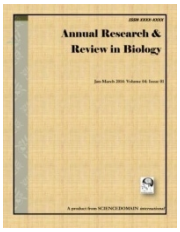


Annual Research & Review in Biology
4(1): 312-324, 2014

SCIENCEDOMAIN *international*
www.sciencedomain.org

Gene Encoding and Bioinformatics Analysis of Protein Structure of β-Galactosidase from Sunn Pest, *Eurygaster integriceps* Putton

Samin Seddigh[1*] and Maryam Darabi[2]

[1]Department of Plant Protection, Varamin-Pishva Branch, Islamic Azad University, Varamin, Iran.
[2]Department of Agronomy and Plant Breeding Sciences, College of Aboureihan, University of Tehran, Tehran, Iran.

Authors' contributions

This work was carried out in collaboration between all authors. Author SS designed the study, performed the field work and statistical analysis, wrote the protocol, and wrote the first draft of the manuscript. Author MD managed the analyses of the study. All authors read and approved the final manuscript.

Research Article

Received 30th July 2013
Accepted 14th September 2013
Published 9th October 2013

ABSTRACT

Aims: To identify the partial sequence of beta-galactosidase (EC 3.2.1.23) enzyme of sunn pest, *Eurygaster integriceps* Putton (Hemiptera: Scutelleridae), which is a key pest of wheat and barley in the wide area of the world and its relationship with other creatures.
Place and Duration of Study: Department of Plant Protection, Varamin-Pishva Branch, Islamic Azad University, Varamin, Iran, and Department of Entomology, Science and Reaserch Branch, Islamic Azad University, Tehran, Iran. Between March 2012 and July 2013.
Methodology: A part of β-galactosidase (βgal) gene was isolated from *E. integriceps* (designated as *Ei*-βgal-JQ889818), containing 328 bp. Nucleotide sequences were translated into 109 amino acids by translation tools. Twenty-six beta-galactosidase protein sequences from twenty-seven insect species, two animal samples including human and mouse, two bacteria samples including *Escherichia coli* and *Synechococcus sp.* and a sample of plants including *Arabidopsis thaliana* were aligned. Homology search was done by BLAST to identify the most similar protein sequences to *Ei*-βgal-JQ889818.
Results: Protein structure analysis revealed that the deduced *Ei*-βgal-JQ889818 had

**Corresponding author: Email: seddigh@iauvaramin.ac.ir, samin.seddigh@gmail.com;*

extensive homology with other insect βgals and contained two catalytic domains of βgals. The predicted 3-D model of *Ei*-βgal-JQ889818 has a typical spatial structure of βgals and is partly similar to βgals. Phylogenetic tree analysis of *Ei*-βgal-JQ889818 showed that there is a close relationship among *Arabidopsis thaliana, Acyrthosiphun pisum* and *Mus musculus.*
Conclusion: Accordingly, βgals should be functional proteins involved in the biosynthesis of lactose and are derived from a common ancestor. This research will lead us to know more about the role of βgal as a digestive enzyme through its phylogenetic relationship.

Keywords: β-Galactosidase; Carbohydrate enzyme; Catalytic domain; Eurygaster integriceps; Phylogenetic tree.

1. INTRODUCTION

Feeding habits of the Hemipteran order is in extreme range from strict phytophagy to exact zoophagy as well as omnivory [1,2]. Some families contain omnivores which depend on the comparative degree of animal versus plant consumption [1]. Many insects, which constitute serious pests of cereals (polysaccharide-rich diet) are dependent on their glycosidases (α-amylase, glucosidases and galactosidases) for survival. The Sunn pest, *Eurygaster integriceps* Puton (Hemiptera: Scutelleridae), is an oligophagous species that feeds on graminous plants including wheat and barley. It is a major pest of cereals in a wide area of the world including Middle East [3]. The Sunn pest infestations in some regions of the earth are so extensive which causes 100% crop loss in the absence of controlling measures [4]. They cause serious damage to crops by feeding on leaves, stems and grains. Sunn pest penetrates plant tissues with their stylets and injects digestive enzymes through the salivary canal to make liquid the plant tissue into a nutrient-rich slurry [5,6]. Proteolytic and amylolytic enzymes, injected into wheat grains, demolish their gluten and lessen the baking quality of flour [7,8,6,4]. Pesticide usage is the main method of the insect control which is used either against adults or nymphs. In recent years, different methods based on molecular approaches are used in order to control insect pest especially the first-generation of crop expressing BT toxin have been successful [9,10]. These approaches are use of digestive enzymes inhibitors, BT toxins alone or in combination with α-amylase or proteinase inhibitors, glycosidae inhibitors, lectins, and RNA interference technology [11,4,12].

Many insect species, which constitute serious pests of cereals (polysaccharide-rich diet) are depended on their carbohydrate-active enzymes for their survival. In insects, different type of digestive glycosidases including α- and β-glucosidase, α- and β-galactosidase, α- and β-mannosidase, β-fructofuranosidase and some other carbohydrate hydrolyzing enzymes are known to exist [13,14,15,16,17].

β-Galactosidase (EC 3.2.1.23) is a hydrolase enzyme that catalyzes the hydrolysis of β-galactosides (like lactose) into monosaccharides (such as galactose and glucose), where the glucosidic group on the non-reducing β-D-galactose residue is replaced by a hydroxyl group (an acceptor group). β-Galactosidase is also known to catalyze the transglycosylation of sugars i.e. when a sugar moiety is the acceptor instead of the water molecule. Latter on this will cause the synthesis of new oligosaccharides [18]. βGal, the product of the lacZ gene of *Escherichia coli*, is the most extensively reporter gene which is used in the study of genetics, cell and molecular biology [19]. The 1024 amino acid residues of *E. coli* β-galactosidase were sequenced in 1970 by Fowler and Zabin for the first time [20]. βGal monomer constructed from five different domains arranged around a central alpha/beta

barrel [21]. Domain one is a jelly-roll type barrel; domain two and four are a fibronectin type III-like barrels, domain five a β-sandwich, while the central domain three is a TIM-type barrel. The third domain contains the active site [22]. The active site is shaped from residues of four of the domains into a pocket which complements the relatively tiny size of lactose. Beta-galactosidase can be split in two peptides, LacZα and LacZΩ, none of which is active by itself but both spontaneously reassemble into a functional enzyme [21].

