Gene 575 (2016) 478–487

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Research paper

In silico identification of BESS-DC genes and expression analysis in the silkworm, *Bombyx mori*

Zhongchen Rao [a], Jun Duan [b], Qingyou Xia [b], Qili Feng [a,*]

[a] Guangzhou Key Laboratory of Insect Development Regulation and Application Research, School of Life Sciences, South China Normal University, Guangzhou, 510631, China
[b] State Key Laboratory of Silkworm Genome Biology, Southwest University, 400716, China

ARTICLE INFO

Article history:
Received 27 January 2015
Received in revised form 10 September 2015
Accepted 11 September 2015
Available online 16 September 2015

Keywords:
BESS domain
Protein–protein interaction
Bioinformatics
Expression
Bombyx mori

ABSTRACT

BESS domain is a protein binding domain that can interact each other or with other domains. In this study, 323 BESS domain containing (BESS-DC) proteins were identified in 3328 proteomes. These BESS-DC genes pertain to 41 species of five phyla, most of which are arthropod insects. A BESS domain contains two α-helixes linked by a coil or β-turn. Phylogenetic tree and architecture analysis show that the BESS domain seems to generate along with the DNA-binding MADF domain. Two hundred thirty three BESS-DC genes (71.1%) contain at least one MADF domain, while 59 genes (18.2%) had only the BESS domain. In addition to BESS and MADF domains, some of genes also contain other ligand binding domains, such as DAO, DUS and NAD_C. Nineteen genes (5.8%) are associated with other DNA binding domains, such as Myb and BED. The BESS-DC genes can be divided into 17 subfamilies, eight of which have more than one clade. In *Bombyx mori*, 12 BESS-DC genes that do not contain intron in the BESS domain region were localized to eight chromosomes. Real-time PCR results showed that most of the *B. mori* BESS-DC genes highly expressed from late larval stage to adult stage. The results of sequence comparison and evolution analyses suggest a hypothesis that the BESS-DC genes may play a role in central nervous system development, long term memory and metamorphosis of insects of different phyla.

1. Introduction

A protein domain is a conserved part of a protein sequence and can function independently from the rest of the protein sequence. A domain forms a compact three-dimensional structure that usually is stable and properly folded independently. Many proteins may consist of several structural domains and a domain may appear in a variety of different proteins. Domains can serve as modules for building up large assemblies, such as virus particles or muscle fibers. Domains also can provide specific catalytic or binding sites for substrates, as found in enzymes or regulatory proteins. The domains that act as a specific binding site can be divided into several classes based on nature of the ligands they bind with, such as nucleic acid binding, protein binding or iron binding.

BESS domain is a kind of specific protein-binding domain, named after three proteins that originally defined the domain: *BE*AF32 (Boundary element associated factor 32) (Zhao et al., 1995), *S*u(var)3–7 (Reuter et al., 1990) and *S*tonewall (Clark and McKearin, 1996). A BESS domain usually contains 40 amino acid residues and composes of two or three α-helices, which might be associated with a myb/SANT HTH domain (Bhaskar and Courey, 2002). A BESS domain in a protein is involved in a variety of protein–protein interaction, for example, interaction with Dorsal and TBP-associated factors (Cutler et al., 1998). The proteins that contain BESS domain(s) are usually called BESS domain containing (BESS-DC) proteins.

BESS domain is firstly found in *Drosophila* and believed to be specific in *Drosophila* (Cutler et al., 1998; England et al., 1992). But in pfam online research and our preliminary study on *Bombyx mori*, the BESS-DC genes were also found in other species. BESS-DC proteins are involved in many biology processes, for example, nurse cells and oocyte specialization (Clark and McKearin, 1996), eye development (Duong et al., 2008), immune response (Ratnaparkhi et al., 2008), lethal hybrid rescue (Maheshwari and Barbash, 2012; Chatterjee et al., 2007; Brideau et al., 2006), chromosomes boundary (Cuvier et al., 1998; Pathak et al.,

Abbreviations: BESS, *BE*AF32 (Boundary element associated factor 32), *S*u(var)3–7 and *S*tonewall; BESS-DC, BESS-domain containing; MADF, *m*yb/SANT-like domain in *Adf*-1; NLS, nuclear localization signal; NES, nuclear export signal; HP 1, heterochromatin protein 1; qRT-PCR, quantitative real time-polymerase chain reaction; ER, endoplasmic reticulum; BLAST, The Basic Local Alignment Search Tool.

* Corresponding author.
E-mail address: qlfeng@scnu.edu.cn (Q. Feng).

http://dx.doi.org/10.1016/j.gene.2015.09.024
0378-1119/

2007) and long-term memory (DeZazzo et al., 2000). In this study, genes that encode BESS-DC proteins were identified by in silico approach in different organisms with completely sequenced proteomes (UniProt, 2010). Expression patterns of some of these BESS-DC genes in *B. mori* were analyzed to explore their possible functions.

2. Result

2.1. Identification of BESS-DC proteins in organisms with complete proteomes

Through comprehensive searches in 3328 complete proteomes deposited in Uniprot (www.uniprot.org), a total of 323 independent BESS-DC proteins encoded by 323 genes were found (Supplement Table 1) by using HMMsearch. These BESS-DC genes are derived from 41 eukaryota species of 20 families of five different phyla (Fig. 1). The BESS-DC genes from the arthropod phylum accounted for 94% of the total genes, most of which are from insecta, including the orders of Diptera, Hymenoptera, Lepidoptera, Coleoptera and Hemiptera. Eighty two percent (252) BESS-DC genes are from the order of Diptera. Thirteen genes are from the chordata phylum. Two, one and one BESS-DC genes are from echinodermata, porifera and nematode, respectively.

