERIAL AND METHODS

Cloning of *CaAB* gene:

Cicer arietinum ARIDBRIGHT transcription factor was cloned following 3Prime and 5Prime RACE protocol.

Rapid amplification of cDNA ends (RACE)

This is a procedure for amplification of nucleic acid sequences from a messenger RNA template between defined internal site and either the 32 or the 52 end of the mRNA. The entire procedure consists of following steps

First-Strand cDNA Synthesis

1. For each reaction (including the Control Human Placental Poly A+ RNA), combine the following in a sterile 0.5 ml microcentrifuge tube:
1 ìg (1–4 ìl) RNA sample (poly A+)
1 ìl cDNA Synthesis Primer (10 ìM)
2. Add sterile H2O to a final volume of 5 ìl.
3. Mix contents and spin the tube briefly in a microcentrifuge.
4. Incubate the tube at 70°C for 2 min.
5. Cool the tube on ice for 2 min.
6. Spin the tube briefly to collect the contents at the bottom.
7. Add the following to each reaction tube:
2 ìl 5X First-Strand Buffer
1 ìl dNTP Mix (10 mM)
1 ìl [á-32P]dCTP (1 ìCi/il)*
1 ìl AMV Reverse Transcriptase (20 units/ìl)
10 ìl Total volume
* [á-32P]dCTP is optional.
8. Mix the contents of the tube by gently pipetting.
9. Spin the tube briefly to collect the contents at the bottom.
10. Incubate the tube at 42°C for 1 hr in an air incubator.
11. Place the tube on ice to terminate first-strand synthesis.
12. Proceed directly to second-strand synthesis.

Second-Strand cDNA Synthesis

Note: All components and reaction vessels should be prechilled on ice.
1. For each reaction (including the positive control), combine the following components in the reaction tube from Step A.11:
(10 ìl First-strand reaction)
48.4 ìl Sterile H2O
16 ìl 5X Second-Strand Buffer
1.6 ìl dNTP Mix (10 mM)
4 ìl 20X Second-Strand Enzyme Cocktail
80 ìl Total volume
2. Mix contents thoroughly with gentle pipetting.
3. Spin the tube briefly to collect the contents at the bottom.
4. Incubate the tube at 16°C for 1.5 hr.
5. Add 2 ìl (6 units) of T4 DNA Polymerase and mix thoroughly with gentle pipetting.
6. Incubate the tube at 16°C for 45 min.
7. Add 4 ìl of the EDTA/glycogen Mix to terminate second-strand synthesis.

Table 1.1Ttranscription factor types involved in defense against pathogens.

Transcription factor type	Size of Arabidopsis family	Key features	Consensus core motif of binding sites
ERF	56	One ERF-DNA binding domain	GCCGCC (GCC box)
R2R3 Myb	125	Two repeats of Myb domain (R2 and R3)	Type I: (T/C)AAC(T/G)G Type II: G(G/T)T(A/T)G(G/T)T
TGA bZIP	10	One basic DNA binding domain; leucine zipper protein dimerization motif Ankyrin repeat domain; BTB-POZ domain	TGACGTCA (TGA box), this motif usually occurs as direct repeats (e.g. in as1-like elements)
NPR1	6	Whirly domain	No DNA binding sites
Whirly	3	One or two WRKY DNA	GTCAAAAA/T
WRKY	74	binding domains	(T)GACC/T (W box)