Despite the fact that some functions of β-galactosidases have been investigated before, limited information is available on this protein structure in creatures. Recently, bioinformatics and genomic tools have revolutionized the studies of the metabolism in different organisms [23,24,25]. By considering the significance of carbohydrate digestion as a target for sunn pest control, a study on their digestive enzymes could surely be crucial in adopting of new control procedures. So, the aim of the current study was to identify a part of βgal protein of *E. integriceps* and the phylogenetic relationship of βgals by different tools to gain a better understanding of the digestive physiology of the insect. This knowledge will optimistically lead to new and winning management strategies for controlling of this pest.

2. MATERIALS AND METHODS

2.1 Insect Samples and Growth Condition

Adult insects (*Eurygaster integriceps*) were collected from the Pakdasht wheat farm of Tehran Province, Iran. They maintained on wheat kernels in the laboratory at 27±2 ͦc under a 14h light: 10h dark (LD 14:10) photoperiod [4].

2.2 DNA Isolation

Adults of sunn pest were used for DNA isolation in all experiments. Total DNA of *E. integriceps* was extracted using the modified CTAB method [26] with slight modification. The qualities of the extracted DNA were checked by agarose gel electrophoresis and Nano-Drop spectrophotometer (Model Thermo Scientific 1000). After the extraction, the DNA samples were stored at -20°C for subsequent usage.

2.3 Isolation of β-galactosidase Gene Fragment

Four degenerate oligonucleotide primers were designed and synthesized based on the conserved DNA sequence regions of all insects βgal by AllelIID 6.0 as follows:

βgal F1- 5′- CGGAATTCACNTAYGTNGAR-3′
βgal F2- 5′- AATTCCARGTNGARAAYGARTAY-3′
βgal R1- 5′- AAGCTTNCCNCCRTARAACATRTA-3′
βgal R2- 5′- AAGCTTNGCRTCRTARTCRTA-3′

PCR was performed to amplify the fragment of βgal gene with the two degenerate primers according to the protocol of CinnaGen Master Mix PCR Kit. The PCR program was carried out at 94°C for 5 min followed by 35 cycles of amplification (1 min of denaturation at 94°C, 1 min of annealing 52°C, 1 min of extension at 72°C). After the final cycle, the amplification was extended for 10 min at 72°C. The PCR products were separated on 1.2% agarose gel stained with ethidium bromide.

A pair of primers, 18S-F (5′-ATTGAGGTCTTCGGAGTG-3′) and 18S-R (5′-GATTTCGGTCATCTTGCG-3′), were also designed by AllelIID 6.0 as an internal control. The template was denatured at 94°C for 5 min, followed by 35 cycles of amplification (94°C for 1 min, 50°C for 1 min and 72°C for 1 min) and by extension at 72°C for 10 min. The products were separated on 1.2% agarose gel stained with ethidium bromide, and the amplified products were sent to CinnaGen Company for DNA sequencing.

2.4 Data Collection and Analysis

The βGal protein sequence belonging to the sunn pest which was identified in this study was compared with sequences of two animal species, including *Homo sapiens* (Hominidae) and *Mus musculus* (Muridae), a plant species, *Arabidopsis thaliana* (Brassicaceae), and two bacterial species, *Escherischia coli* (Enterobacteriaceae) and *Synechococcus sp.* (Synechococcaceae) which were downloaded from NCBI (http://www.ncbi.nlm.nih.gov) (date received: May 2013) in FASTA and GenBank format. The number of data given was consisted of 26 insects' protein data, two animals, one plant and two bacterial protein data, which were related to β-galactosidase enzyme. They are listed in Table 1.

The homology-based 3-D structural modeling of βgal was accomplished by EsyPred 3-D, and Web Lab Viewer Lite 4.0 was used for 3-D structure displaying. Functional domains of *Ei*-βgal-JQ889818 were analyzed by ProDom. Homology search was carried out online by BLAST (Basic Local Alignment Search Tool) at the NCBI website (http://www.ncbi.nlm.nih.gov). Nucleotide Sequence was translated to amino acid by Translate tool in EXPASY webserver (http://www.expasy.org). The sequence alignment of βgals was performed by CLUSTALW (MegAlign) using default parameters. To investigate the evolutionary relationships among different βgal proteins, a phylogenetic tree was constructed based on the deduced amino acid sequence of *Ei*-βgal-JQ889818 and other βgals from different organisms including insects, human, plants, bacteria and animals using the Mega 5 program from aligned sequences. Maximum parsimony method (MP) was used to construct the tree.

Table 1. The thirty two species and their family which were analyzed in this study. The abbreviations of each sequence were used for better access to the species

Index	Scientific name	Family	Abbreviation
1	*Aedes aegypti*	Culicidae	*Aa*-βgal
2	*Acyrthosiphun pisum*	Aphididae	*Ap*-βgal
3	*Acromyrmex echinator*	Formicidae	*Ae*-βgal
4	*Anopheles darlingi*	Culicidae	*Ad*-βgal
5	*Anopheles gambiae str. Pest*	Culicidae	*Ag*-βgal
6	*Apis melifera*	Apidae	*Am*-βgal
7	*Arabidopsis thaliana*	Brassicaceae	*At*-βgal
8	*Bombyx mori*	Bombycidae	*Bm*-βgal
9	*Camponatus floridanus*	Formicidae	*Cf*-βgal
10	*Culex quinquefasciatus*	Culicidae	*Cq*-βgal
11	*Drosophila ananassae*	Drosophilidae	*Da*-βgal
12	*Drosophila erecta*	Drosophilidae	*De*-βgal

