Indian Journal of Biotechnology
Vol 7, April 2008, pp 183-187

Characterization and genomic organization of esterase gene in silkworm, *Bombyx mori* L.

K M Ponnuvel*, K Ashok Kumar, D Velu, P Somasundaram, R K Sinha and C K Kamble

Biotechnology Laboratory, Central Sericulture Germplasm Resources Centre, Thally Road, Hosur 635 109, India

Received 12 May 2006; revised 18 June 2007; accepted 6 October 2007

N-terminal amino acid sequence of esterase gene was blast searched with *Bombyx mori* EST (expressed sequence tag) database. EST clone fbpv0006 showed 95% homology with N-terminal sequence of esterase protein. Further, the genomic organization of exons and introns were identified using the EST clone sequence. The results on genomic organization of esterase gene indicated that the blood esterase gene possesses two exons with varying length of 192 and 524 bp, and a long intron of 2124 bp length present between these two exons. The primers were designed to the intron and exon regions and the fragments were PCR amplified using genomic DNA from silk moth as template. The amplicon showed a mol wt 799 bp in the intron region and 547 bp covering the exon regions. When the esterase genes from two multivoltine (Pure Mysore & PMX) and bivoltine (NB4D2 & CSR-19) races, possessing contrasting characters to thermal tolerance, were PCR amplified and products were sequenced. The two sequences showed 97% homology with 3% mismatch through pair blast analysis. Furthermore, per cent identity per exon, number of gaps per exon, overall per cent identity, per cent coverage of the mRNA, number of splice donor sites, presence or absence of splice donor and acceptor sites for each exon were inferred through spidey programme. Also, the phylogenetic relationship of *B. mori* esterase gene sequence was compared with other organisms through Clustal W analysis and a dendogram was projected. The projection showed that majority of insect esterases formed a major separate cluster, while *B. mori* esterase clustered with aphid insect esterases.

Keywords: *Bombyx mori*, exons, introns, BLAST, cDNA

Introduction

The domesticated silkworm, *Bombyx mori* is one of the genetically well-characterized insects, next only to the fruit fly *Drosophila*, and recently has emerged as lepidopteron molecular model system[1]. The genetic studies of this species include more than 400 mutations, which have been mapped to >200 loci, comprising of 28 linkage groups[2,3]. Mita *et al*[4] analyzed many cDNA libraries prepared from various tissues and at different developmental stages that cover almost entire set of *Bombyx* genes, comprising 35,000 expressed sequence tags (ESTs) from 36 cDNA libraries. By virtue of a long history of rearing for commercial purpose, silkworm has been the subject of research interest, resulting in careful collection, cataloguing and maintenance of various silkworm genetic stocks of considerable scientific and economic interest. A vast array of distinct geographical races and inbred lines are available that represent variations for a number of qualitative and quantitative traits of basic biological and economic interests, such as silk quality, fecundity, pathogen resistance and heat tolerance.

In insects, esterases have been implicated in resistance to organophosphate insecticides and their banding patterns on electrophoretic gels have been shown to provide an extremely reliable method for determining biotypes[5]. Based on the substrate inhibitor specificity, esterases are classified into three groups, viz. carboxylesterases (CE), arylesterases (ArE) and cholineesterase (ChE)[6]. CE isozyme acts vigorously on various simple esters of low molecular size, but slowly on phenolic esters, and has an important role in synthesis and transport of cuticular wax. In addition, it also has a role in the detoxification of substances introduced from the external environment, such as insecticides[7]. Insecticidal resistance of CE is attributed to the fact that it acts as immune defense molecule against bacteria and degrades insecticides[8]. ArE is a small group of isozymes which function in the degradation of aromatic esters including phenolic esters. On the other hand, ChE is an important regulatory enzyme that is responsible for controlling the neural transmission on synapses, where it acts to hydrolyse the excitatory of neuro transmitter, acetylcholine.

*Author for correspondence:
Tel: 91-4344-221148; Fax: 91-4344-220520
E-mail: kmpvel@yahoo.com

Also ChE is targeting on organo-phosphorus and carbamate insecticides, as these toxic compounds readily inhibit it.

Arai *et al* [9] has purified and characterized major esterase BesB from the haemolymph of *B. mori*. Though, esterases are primarily involved in digestion of nutrition materials and reproduction, the relationship between thermotolerance and heat-stable esterase has also been studied. The biological organization of differential gene regulation was related to variations existing in DNA sequence, mainly in the regulatory and introns of a candidate gene[3]. Therefore, in the present study, organization of esterase gene and variation in the esterase gene sequences were analyzed in selected thermo-tolerant and thermo-susceptible silkworm races.