8. Add 100 ìl of phenol:chloroform:isoamyl alcohol (25:24:1).
9. Vortex thoroughly.
10. Spin the tube in a microcentrifuge at 14,000 rpm for 10 min to separate phases.
11. Carefully transfer the top aqueous layer to a clean 0.5 ml microcentrifuge tube Discard the interface and lower phase.
12. Add 100 ìl of chloroform:isoamyl alcohol (24:1) to the aqueous layer and vortex thoroughly.
13. Spin the tube in a microcentrifuge at 14,000 rpm for 10 min to separate phases.
14. Remove the top aqueous layer and place in a clean 0.5 ml microcentrifuge tube.
15. Add one-half volume of 4 M ammonium acetate (e.g., if you recovered H"70 ìl at Step 14, add 35 ìl of 4 M ammonium acetate).
16. Add 2.5 volumes of room temperature 95% ethanol. (e.g., if your volume at Step
15 was H"105 ìl, add
17. 263 ìl of 95% ethanol.)
18. Vortex the mixture thoroughly.
19. Spin the tube immediately in a microcentrifuge at 14,000 rpm at room temperature for 20 min.
20. Remove the supernatant carefully.
21. Gently overlay the pellet with 300 ìl of 80% ethanol.
22. Spin in a microcentrifuge at 14,000 rpm for 10 min.
23. Carefully remove the supernatant.
24. Air dry the pellet for approximately 10 min to evaporate residual ethanol.
25. Dissolve the precipitate in 10 ìl of H2O and store at –20°C.

C. Adaptor Ligation

Note: Allow 5X DNA Ligation Buffer to completely thaw at room temperature and keep it at room temperature for 30 min before use.
1. For each reaction (including the positive control), combine the following reagents in a 0.5 ml microcentrifuge test tube at room temperature and in the order shown:
5 ìl ds cDNA
2 ìl Marathon cDNA Adaptor (10 ìM)
2 ìl 5X DNA Ligation Buffer
1 ìl T4 DNA Ligase (1 unit/ìl)
10 ìl Total volume
2. Mix by vortexing and spin briefly in a microcentrifuge.
3. Incubate at either:
• 16°C overnight; or
• room temperature (19–23°C) for 3–4 hr.
4. Heat at 70°C for 5 min to inactivate the ligase.
5. Using the following guidelines, dilute your adaptor-ligated ds cDNA to a concentration which is suitable for subsequent RACE PCR procedures (H"0.1ig/ml).
6. Dilute 1 ìl of the positive control reaction mixture with 250 ìl of Tricine-EDTA Buffer.
7. Store the undiluted adaptor-ligated cDNA at –20°C for future use.
8. Heat the diluted ds cDNA at 94°C for 2 min to denature the ds cDNA.

9. Cool the tube on ice for 2 min.
10. Briefly spin the tube in a microcentrifuge to collect the contents in the bottom of the tube. Store at –20°C until ready for RACE PCR. At this stage, you essentially have a library of adaptor-ligated ds cDNA. The RACE reactions in Section E use only a fraction of this material for each RNA of interest.

Rapid Amplification of cDNA Ends (RACE)

1. Prepare enough PCR master mix for all of the PCR reactions plus one additional tube. The same master mix can be used for both 5'- and 3'-RACE reactions. For each 50 ìl reaction, mix the following reagents:
36 ìl H2O
5 ìl 10X cDNA PCR Reaction Buffer
1 ìl dNTP Mix (10 mM)
1 ìl Advantage 2 Polymerase Mix (50X)
43 ìl Final volume
Mix well by vortexing (without introducing bubbles) and briefly spin the tube in a microcentrifuge.

2. For 5'-RACE: To the master mix 1.0 ìl each AP1 and AB-5-1 were added along with 5.0ul of Adaptor ligated cDNA
3. For 3'-RACE: To the master mix 1.0 ìl each AP1 and AB-3-1 were added along with 5.0 ìl of Adaptor ligated cDNA
Both the reaction mix was subjected to PCR amplification as suggested in the manual

Cloning of DNA fragments

The fragments obtained as a result of PCR, were either cleaned with PCR purification kit and ligated to pGEM-T Easy vector or electrophoresed on the 1.0% agarose/EtBr gel, eluted from the agarose gel and ligated to the T-vector.