13	*Drosophila grimshawi*	Drosophilidae	*Dg*-βgal
14	*Drosophila melanogaster*	Drosophilidae	*Dm*-βgal
15	*Drosophila mojavensis*	Drosophilidae	*Dmo*-βgal
16	*Drosophila persimilis*	Drosophilidae	*Dp*-βgal
17	*Drosophila pseudoobscura*	Drosophilidae	*Dps*-βgal
18	*Drosophila sechellia*	Drosophilidae	*Ds*-βgal
19	*Drosophila simulans*	Drosophilidae	*Dsi*-βgal
20	*Drosophila virilis*	Drosophilidae	*Dv*-βgal
21	*Drosophila willistoni*	Drosophilidae	*Dw*-βgal
22	*Drosophila yakuba*	Drosophilidae	*Dy*-βgal
23	*Escherschia coli*	Enterobacteriaceae	*Ec*-βgal
24	*Eurygaster integriceps*	Scutelleridae	*Ei*-βgal
25	*Glossina moristans moristans*	Glossinidae	*Gmm*-βgal
26	*Harpegnathos saltator*	Formicidae	*Hsal*-βgal
27	*Homo sapiens*	Homonidae	*Hs*-βgal
28	*Mus musculus*	Muridae	*Mm*-βgal
29	*Nasonia vitripennis*	Pteromalidae	*Nv*-βgal
30	*Pediculus humanus corporis*	Pediculidae	*Ph*-βgal
31	*Synechococcus sp.*	Synechococcaceae	*Ss*-βgal
32	*Tribolium castaneum*	Tenebrionidae	*Tc*-βgal

3. RESULTS AND DISCUSSION

3.1 βGal Identification

Total DNA isolated from *E. integriceps* and subjected to the PCR using two pairs of degenerate primers named βgal F1 and βgal F2 as forward primers and βgal R1 and βgal R2 as reverse primers. Results showed that a pair of βgal F2 and βgal R1 primers amplified 328-bp product of the gene of interest (βgal) (Fig. 1). The obtained sequence was deposited to NCBI with an accession number of JQ889818.

3.2 Characterization of the βgal Protein

Sequence comparisons performing BLAST search in GenBank database (http://www.ncbi.nih.gov) and multiple alignment analysis revealed that *Ei*-βgal-JQ889818 had high homology with a bacteria βgal, *Paenibacillus sp* but fewer homologies with an insect species, *Tribolium castaneum* and an animal species, *Rattus norvegicus*. *Ei*-βgal-JQ889818 which contained 328 bp was translated into 109 amino acids by Translate tool (Fig. 2). Results analysis by ProDom revealed that the catalytic regions of *Ei*-βgal-JQ889818 consisted of two domains: PD003386 with 95bp, between amino acids 6-101 and PD328308 with 65bp, between amino acids 40-105 (Fig. 2).

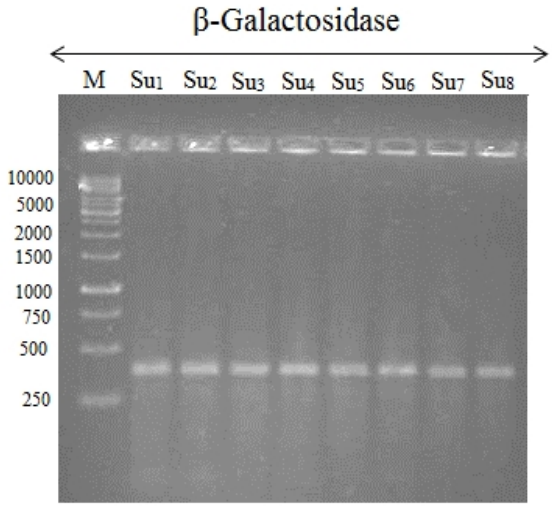
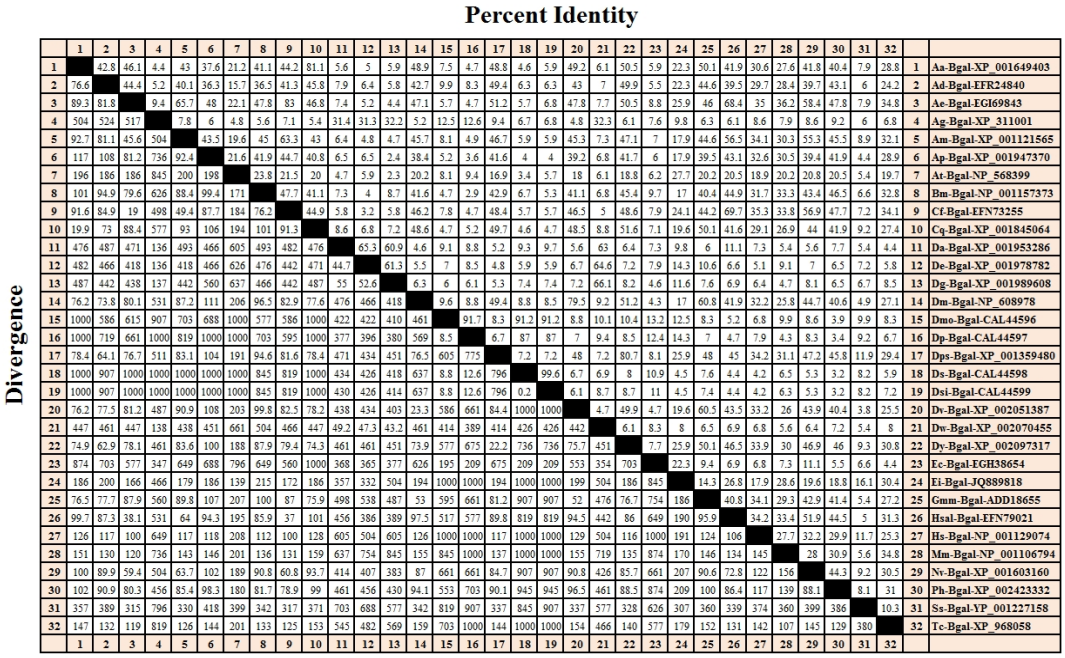