Materials and Methods

Silkworm Races

B. mori races, Pure Mysore, PMX, NB_4D_2 and CSR 19, which have shown contrasting response to thermal stress, were selected in the present study[10]. The former two races are multivoltine, while the later two come under bivoltine group.

DNA Sequences and Database Retrieval

B. mori esterase has already been purified and characterized, and its N-terminal sequence was determined as ESPRVTVKHGTLGSKPKTYSGY EYYFLQ[9]. In the present study, the N-terminal sequence of esterase protein was blast (tblastn) searched with 'Silkbase', a *B. mori* EST database of Tokyo Universit (http://www.ab.a.u-tokyo.ac.jp/-silkbase)[11], and corresponding EST sequence of esterase gene was retrieved. The obtained full-length translated protein sequence of esterase gene was analysed for the presence of esterase domain using conserved domain search.

Identification of Exons and Introns

Spidey programme, a tool for mRNA and genomic alignment, was used to elucidate the details on the number of exons, introns, donor, acceptor sites and the gaps and the length of exons in the esterase gene of *B. mori*[12]. In this analysis, all parameters were default unless otherwise specified. Spidey takes a single genomic sequence and a set of mRNA accession numbers or FASTA sequences as input. All processing was done at the rate of one mRNA sequence at a time.

The data was examined carefully to get the per cent identity per exon, the number of gaps per exon, the overall per cent identity, the percent coverage of the mRNA, presence of an aligning or non-aligning poly-A tail, number of splice donor sites, the presence or absence of splice donor and acceptor sites for each exon.

Primer Designing

The up and down gene-specific primers were designed for available gene sequences of esterase, using the software programme of primer3 (http://frodo.wi.mit.edu/cgibin/primer3). The primers were designed for the exon as well as intron region of esterase gene.

PCR Condition and Analysis of Amplified Product

The genomic DNA was isolated from silk moths using standard protocols[13] and used as template DNA in PCR reactions. The reaction was done in an MJ research thermal cycler, PTC200, using 20 μL reaction mixture containing 50-100 ng of genomic DNA as template, 2.0 μL of 10× PCR buffer (MBI Fermentas), 0.2 m*M* dNTPs, 1.5 m*M* $MgCl_2$, 66 ng of forward and reverse primer each and 0.3 U of *Taq* DNA polymerase (MBI Fermentas). The PCR reaction was performed with two sets of primers (intron region 792F & 1555R and exon region 2086F & 3305R). The PCR schedule was 94°C for 3 min, followed by 30 cycles of 94°C for 30 sec, 56°C for 30 sec, 72°C for 2 min and a final extension of 7 min at 72°C. The PCR products were resolved on 1.5% agarose gel in Tris-acetic acid/EDTA buffer (1× TAE) and electrophoresis was carried out with a constant voltage of 80 V in parallel with molecular weight markers[13].

DNA Sequencing

The PCR amplified products were purified through Gel-spin column (Bangalore Genei). The sequencing primers were gene specific positive sense primer. The purified DNA was diluted with TE buffer (*p*H 8.0) to adjust the concentration. The amplified DNA fragment was cloned in TA cloning vector using manufacturer's protocol (Bangalore Genei). The sequencing reaction was performed using M13 and gene specific primer at Bangalore Genei.

Results and Discussion

The molecular mass of blood esterase of *B. mori* was estimated as 58000 kDa and N-terminal amino acid sequence was identified as SPRVTVKHGT-LGSKPKTYSGYEYYFLQ by Arai *et al*[9]. In the present study, the N-terminal sequence was blast

(tblastn) searched with *B. mori* EST database (http://www.ab.a.u-tokyo.ac.jp/silkbase). It was found that clone fbpv 0006 had maximum homology with N-terminal protein sequence of the blood esterase. The complete cDNA sequence of this clone was translated into protein (http://www.expasy.org) and the protein was analysed for the presence of esterase domain using conserved domain search programme (http://www.ncbi.nih.nlm.gov). Two prominent domains of esterase, lipase and carboxyl esterase were predicted in the putative translated protein sequence of *B. mori* (Fig. 1). These esterase and lipase enzymes act on carboxylic esters and the catalytic apparatus involves three residues (catalytic triad)—a serine, a glutamate and a histidine. These catalytic residues are responsible for the nucleophilic attack on the carbonyl carbon atom of the ester bond. In contrast with other α/β hydrolase fold family members, p-nitrobenzyl esterase and acetylcholine esterase have a Glu instead of Asp at the active site carboxylate[14]. The corresponding cDNA clone showed maximum homology to esterase protein of *Schizaphis graminum* (aphid). These findings indicate that cDNA is esterase gene of *B. mori* and this esterase belongs to the family of carboxylic ester hydrolases.