2.2. Phylogenetic and evolutionary analysis of the BESS-DC protein family

The predicated BESS domain sequences (37 amino acids each) were used to construct phylogenetic tree. The results show that 325 BESS domain sequences (2 genes contain more than one BESS domain) could be grouped into 17 subfamilies, out of which eight subfamilies could be further classified into several clades (Table C1).

Generally, a homologous gene from the closest species that is not analyzed is selected as the best out-group in phylogenetic tree construction. In our study, however, all the BESS-DC genes are considered as in-group genes, out-group does not exist theoretically because they were predicted from the global species with complete proteome. Therefore, the gene that meets the following two conditions was chosen as presumed out-group in our analysis: 1) a gene from a species that has a farther evolution distance than any other species; 2) the gene cannot be grouped into the other 17 subfamilies. Thus, the gene H3ILF2 from echinodermata was selected as an out-group.

Phylogenetic tree analysis indicated that BESS-1 subfamily formed a binary monophyletic group with BESS-2 subfamily (Fig. 2), both of which are from Diptera and Lepidoptera (Table C1). Similarly, BESS-4 and BESS-5 subfamilies formed another binary monophyletic group, both of which are from Diptera. The BESS-3 subfamily is also from Diptera and might be a common ancestor for BESS-1/BESS-2 and BESS-4/BESS-5 subfamily groups. The above subfamilies form a ternary tree with BESS-6 and BESS-7 subfamilies. BESS-8, BESS-9 and BESS-10 subfamilies form a ternary tree with a bootstrap value of 62 and these three subfamilies are from Diptera. Most of the members from the above 10 subfamilies (BESS-1 to BESS-10, except BESS-6) are found in either Diptera or Lepidoptera. The members of BESS-6 are present in Coleoptera and Lepidoptera. The members of the subfamily BESS-11 were found in Diptera, Lepidoptera, Coleoptera and Hymenoptera. The

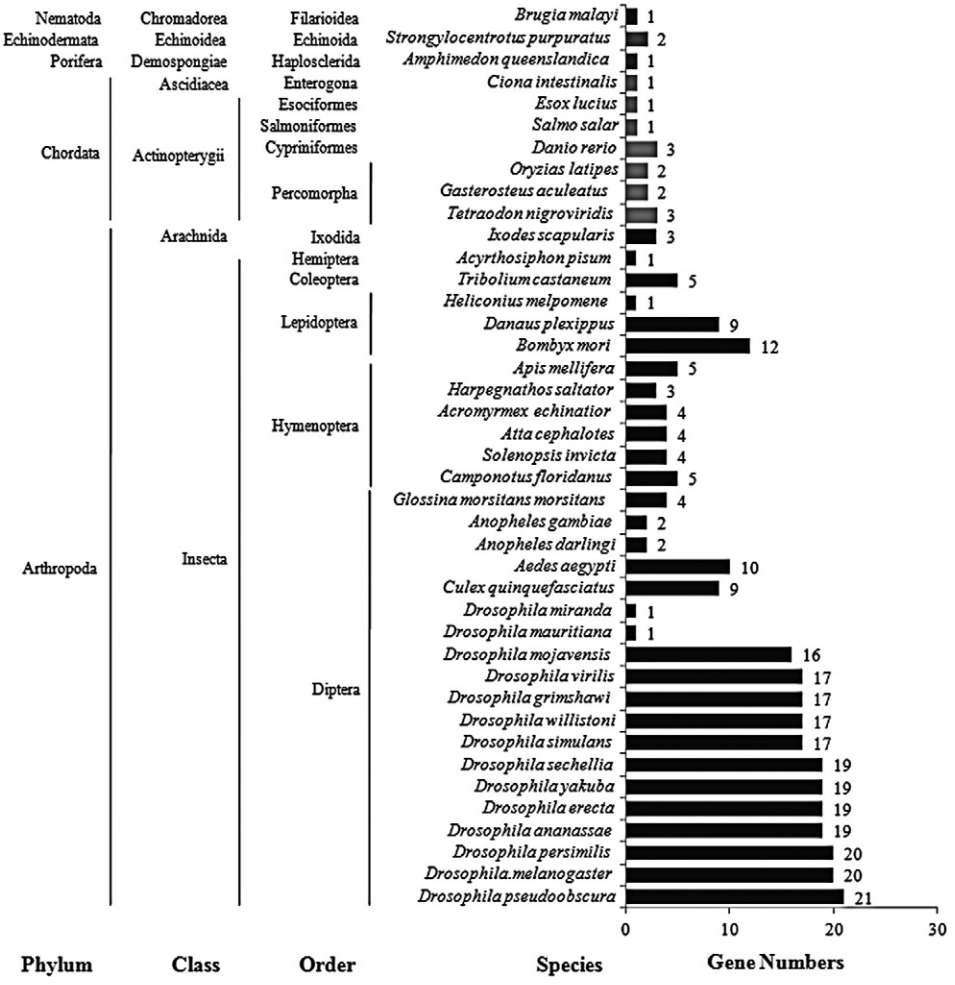

Fig. 1. Distribution of BESS-DC genes. BESS domain containing genes distribute to five phyla of eukaryota, seven different classes, 14 orders and 41 species. Most (94%) of the BESS-DC genes belong to arthropoda.