Fig. 1. DNA isolated from *Eurygaster integriceps* by the degenerate primers βgal F2 and βgal R1. M is the Ladder, Su_1-Su_8 are *E. integriceps* βgals samples repetitions.

```
acagccacatatgaccaccttggggacgggatgctgcaacgagggatagacgtcccgctg
 T  A  T  Y  D  H  L  G  D  G  M  L  Q  R  G  I  D  V  P  L
attacatgtgttggcggagcggaaggaacgattgaaggagcgaacttctggtctggtgcg
 I  T  C  V  G  G  A  E  G  T  I  E  G  A  N  F  W  S  G  A
gatggacattatgcgaatttgcgtgcgaaacagccggatacaccaaagatggttactgaa
 D  G  H  Y  A  N  L  R  A  K  Q  P  D  T  P  K  M  V  T  E
ttctggaccggatggtttgaaaattggggaggaccttcggccatccagaagacggcttct
 F  W  T  G  W  F  E  N  W  G  G  P  S  A  I  Q  K  T  A  S
ctgctggaccgaagaatcatggaaattctgagagctggatataccggaatcagttactat
 L  L  D  R  R  I  M  E  I  L  R  A  G  Y  T  G  I  S  Y  Y
atgttctacggaggaaagcttaatacaa
 M  F  Y  G  G  K  L  N  T
```

Fig. 2. The partial nucleotide and protein sequence of *Ei*-βgal-JQ889818. ProDom indicated two catalytic domains: PD003386 (marked with yellow) and PD328308 (underlined with red).

Protein BLAST revealed that *Ei*-βgal-JQ889818 belongs to the Glyco-hydro-42 superfamily and was known as βgals family. On the amino acid scale, *Ei*-βgal-JQ889818 was 30.4% identical to *Tribolium castaneum* (*Tc*-βgal-XP_968058) and 4.5% identical to *Drosophila sechellia* (*Ds*-βgal-CAL44598) and *Drosophila simulans* (*Dsi*-βgal-CAL44599) as the highest and lowest identical, respectively. Multiple alignments revealed that βgal protein in different organisms had high similarity with each other (Fig. 3). The most identities are shown in Table 2.