Ligation

Most of the PCR amplified fragments were cloned in pGEM-T Easy vector
1. Vector, ligation buffer were thawed on the ice and contents centrifuged briefly.
2. In a 0.5 ml tube (known to have low-DNA binding capacity) ligation was set by adding 5 ìl 2X

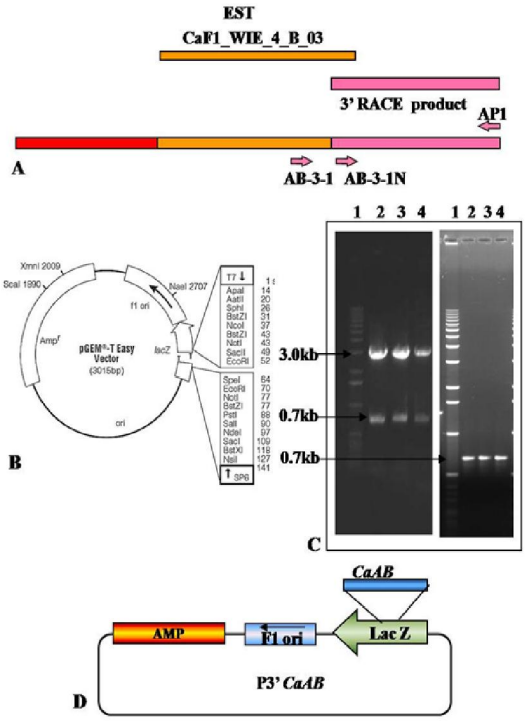

Figure 1.1. 3' RACE of *Ca*AB. (A) Flow diagram depicting 3' RACE procedure. Oligo dT adapter primer (AP1) was used to bind the polyA tail of the mRNA. Reverse transcriptase generates single stranded cDNA using mRNA as template. Gene specific primers (AB-3-1 , AB-3-1N) and AP1 primer specific to 3' adapter were used to amplify the target 3'cDNA end sequence; (B) Map of pGEM-T Easy vector used to clone the amplified product; (C) 1% agarose/EtBr gels confirming 3' RACE product of *Ca*AB by NotI digestion and colony PCR. Lane 1 represents 1Kb ladder and lane 2,3,4 three representative clones; (D) Diagrammatic representation of 3' RACE product of *Ca*AB cloned in pGEM-T Easy vector and the construct named as p3' *Ca*AB

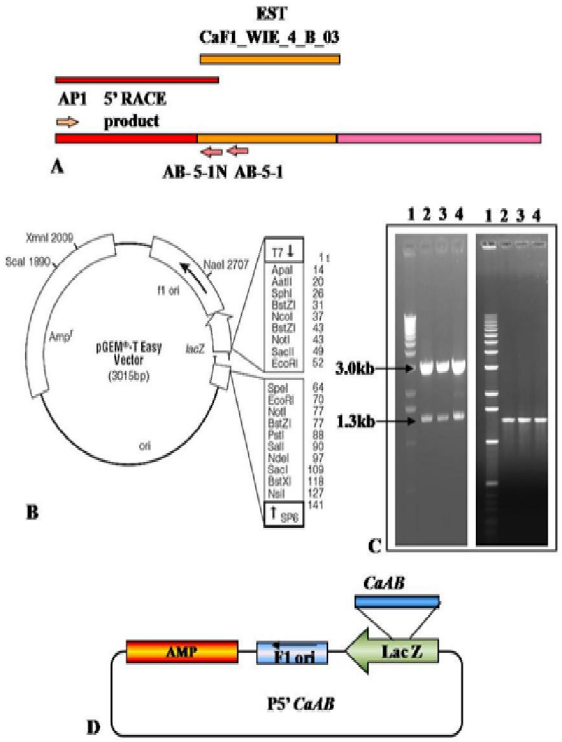

Figure 1.2. 5' RACE of *Ca*AB. (A) Flow diagram depicting 5' RACE procedure. AP1 was used to bind the 5'end of cDNA. Gene specific primers (AB-5-1 , AB-5-1N) and AP1 primer specific to 5' adapter were used to amplify the target 5'cDNA end sequence; (B) Map of pGEM-T Easy vector used to clone the amplified product; (C) 1% agarose/EtBr gels confirming 5' RACE product of *Ca*AB by NotI digestion and colony PCR. Lane 1 represents 1Kb ladder and lane 2,3,4 three representative clones; (D) Diagrammatic representation of 5' RACE product of *Ca*AB cloned in pGEM-T Easy vector and the construct named as p5' *Ca*AB.