The available cDNA sequence (730 bp) of blood esterase gene was compared with *B. mori* genomic sequence by basic alignment search tool (BLAST). It was found that a single genomic contig (Acc. No. BAAB01009617) having the length of 3562 bp had 91.8% homology with the esterase gene (score e=0). The contig was analysed for genomic organization of the esterase gene. Based on spidey analysis, the maximum length of exon (523 bp) was found in the region of 2806-3329 bp (exon-II) in the genomic contig and the minimum exon fragment length of 192 bp (exon-I) was found at 5′ region of the contig

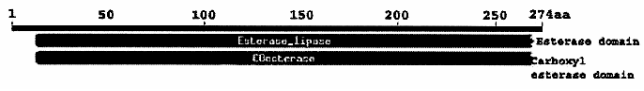

Fig. 1—Conserved functional domains of esterase enzyme in *B. mori*

(Fig. 2) There were 7 to 52 mismatches in the exon-I and exon- II regions, and 4 gapes were found in the genomic DNA (Table 1). It indicates that the above contig of genomic DNA transcribed the corresponding mRNA of the gene. The untranslated region (UTR) analysis of full length cDNA showed 5′ UTR of 118 nucleotide (nt) length and 3′ UTR of 21 nt length (Table 2).

The primers were designed for the predicted exons as well as intron regions of esterase genomic contig (Table 3). The PCR was performed with two sets of primers, which were binding site at intron region (792F-1555R) and exon region (2806F-3305R) of the genomic contig. The intron and exon regions were amplified with sequence specific primers. As a result two PCR fragments with different molecular size were observed. The primers designed in intron region yield the PCR product of 763 bp, whereas the exon region primers yield 532 bp (Fig. 3). Of four silkworm races selected for the identification of esterase gene, two belong to multivoltine and the other two to bivoltine group. Among two multivoltine races, Pure Mysore was selected as thermo-tolerant race and PMX as thermo-susceptible race. Similarly, NB4D2 (bivoltine) race was selected as thermo-susceptible breed and CSR-19 as thermo-tolerant breed. The esterase gene was amplified using gene specific primers and results indicate that there was no difference in the size of the PCR fragments in all the four races selected. The esterase gene sequences of thermo-tolerant as well as thermo-susceptible breeds were compared using pair BLAST analysis. The pair BLAST analysis showed 97% homology between these two races with 3% mismatch that indicates that this sequence variation probably contributes for the differential expression of esterase gene in *B. mori*. This hypothesis needs to be confirmed through northern blot/RT-PCR analysis in future studies. It was observed that the characteristic feature of carboxylesterase gene is having two exons with intervening introns of 2124 bp length (Fig. 2). Usually, long introns tend to accumulate insertions, such as transposable elements and repetitive

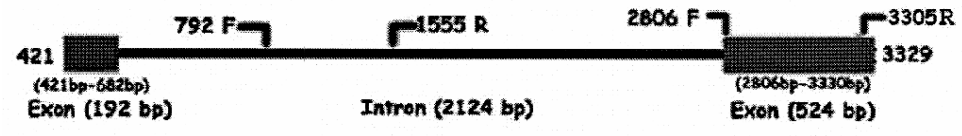

Fig. 2—Genomic organization of *B. mori* esterase gene in the genomic contig (Acc. No.: BAAB01009617). R and F indicate binding site of reverse and forward primers

Table 1—Details on genomic co-ordinates, exons, percent identity, gaps, donor sites of *B. mori* esterase gene

Exon	Position in genomic coordinate	Position in mRNA coordinate	Exon fragment length	Identity (%)	Mismatches	Gaps	Donor Site	Acceptor site
I	491-682	1-192	192	96.4	7	0	D	-
II	2806-3329	193-716	524	90.1	52	4	-	A

Table 2—Details of cDNA, genomic contig, conserved domain and ORF in *B. mori* carboxyl esterase gene

Target genes	Total cDNA length (bp)	Corresponding genomic DNA contig length (bp)	Domain	5′ UTR Position/length bp	3′ UTR Position/length bp	Amino acid length (AA)
Carboxylesterase gene	730	3562	Esterase lipase Carboxyl esterase Carboxyl esterase type B	1-118 (118)	21(709-730)	196