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	
1		42.8	46.1	4.4	43	37.6	21.2	41.1	44.2	81.1	5.6	5	5.9	48.9	7.5	4.7	48.8	4.6	5.9	49.2	6.1	50.5	5.9	22.3	50.1	41.9	30.6	27.6	41.8	40.4	7.9	28.8	1	Aa-Bgal-XP_001649403
2	76.6		44.4	5.2	40.1	36.3	15.7	36.5	41.3	45.8	7.9	6.4	5.8	42.7	9.9	8.3	49.4	6.3	6.3	43	7	49.9	5.5	22.3	44.6	39.5	29.7	28.4	39.7	43.1	6	24.2	2	Ad-Bgal-EFR24840
3	89.3	81.8		9.4	65.7	48	22.1	47.8	83	46.8	7.4	5.2	4.4	47.1	5.7	4.7	51.2	5.7	6.8	47.8	7.7	50.5	8.8	25.9	46	68.4	35	36.2	58.4	47.8	7.9	34.8	3	Ae-Bgal-EGI69843
4	504	524	517		7.8	6	4.8	5.6	7.1	5.4	31.4	31.3	32.2	5.2	12.5	12.6	9.4	6.7	6.8	4.8	32.3	6.1	7.6	9.8	6.3	6.1	8.6	7.9	8.6	9.2	6	6.8	4	Ag-Bgal-XP_311001
5	92.7	81.1	45.6	504		43.5	19.6	45	63.3	43	6.4	4.8	4.7	45.7	8.1	4.9	46.7	5.9	5.9	45.3	7.3	47.1	7	17.9	44.6	56.5	34.1	30.3	55.3	45.5	8.9	32.1	5	Am-Bgal-XP_001121565
6	117	108	81.2	736	92.4		21.6	41.9	44.7	40.8	6.5	6.5	2.4	38.4	5.2	3.6	41.6	4	4	39.2	6.8	41.7	6	17.9	39.5	43.1	32.6	30.5	39.4	41.9	4.4	28.9	6	Ap-Bgal-XP_001947370
7	196	186	186	845	200	198		23.8	21.5	20	4.7	5.9	2.3	20.2	8.1	9.4	16.9	3.4	5.7	18	6.1	18.8	6.2	27.7	20.2	20.5	18.9	20.2	20.8	20.5	5.4	19.7	7	At-Bgal-NP_568399
8	101	94.9	79.6	626	88.4	99.4	171		47.7	41.1	7.3	4	8.7	41.6	4.7	2.9	42.9	6.7	5.3	41.1	6.8	45.4	9.7	17	40.4	44.9	31.7	33.3	43.4	46.5	6.6	32.8	8	Bm-Bgal-NP_001157373
9	91.6	84.9	19	498	49.4	87.7	184	76.2		44.9	5.8	3.2	5.8	46.2	7.8	4.7	48.4	5.7	5.7	46.5	5	48.6	7.9	24.1	44.2	69.7	35.3	33.8	56.9	47.7	7.2	34.1	9	Cf-Bgal-EFN73255
10	19.9	73	88.4	577	93	106	194	101	91.3		8.6	6.8	7.2	48.6	4.7	5.2	49.7	4.6	4.7	48.5	8.8	51.6	7.1	19.6	50.1	41.6	29.1	26.9	44	41.9	9.2	27.4	10	Cq-Bgal-XP_001845064
11	476	487	471	136	493	466	605	493	482	476		65.3	60.9	4.6	9.1	8.8	5.2	9.3	9.7	5.6	63	6.4	7.3	9.8	6	11.1	7.3	5.4	5.6	7.7	5.4	4.4	11	Da-Bgal-XP_001953286
12	482	466	418	136	418	466	626	476	442	471	44.7		61.3	5.5	7	8.5	4.8	5.9	5.9	6.7	64.6	7.2	7.9	14.3	10.6	6.6	5.1	9.1	7	6.5	7.2	5.8	12	De-Bgal-XP_001978782
13	487	442	438	137	442	560	637	466	442	487	55	52.6		6.3	6	6.1	5.3	7.4	7.4	7.2	66.1	8.2	4.6	11.6	7.6	6.9	6.4	4.7	8.1	6.5	6.7	8.5	13	Dg-Bgal-XP_001989608
14	76.2	73.8	80.1	531	87.2	111	206	96.5	82.9	77.6	476	466	418		9.6	8.8	49.4	8.8	8.5	79.5	9.2	51.2	4.3	17	60.8	41.9	32.2	25.8	44.7	40.6	4.9	27.1	14	Dm-Bgal-NP_608978
15	1000	586	615	907	703	688	1000	577	586	1000	422	422	410	461		91.7	8.3	91.2	91.2	8.8	10.1	10.4	13.2	12.5	8.3	5.2	6.8	9.9	8.6	3.9	9.9	8.3	15	Dmo-Bgal-CAL44596
16	1000	719	661	1000	819	1000	1000	703	595	1000	377	396	380	569	8.5		6.7	87	87	7	9.4	8.5	12.4	14.3	7	4.7	7.9	4.3	8.3	3.4	9.2	6.7	16	Dp-Bgal-CAL44597
17	78.4	64.1	76.7	511	83.1	104	191	94.6	81.6	78.4	471	434	451	76.5	605	775		7.2	7.2	48	7.2	80.7	8.1	25.9	48	45	34.2	31.1	47.2	45.8	11.9	29.4	17	Dps-Bgal-XP_001359480
18	1000	907	1000	1000	1000	1000	1000	845	819	1000	434	426	418	637	8.8	12.6	796		99.6	6.7	6.9	8	10.9	4.5	7.6	4.4	4.2	6.5	5.3	3.2	8.2	5.9	18	Ds-Bgal-CAL44598
19	1000	907	1000	1000	1000	1000	1000	845	819	1000	430	426	414	637	8.8	12.6	796	0.2		6.1	8.7	8.7	11	4.5	7.4	4.4	4.2	6.3	5.3	3.2	8.2	7.2	19	Dsi-Bgal-CAL44599
20	76.2	77.5	81.2	487	90.9	108	203	99.8	82.5	78.2	438	434	403	23.3	586	661	84.4	1000	1000		4.7	49.9	4.7	19.6	60.5	43.5	33.2	26	43.9	40.4	3.8	25.5	20	Dy-Bgal-XP_002051387
21	447	461	447	138	438	451	661	504	466	447	49.2	47.3	43.2	461	414	389	414	426	426	442		6.1	8.3	8	6.5	6.9	6.8	5.6	6.4	7.2	5.4	8	21	Dw-Bgal-XP_002070455
22	74.9	62.9	78.1	461	83.6	100	188	87.9	79.4	74.3	461	461	451	73.9	577	675	22.2	736	736	75.7	451		7.7	25.9	50.1	46.5	33.9	30	46.9	46	9.3	30.8	22	Dy-Bgal-XP_002097317
23	874	703	577	347	649	688	796	649	560	1000	368	365	377	626	195	209	675	209	209	553	354	703		22.3	9.4	6.9	6.8	7.3	11.1	5.5	6.6	4.4	23	Ec-Bgal-EGH38654
24	186	200	166	466	179	186	139	215	172	186	357	332	504	194	1000	1000	194	1000	1000	199	504	186	845		14.3	26.8	17.9	28.6	19.6	18.8	16.1	30.4	24	Ei-Bgal-JQ889818
25	76.5	77.7	87.9	560	89.8	107	207	100	87	75.9	498	538	487	53	595	661	81.2	907	907	52	476	76.7	754	186		40.8	34.1	29.3	42.9	41.4	5.4	27.2	25	Gmm-Bgal-ADD18655
26	99.7	87.3	38.1	531	64	94.3	195	85.9	37	101	456	386	389	97.5	517	577	89.8	819	819	94.5	442	86	649	190	95.9		34.2	33.4	51.9	44.5	5	31.3	26	Hsal-Bgal-EFN79021
27	126	117	100	649	117	118	208	112	100	128	605	504	605	126	1000	1000	117	1000	1000	129	504	116	1000	191	124	106		27.7	32.2	29.9	11.7	25.3	27	Hs-Bgal-NP_001129074
28	151	130	120	736	143	146	201	136	131	159	637	754	845	155	845	1000	137	1000	1000	155	719	135	874	170	146	134	145		28	30.9	5.6	34.8	28	Mm-Bgal-NP_001106794
29	100	89.9	59.4	504	63.7	102	189	90.8	60.8	93.7	414	407	383	87	661	661	84.7	907	907	90.8	426	85.7	661	207	90.6	72.8	122	156		44.3	9.2	30.5	29	Nv-Bgal-XP_001603160
30	102	90.9	80.3	456	85.4	98.3	180	81.7	78.9	99	461	456	430	94.1	553	703	90.1	945	945	96.5	461	88.5	874	209	100	86.4	117	139	88.1		8.1	31	30	Ph-Bgal-XP_002423332
31	357	389	315	796	330	418	399	342	317	371	703	688	577	342	819	907	337	845	907	337	577	328	626	307	360	339	374	360	399	386		10.3	31	Ss-Bgal-YP_001227158
32	147	132	119	819	126	144	201	133	125	153	545	482	569	159	703	1000	144	1000	1000	154	466	140	577	179	152	131	142	107	145	129	380		32	Tc-Bgal-XP_968058
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	

Fig. 3. The degree of percentage sequence identity of residues across the thirty-two different organisms' protein sequences.