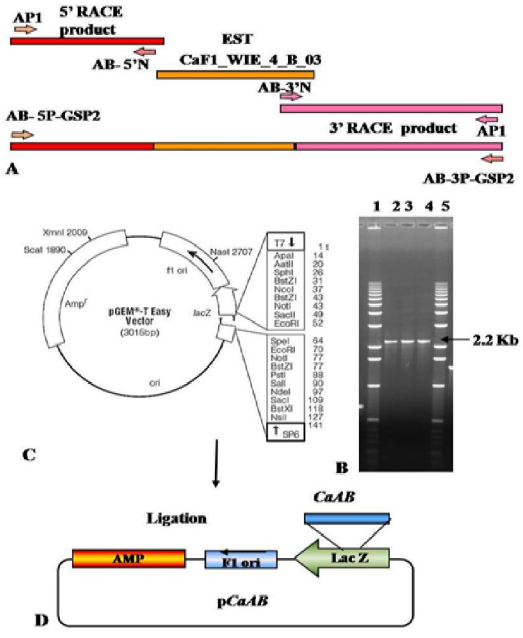

Figure 1.3. Isolation of full length cDNA clone of *CaAB*. (A) Schematic representation of the alignment of sequence of partial clone (EST) and that of 3'and 5'RACE products and the primer position for full length cloning; (B) 1% agarose/EtBr gel showing full length cDNA clone of *CaAB*. Lane 1,5 indicates 1Kb ladder and lane 2,3,4 PCR product; (C) Map of pGEM-T Easy vector used for cloning the PCR product; (D) Diagrammatic representation of full length clone of *CaAB* cloned in pGEM-T Easy vector and named as p*CaAB*.

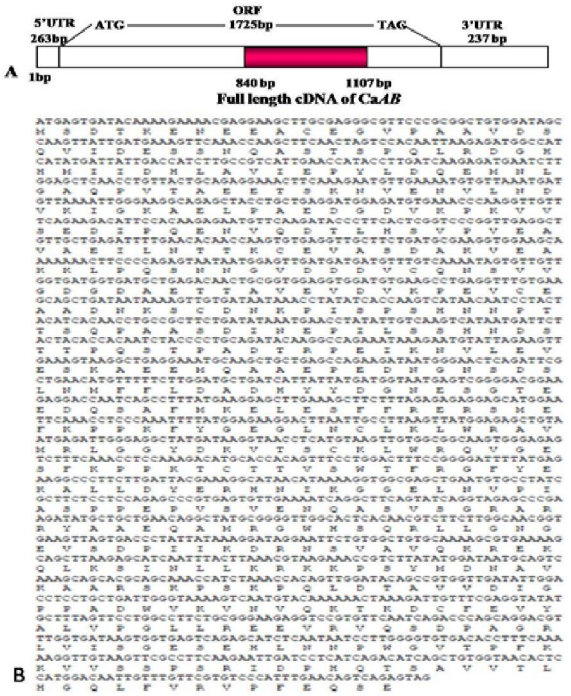

Figure 1.4. cDNA and deduced amino acid sequence of *CaAB*. (A) Diagrammatic representation of gene structure of *CaAB* cDNA; (B) Nucleotide and the amino acid sequence.

ligation buffer, 1 ìl of pGEM-T Easy Vector (50 ng) and PCR product in a molar ratio of 1:3 (Vector: Insert)

3. Then 1 ìl of T4 DNA ligase was added, the contents were mixed and incubated at 4ºC overnight. Next day 1- 2 ìl of the ligation mix was transformed using *E.coli* competent cells.

Preparation of Competent Bacterial Cells

For cloning purpose, *E. coli* DH5á bacterial strain were made competent by the Calcium Chloride method and competent Ecoli cells were transformed using standard procedure.

Presence of the insert

The presence of the insert in the clone was confirmed by the colony PCR by using either gene specific primers or primers compatible with cloning vector.