Table 3—Details of genomic contig, primer sequence and its binding location in the contig and amplicon size of *B. mori* esterase gene

No.	Target genes	Primer sequence	Length	Primer binding location in the contig (Acc.No: BAAB01009617)	Amplicon size
1	Intron 792 F	CTCATCCACAGCGTTCTCAA	20	792 bp	763 bp
	Intron 1555 R	AGTCCACAACCACCAGAGTCTT	20	1555 bp	
2	Exon 2806 F	TGTTGAAATAATCCTCTAGACGAAGG	22	2806 bp	532 bp
	Exon 3305 R	TGTTGAAATAATCCTCTAGACGAAGG	26	3338 bp	

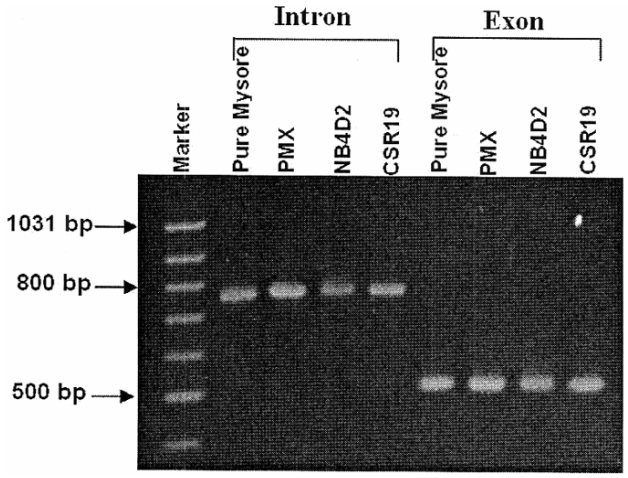

Fig. 3—PCR amplified product of different silkworm races esterase gene in 1.5% agarose gel

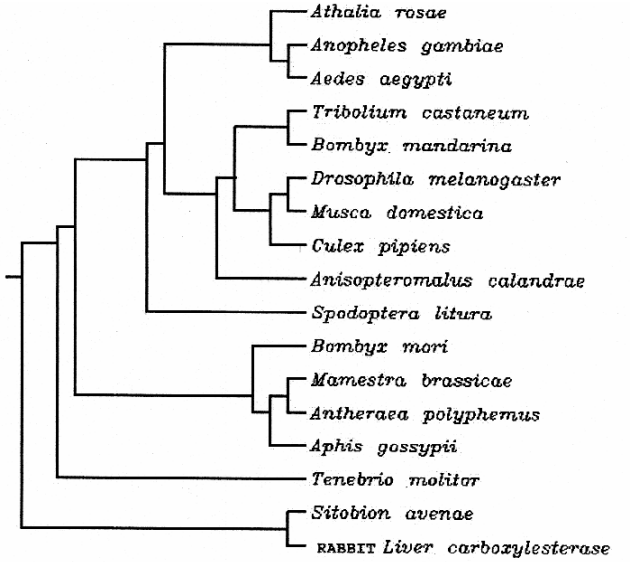

Fig. 4—Phylogenetic relationship of *B. mori* esterase gene with other organisms

sequences, and tend to increase polymorphism in term of size than short introns[15]. Thus, Spidey analysis was useful in determining the actual length of coding and non-coding regions of esterase genes. In the present study, UTR in mRNA sequence was determined and esterase gene possessed precise 5′ and 3′ UTRs.

The 14 kb region of genomic DNA encoding the wild-type *white* eye (*w*1) colour gene from the medfly, *Ceratitis capitata* was cloned and characterized at the molecular level[16]. Comparison of the intron-exon organization of this locus among several dipteran insects revealed distinct organizational patterns that are consistent with the phylogenetic relationships of these flies and the

dendrogram of the predicted primary amino acid sequence of the *white* loci. Similarly, the phylogenetic relationship of *B. mori* esterase gene sequence was compared with other organisms through Clustal W analysis. The results indicate that majority of insect esterases formed a major separate cluster, while *B. mori* esterase clustered with aphid insect esterases. The rabbit liver carboxylesterase was placed as an out group member and grouped with grain aphid (*Sitobion avenae)* in the phylogenetic tree (Fig. 4). These findings indicate that the esterase gene was conserved across the insect taxa.