Table 2. The percentage identity between βgal protein sequences (above 80%) in the multiple alignments of different insect species

Index	Characteristics	Identity (%)
1	*Ae*-βgal-EGI69843, *Cf*-βgal-EFN73255	83
2	*Cq*-βgal-Xp_001845064, *Aa*-βgal-Xp_001649403	81.1
3	*Dmo*-βgal-CAL44596, *Dp*-βgal-CAL44597	91.7
4	*Dp*-βgal-CAL44597, *Ds*-βgal-CAL44598	87
5	*Dmo*-βgal-CAL44596, *Ds*-βgal-CAL44598	91.2
6	*Dmo*-βgal-CAL44596, *Dsi*-βgal-CAL44599	91.2
7	*Dp*-βgal-CAL44597, *Dsi*-βgal-CAL44599	87
8	*Ds*-βgal-CAL44598, *Dsi*-βgal-CAL44599	99.6
9	*Dps*-βgal-XP_001359480, *Dy*-βgal-XP_002097317	80.7

Phylogenetic tree constructed based on deduced amino acid sequences of *Ei*-βgal-JQ889818 and other βgals from different organisms revealed that βgals were derived from an ancestor protein and evolved into different groups including insects, animals, plants and bacteria βgal groups. According to the phylogenetic tree, *Ei*-βgal-JQ889818 belongs to the insect group and has the closest relationship with βgal from *Arabidopsis thaliana, Acyrthosiphun pisum* and *Mus musculus* (Fig. 4).

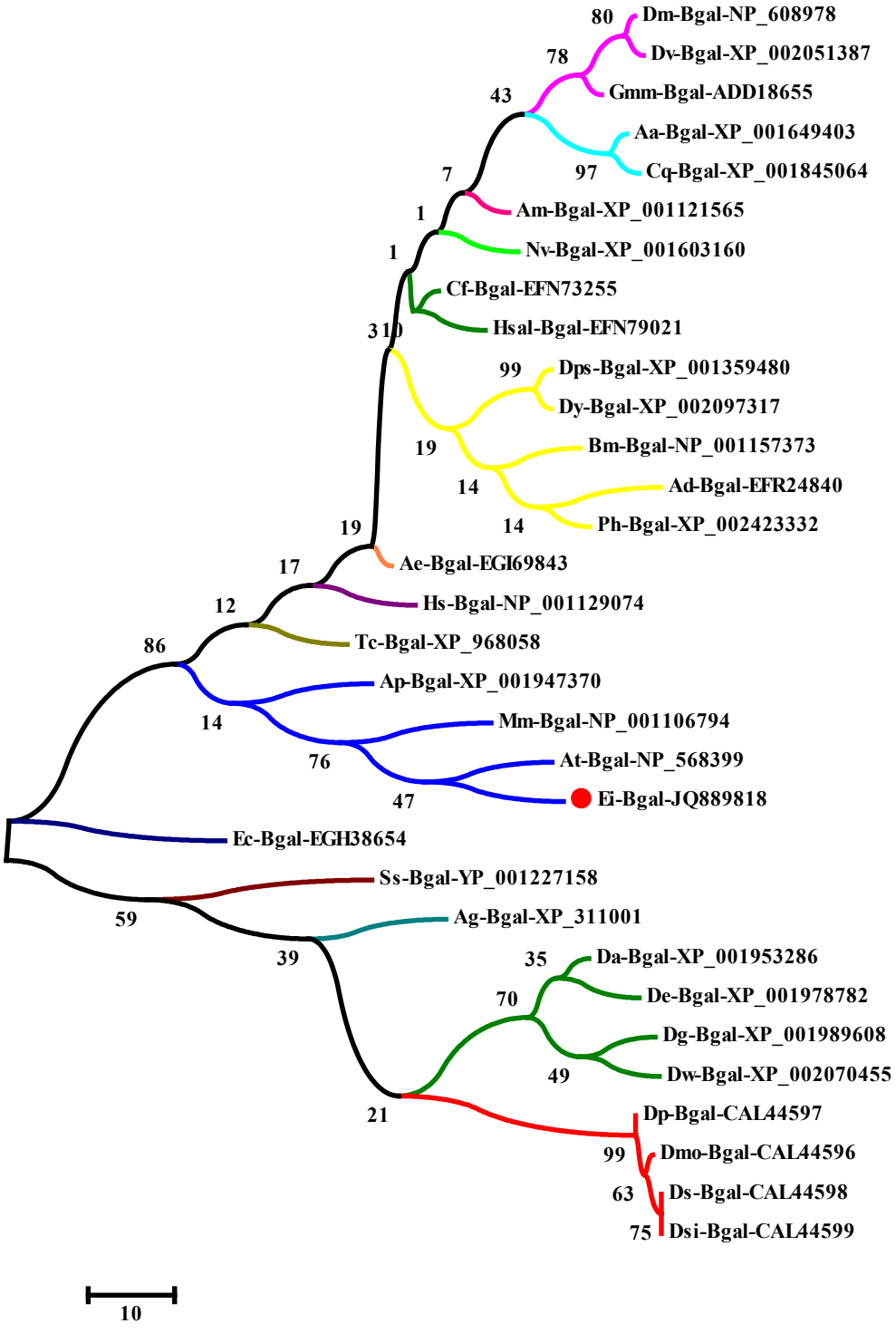


Fig. 4. The phylogenetic tree of insects, bacteria, *Arabidopsis thaliana*, human and mouse constructed by MEGA 5 program (using the CLUSTALW method).

The Maximum Parsimony (MP) method was used to construct the tree. The new βgal protein sequence of *Eurygaster integriceps* was marked by red circle. The percentage of 500 bootstrap replicates was given at each node. The accession number of each species is followed by their abbreviations.

The homology-based 3-D structural modeling of *Ei*-βgal-JQ889818 was predicted by ESYPred 3-D (Automated homology modeling program using neural networks) (http://www.fundp.ac.be/sciences/biologie/urbm/bioinfo/esypred/) and compared with swiss-modeling (http://swissmodel.expasy.org/) [27,28,29] on the basis of *E.coli* crystal structure and displayed by Web Lab Viewer Lite. Molecular modeling results showed that *Ei*-βgal-JQ889818 has a simple spatial architecture (Fig. 5).