DNA Purification System

1. 1-5.0 ml (high-copy-number plasmid) or 10ml (low-copy-number plasmid) of bacterial culture was harvested by centrifugation for 5 minutes at 10,000 x g in a tabletop centrifuge. The supernatant was discarded and the tubes were inverted on a paper towel to remove excess media.

2. 250ìl of Cell Resuspension Solution was added and the cells were completely resuspended by vortexing or pipetting. It is essential to thoroughly resuspend the cells and transfer them to 1.5 ml microcentrifuge tube.

3. After resuspension 250ìl of Cell Lysis Solution was added and mixed by inverting the tube 4 times (do not vortex). Then the cells were incubated until the cell suspension become clear. (approximately 1-5 minutes).

4. 10ìl of Alkaline Protease Solution was added and mixed by inverting the tube 4 times and incubated for 5 minutes at room temperature. Alkaline protease inactivates endonucleases and other proteins released during the lysis of the bacterial cells that can adversely affect the quality of the isolated DNA.

5. Then 350ìl of Neutralization Solution was added and immediately mixed by inverting the tube 4 times.

6. The bacterial lysate was centrifuged at maximum speed (around 14,000 × g) in a microcentrifuge for 10 minutes at room temperature. Meantime the plasmid DNA purification unit was assembled by inserting one Spin Column into one 2ml Collection Tube for each sample. Then the cleared lysate (approximately 850ìl,) was transferred to the prepared Spin Column by decanting. Further, the supernatant was centrifuged at maximum speed in a microcentrifuge for 1 minute at room temperature.

And the flow through was discarded from the Collection Tube.

7. The Spin column was reinserted into the collection tube and 750ìl of columnwash solution was added and centrifuged at maximum speed for 1 minute at room temperature. The wash step was repeated using 250ìl of Column wash Solution.

8. Then the spin column was transferred to a new, sterile 1.5ml microcentrifuge tube and the plasmid DNA was eluted by adding 30ìl of nuclease-free water to the Spin Column and centrifuged at maximum speed for 1 minute at room temperature in a microcentrifuge. Further the eluted DNA was stored at -20 ºC.

9. The plasmids isolated were checked on the 0.8% agarose gel and were sequenced after confirming by Not1 digestion.

RESULTS AND DISCUSSION

Cloning of full length cDNA of *CaAB* gene

Full length cDNA of *CaAB* gene was amplified by 5' and 3' RACE using RACE kit (Clonetech Marathon Race) and following manufacturer's instructions. For 3' RACE, cDNA was synthesized using oligo dT primer from the kit (AP1 primer). Using this cDNA as a template, 3' end of *CaAB* gene was amplified using AP1 primer of the kit and the gene specific primer AB3-1 designed from the 3prime end sequence of the EST (CaF1_WIE_04_B_03). The product of primary amplification was subjected to secondary pcr using primer combination AP1 and nested 3prime gene specific primer AB3-1N. The amplified product of approximately 0.7 kb was run on the gel, eluted and subsequently cloned in pGEM-T Easy vector as shown in Figure1.1

Three representative clones which are colony pcr and digestion positive (Figure 1.1) were later on sequenced by standard procedures of sequencing and confirmed by the presence of overlap with the sequence of partial clone.

For 5' RACE two gene specific primers were designed namely AB5-1, and AB5-1N from the 5prime end sequence of the EST (CaF1_WIE_04_B_03). Using adapter ligated cDNA as a template primary pcr was carried out with the primer combinations AP1 and AB5-1.The primary PCR product was diluted (1:100) and 2ìl was taken for the secondary PCR using the primer combination AP1 and nested gene specific primer AB5-1N. The amplified product of ~1.3kb was eluted from gel and cloned in pGEM-T Easy vector(Figure1.2).