Goldsmith *et al*[18] reported that UTR posses conserved regions and microsatellite domains, which are useful to determine the genetic variability in silkworm germplasm stocks. Generally, there are greater numbers of introns in *Bombyx* genes (and those of other lepidopterans) than in *Drosophila*. Komoto *et al*[18] compared the exon-intron structures of XDH genes among metazoans and pointed out the possibility of introns loss in *Drosophila* and other metazoans. It is common to find that the introns of *Bombyx* genes were longer than orthologues in *Drosophila*. The *Bombyx* specific situations cause problems in gene finding, because it is difficult to identify the exon-intron boundaries and promoters correctly and to annotate the genes[18]. In the present study, promoter regions, exons and introns in the genomic region were identified. In turn, this information would help for molecular characterization of silkworm germplasm with unique characters.

References

1 Nagaraju J & Goldsmith M R, Silkworm genomics—Progress and prospectus, *Cur Sci,* 83 (2002) 415-425.

2 Yasukochi Y, A dense genetic map of the silkworm, *Bombyx mori* covering all chromosomes based on 1018 molecular markers, *Genetics,* 150 (1998) 1513-1525.

3 Mita K, Kasahara M, Sasaki S, Nagayasu Y *et al*, The genome sequence of silkworm, *Bombyx mori*, *DNA Res,* 11 (2004) 27-35.

4 Mita K, Morimyo M, Okano K, Koike Y, Nohata J *et al*, The construction of an EST database for *Bombyx mori* and its application, *Proc Natl Acad Sci, USA*, 100 (2003) 14121-14126.

5 Balakirev E S, Chechetkin V R, Lobzini V V & Ayala F J, DNA polymorphism in the β-esterase gene cluster of *Drosophila melanogaster*, *Genetics*, 164 (2003) 533-544.

6 Yoo C M, Bak C B & Lee H C, Substrate and inhibitor specificities of esterase in *Lucilia illustris*, *Korean J Zool*, 39 (1996) 190-197.

7 Yoshitake N & Eguchi M, Distribution of blood esterase types in various strains of the silkworm *B. mori* L (in Japanse), *Jpn Seric Sci*, 34 (1965) 95-101.

8 Shiotsuki T & Kato Y, Induction of carboxylesterase isozymes in *Bombyx mori* by *Escherichia coli* infection, *Insect Biochem Mol Biol*, 29 (1999) 731-736.

9 Arai H, Okido, T, Fujii H & Doira H, Purification and characterization of a major esterase BesB from haemolymph of the silkworm, *Bombyx mori*, *J Seric Sci Jpn,* 89 (2000) 121-130.

10 Koundinya P R, Kumaresan P, Sinha R K & Thangavelu K, Screening of promising germplasm of polyvoltine silkworm *(Bombyx mori)* for thermo-tolerance, *Indian J Seric*, 42 (2003) 67-71.

11 Altschul M D, Madden T I, Schaffer A A, Zhang J, Zhang Z *et al*, Gapped BLAST and PSI-BLAST: A new generation of protein database search programs, *Nucleic Acids Res*, 25 (1997) 3389-3402.

12 Wheelan S J, Church D M & Ostell J M, Spidey—A tool for mRNA to genomic alignments, *Genome Res,* 11 (2001) 1952-1957.

13 Sambrook J, Fritsch E F & Maniatis T, *Molecular cloning: A laboratory manual*, 2nd ed, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY) 1989.

14 Chattopadhyay G K, Sengupta A K, Verma S K, Sen S K & Saratchandra B, Esterase isozyme polymorphism, specific and non-specific esterase, syngenic lines development and natural occurrence of a thermostable esterase in the tropical silkworm *Bombyx mori, Insect Biochem Mol Biol*, 31 (2001) 1191-1199.

15 Benson G, Tandem repeats finder: A program to analyze DNA sequences, *Nucleic Acids Res*, 27 (1999) 573-580.

16 Gomulski L M, Pittas R J, Costa S, Saccone G, Torti C *et al*, Genomic organization and characterization of the white locus of the *Ceratitis capitata*, *Genetics*, 157 (2001) 1245-1255.

17 Goldsmith M R, Shimada T & Abe H, Genetics and genomics of the silkworm *Bombyx mori*, *Annu Rev Entomol*, 50 (2005) 71-100.

18 Komoto N, Yakuchiro K & Tamura T, Structure and expression of tandemly duplicated xanthine dehydrogenase genes of the silkworm *Bombyx mori, Insect Mol Biol,* 8 (1999) 73-76.