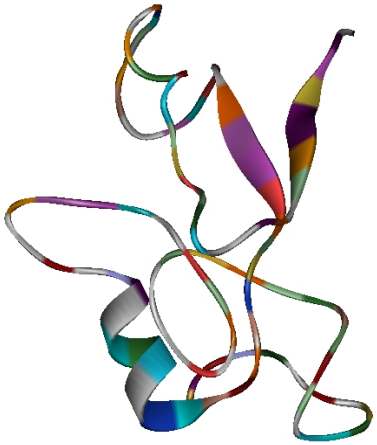


Fig. 5. 3-D Structure of *Ei*-βgal-JQ889818 established by homology-based modeling using EsyPred 3-D. The α-helix is shown with helix-shaped, the beta sheet with wide ribbon-shaped and the random coil with line-shaped.

β-Galactosidase (βgal, EC 3.2.1.23) is an eukaryotic hydrolase localized in the lysosome [30]. In the present study, we isolated and characterized a partial sequence of βgal gene of *E. integriceps*, a key pest of wheat and barley. This gene size is not the same in different insects so that in *Acyrthosophon pisum* a 172bp βgal was known whereas 1745bp βgal was found in *Tribolium castaneum*. In this study the 328bp βgal was identified which codes 109 amino acids in Sunn pest. This may be larger in *E. integriceps*, but there are no witnesses to say the exact size of βgal gene because the sunn pest genome has not identified yet. Beta-galactosidase is an exoglycosidase that cleaves β-linked terminal galactosyl residues from a variety of natural and artificial substrates [31]. β-Galactosidase assay is used frequently in genetics, molecular biology for a blue white screen, and other life sciences. The βgal reporter gene is vastly used in the biological research. Although originally identified from *E. coli*, this reporter is functional in many other organisms such as yeast, *Caenorhabditis elegans*, *Drosophila* and mammals. Other advantages of the βgal system contain the stability of βgal enzyme, easy activity assay procedure, and availability of a broad range of substrates [19]. Protein BLAST showed that *Ei*-βgal-JQ889818 belongs to βgal family and shared significant sequence similarity with βgal from other organisms. In *E. coli*, the biologically active βgal enzyme exists as a tetramer of four identical subunits [32]. In this study, the 3-D structure of *Ec*-βgal has been known, and the structural basis for its reaction mechanism has been reported. 3-D structural modeling of *Ei*-βgal-JQ889818 which has a

simple spatial architecture was similar to one part of *E. coli* βgal whose catalytic portion and activity was determined before [20,22]. Two functional domains has been found in *E. integriceps* βgal gene (*Ei*-βgal-JQ889818), which was confirmed by Gene Ontology as one of the most crucial subunits of beta-galactosidase protein. So, Gene Ontology (http://www.geneontology.org/) results strongly suggest that *Ei*-βgal-JQ889818 is a functional β-galactosidase protein in *E. integriceps* which involved in the catalyzing the hydrolysis of terminal, non-reducing beta-D-galactose residues in beta-D-galactosides.

Phylogenetic tree constructed based on the amino acid sequences and the sequence similarities were found. Different organisms' βgal were close to each other and it can be used in future studies. For instance phylogenetic analysis revealed that *E. integriceps* βgal has a close relationship with some bacteria's βgal. So, in order to improve a functional role it is better to isolate this enzyme from these two creatures and compare their kinetic parameters in order to achieve better understanding of the characterization of the βgal from two different phylogenetic groups, i.e. insect and bacteria. Fig. 4 showed that βgals were derived from an ancestor protein in all organisms and evolved into different groups.

Digestive system of the Sunn pest can be a proper target to apply new managing strategies based on interruption in the digestion process, i.e. enzyme inhibition and toxin delivery. Thus, study of the digestive system and enzymes of the Sunn pest can provide more information regarding phytophagus digestive physiology. Further, and make this possible to find new digestion related controlling methods [33]. This study can lead us to new strategies of sunn pest control such as transgenic insects.

4. CONCLUSION

The partial sunn pest βgal enzyme was identified and deposited to NCBI with the JQ889818 accession number. It was compared with twenty-six beta-galactosidase protein sequences from twenty-seven insect species, two animal samples including human and mouse, two bacteria samples including *Escherichia coli* and *Synechococcus sp.* and a sample of plants including *Arabidopsis thaliana* by bioinformatics tools. Different computational tools were used in this study. The alignment was performed by MegAlign and the phylogenetic trees were constructed by Molecular Evolutionary Genetics Analysis (MEGA), version 5. Functional domains ad 3-D structure were analyzed by PRODOM and EsyPred 3-D, respectively. Sequence comparison analysis showed that there is a high identity between these species and phylogenetic analysis indicated that there is a relationship among species of different organisms. These results of phylogenetic tree showed βgal have a relationship with each other in different creatures, also it can be proved they were derived from an ancestor gene and evolved into different groups.