Three representative clones which are colony pcr and digestion positive (Figure 1.2) were later on sequenced by standard procedures of sequencing and confirmed by the presence of overlap with the sequence of partial clone. Two primers (AB_3PGSP-2, AB5GSP-2) were designed based on the sequencing information of 3prime and 5prime RACE products and were used to amplify the full length cDNA using adaptor ligated cDNA as a template and this resulted in ~2.2kb band which was subsequently eluted from the gel and cloned in pGEM-T Easy vector. Three representative clones which are colony pcr positive (Figure 1.3) were later on sequenced by standard procedures of sequencing and confirmed by matching the sequences with that of partial clone and RACE products. Both the strands of three positive clones were sequenced and a consensus sequence was obtained after aligning all the sequences using Bio edit program and was subsequently used for Insilco analysis. The sequence so obtained was named as CaAB full length cDNA sequence and was analyzed using NCBI ORF finder and the result showed that it had a largest ORF of 1725 bp, with 5prime un

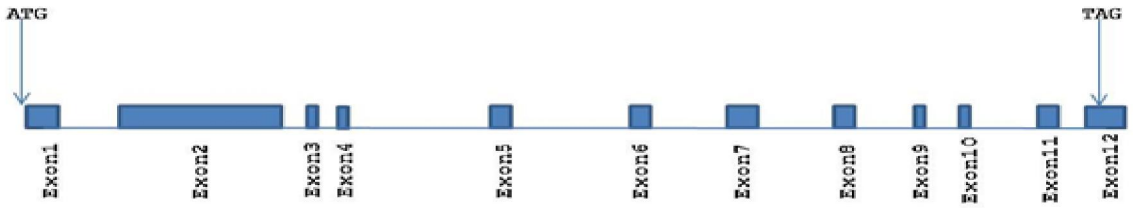

Figure1.5 Genomic Organization of C*aAB*

translated region(UTR) of 263 bp and 3prime un translated region(UTR) of 237bp and with a total length of 2225bp. The translated ORF of CaAB gene showed 574 amino acids (Figure 1.4), with a predicted molecular weight of 60.0kDa. Blast P analysis done at NCBI showed that it had three important domains namely, ARID (AT Rich Interactive Domain)/Bright, Putative Dimerisation Domain and an Alpha Crystalline Domain (ACD) of HSP20 family.

Genomic cloning of *CaAB* gene

Genomic cloning of CaAB gene using end to end primers (AB_3PGSP-2, AB5GSP-2) on WR-315 genomic DNA generated a single product with an approximate size of 6.2kb. Further cloning , sequencing, and alignment with full length cDNA showed that it had 12 Exons with 11 Introns (Figure 1..5).

There is one more intron in 5prime UTR region. Since the amplified product was 6.2kb, primer walking strategy was used to get the full length sequence of genomic clone and three representative clones after confirming with digestion were sequenced on both the strands and a consensus sequence was obtained using BioEdit program Alignment of genomic sequence with cDNA sequence of *CaAB* gene was done using BLAST2 Program available at NCBI site. Analysis of ARID gene of model plant *Arabidopsis thaliana* also showed multiple introns, for example the gene AT1G76510.2 had 14 exons and 13 introns with both ARID and ACD domains.

LITERATURE CITED

Iwahara J and Clubb R T 1999 Solution structure of the DNA binding domain from Dead ringer, a sequence specific AT-rich interaction domain (ARID).*EMBO J* 18: 6084–6094.

Kortschak R D, Tucker PW and Saint R 2000 ARID proteins come in from the desert. Trends Biochem Sci 25*:* 294–299.

Verk MCV, Gatz C and Linthorst HJM 2009 Transcriptional regulation of plant defense responses. In *Plant innate immunity*, Vol. 51, Ed., LCV Loon, Academic Press, London, UK, pp. 398-426.

Webb CF 2001 The transcription factor, Bright, and immune globulin heavy chain expression. *Immunol Res* 24*:* 149–161.

Wilsker D, Patsialou A, Dallas PB and Moran E 2002 ARID proteins: a diverse family of DNA binding proteins implicated in the control of cell growth, differentiation, and development. *Cell Growth Differ* 13: 95–106.

Zhu H, Chen T, Zhu M, Fang Q, Kang H, Hong Z and Zhang A 2008 A novel ARID DNA-binding protein interacts with SymRK and is expressed during early nodule development in *Lotus japonicas*. *Plant Physiol* 148: 337–347.

(Received on 6.09.2016 and revised on 19.10.2016)