ACKNOWLEDGEMENTS

This research was supported by Islamic Azad University, Varamin-Pishva Branch. We sincerely appreciate them.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Cohen AC. Plant feeding by predatory Heteroptera: evolutionary and adaptational aspects of trophic switching. In: Alomar O, Wiedenmann RN, editors. Zoo-phytophagous Heteroptera: implications for life history and integrated pest management. Thomas Say Publications in Entomology. Entomological Society of America, Lanham, Maryland, USA. 1996;1-17.
2. Schaefer CW, Panizzi AR. Economic importance of Heteroptera: A general view. In: Schaefer CW, Panizzi AR, editors. Heteroptera of economic importance. CRC. Boca Raton, FL. 2000;3-8.
3. Critchley BR. Literature review of sunn pest *Eurygaster integriceps* Put. (Hemiptera, Scutelleridae). Crop Prot. 1998;4:271-287.
4. Allahyari M, Bandani AR, Habibi-Rezaei M. Subcellular fraction of midgut cells of the sunn pest *Eurygaster integriceps* (Hemiptera: Scutelleridae): Enzyme markers of microvillar and perimicrovillar membranes. J Insect Physiol. 2010;56:710-717.
5. Hosseininaveh V, Bandani A, Hosseininaveh F. Digestive proteolytic activity in the Sunn pest, *Eurygaster integriceps*. J Insect Sci. 2009;9:70-81.
6. Mehrabadi M, Bandani AR, Saadati F, Ravan S. Sun pest, *Eurygaster integriceps* Putton (Hemiptera: Scutelleridae), digestive α-amylase, α-glucosidase and β-glucosidae. J Asia Pac Entomol. 2009;12:79-83.
7. Hariri G, Williams PC, El-Haramein FJ. Influence of Pentatomoid insects on the physical dough properties and two-layered flat bread baking quality of Syrian wheat. J Cereal Sci. 2000;31:111-118.
8. Rajabi GH. Ecology of cereal Sunn Pest in Iran. Agricultural Research, Education and Extension Organization Press, Tehran, Iran. 2000;343.
9. Christou P, Capell T, Kohli A, Gatehouse JA, Gatehouse AMR. Recent developments and future prospects in insect pest control in transgenic crops. Trends Plant Sci. 2006;11:302-308.
10. Ferry N, Edwards MG, Gatehouse JA, Capell T, Christou P, Gatehouse AMR. Transgenic plants for insect pest control: a forward looking scientific perspective. Transgenic Re. 2006;15(1):13-19.
11. Bellincampi D, Camardella L, Delcour JA, Desseaux V, D'ovidio R, Durand A, et al. Potential physiological role of plant glycosidase inhibitors. Biochim Biophys Acta. 2004;1696:265-274.
12. Rother S, Meister G. Small RNAs derived from longer non-coding RNAs. Biochimie. 2011;93(11):1905-1915.
13. Terra WR, Ferreira C. Comprehensive Molecular Insect Science. In: Gilbert LI, Iatrou K, Gill SS, editors. Biochemistry of digestion. vol. 4. Elsevier, Oxford; 2005.
14. Mendiola-Olaya E, Valencia-Jimenez A, Valdes-Rodriguez S, Delano-Frier J, Blanco-Labra A. Digestive amylase from the larger grain borer, Prostephanus truncates Horn. Comp Biochem Phys B. 2000;126:425-433.
15. Zibaee A, Bandani AR, Kafil M, Ramzi S. Characterization of α -amylase in the midgut and the salivary glands of rice striped stem borer, Chilo suppressalis Walker (Lepidoptera: Pyralidae). J Asia-Pacific Entomol. 2008;11:201-205.
16. Seddigh S, Bandani AR. Comparison of α and β-galactosidase activity in the three cereal pests, *Haplothrips tritici* Kurdjumov (Thysanoptera: Phlaeothripidae), *Rhopalosiphum padi* L. (Hemiptera: Aphididae) and *Eurygaster integriceps* Puton (Hemiptera: Scutelleridae). Mun Ent Zool. 2012;7:904-908.

17. Seddigh S, Masoudi-Nejad A, Tafaghodinia B, Imani S. Comparing carbohydrate enzymes activity in *Eurygaster integriceps* Putton (Hemiptera: Scutelleridae), *Rhopalosiphum padi* L. (Homoptera: Aphididae) and Haplothrips tritici Kurd. (Phlaeothripidae: Thysanoptera) as a complex pest on wheat. Mun Ent Zool. 2012;7:344-351.
18. Yamamoto Y, Saito T, Ajisaka K. Study of the regioselectivity in the transglycosylation to D-galactose derivatives using β-galactosidases of various origins. J Appl Glycosci. 2004;51:335-339.
19. Serebriiskii IG, Golemis EA. Uses of lacZ to study gene function: evaluation of beta-galactosidase assays employed in the yeast two-hybrid system. Anal Biochem. 2000; 285: 1-15.
20. Fowler AV, Zabin I. The amino acid sequence of beta galactosidase. I. Isolation and composition of tryptic peptides. J Biol Chem. 1970;245:5032-5041.
21. Jacobson RH, Zhang XJ, DuBose RF, Matthews BW. Three-dimentionam structure of beta-galactosidase from E. coli. Nature. 1994;369:761-766.
22. Matthews BW. The structure of E. coli beta-galactosidase. C R Biol. 2005;328:549-556.
23. Darabi M, Seddigh S. Conserved motifs identification of 3-hydroxy-3-methylglotaryl-coenzyme A reductase (HMGR) protein in some different species of drosophilidae by bioinformatics tools. Ann Biol Res. 2013;4:158-163.
24. Darabi M, Masoudi-Nejad A, Nemat-Zadeh G. Bioinformatics study of the 3-hydroxy-3-methylglotaryl-coenzyme A reductase (HMGR) gene in Gramineae. Mol Biol Rep. 2012;39:8925-8935.
25. Darabi M, Seddigh S. Phylogenetic study of the 3-hydroxy-3-methylglotaryl-Coenzyme A reductase (HMGR) protein in six different family. Eur J Exp Biol. 2013;3:158-164.
26. Doyle JJ, Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull. 1987;19:11-15.
27. Guex N, Peitsch MC. SWISS-MODEL and the Swiss- Pdb Viewer: an environment for comparative protein modeling. Electrophoresis. 1997;18:2714-2723.
28. Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res. 2003;31:3381-3385.
29. Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modeling. Bioinformatics. 2006;22:195-201.
30. Suzuki Y, Sakuraba H, Oshima A. β-galactosidase deficiency (β-galactosidosis): GM1 gangliosidosis and Morquio B disease. In: Scriver CR, Beaudet AL, Sly WS, editors. The Metabolic and Molecular Bases of Inherited Disease, 7th ed. New York, McGraw Hill. 1995;2801-2810.
31. Lee BY, Han JA, Im JS, Morrone A, Johung K, Goodwin EC, et al. Senescence-associated beta-galactosidase is lysosomal beta-galactosidase. Aging Cell. 2006;5:187-195.
32. Kalnins A, Otto K, Ruther U, Muller-Hill B. Sequence of the lac Z gene of Escherichia coli, EMBO J. 1983;2:593-597.

33. Mehrabadi M, Bandani AR. Secretion and formation of perimicrovillar membrane (PMM) in digestive system of the Sunn pest, *Eurygaster integriceps* (Hemiptera: Scutelleridae) in response to feeding. Arch Insect Biochem. 2011;78:190-200.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sciencedomain.org/review-history.php?iid=287&id=32&aid=2228