Title

Cis-regulatory elements of the cholinergic gene locus in the silkworm *Bombyx mori*

Authors

Kota Banzai^{a, #, *}, Susumu Izumi^a

Affiliations

^a Department of Biological Sciences, Kanagawa University, Hiratsuka, Kanagawa, Japan.

[#] Current Address: Department of Cellular Biology, University of Georgia, Athens, GA, USA

^{*} Corresponding author. E-mail: kotabanzai.bio@gmail.com

Running title

Expression regulation of *Bombyx ChAT* gene

Key words

Cholinergic neuron, gene expression, *ChAT*, *Bombyx mori*

Author contributions

K.B. performed experiments and analyzed data. K.B. and I.S. designed the research and wrote the article.

Acknowledgements

We thank K. Bui and Y. Ishikawa for comments on this manuscript. This work was supported in part by the grant of Research Institute for Integrated Science from Kanagawa University and the Kanagawa University Grant for Joint Research.

Competing financial interests

The authors declare no competing financial interests.

ABSTRACT

Genes of *choline acetyltransferase (ChAT)* and *vesicular acetylcholine transporter (VAChT)* are encoded in the same gene locus, called the cholinergic gene locus. They are essential in cholinergic neurons to maintain their functional phenotype. The genomic structure of the cholinergic gene locus is conserved among invertebrates to mammals. However, it is still inconclusive how cholinergic genes express only in cholinergic neurons in insects. In this study, we analyzed the upstream sequence of cholinergic gene locus in the silkworm *Bombyx mori* to identify specific *cis*-regulatory regions. We found multiple enhancer regions that are localized within 1 kb upstream of the cholinergic gene locus. The combination of promoter assays using small deletions and bioinformatic analysis among insect species illuminates two conserved sequences in the *cis*-regulatory region: TGACGTA and CCAAT, which are known as the cAMP response element and CAAT box, respectively. We found that dibutyryl-cAMP, an analog of cAMP, influences the expression of *ChAT*. Tissue-specific expression analysis of transcriptional factors showed potential candidates that control the cholinergic gene locus expression. Our investigation provides new insight into the regulation mechanism of cholinergic neuron-specific gene machinery in this lepidopteran insect.

INTRODUCTION

Various types of neurons exist in the nervous system. The classification of each neuron is defined by its neurotransmitter phenotype; the expression of the specific genes encoding proteins for synthesis, storage, and release of neurotransmitters. Acetylcholine (ACh) is a major excitatory neurotransmitter in the central nervous system of insects (Gupta, 1987) and the neuromuscular junction in vertebrates (Krnjevic and Miledi, 1958). In insects, cholinergic neurons play a role in wing movements, locomotion, learning, and memory (Gauglitz and Pfluger, 2001, El Hassani et al., 2008). In cholinergic neurons, choline acetyltransferase (ChAT) and vesicular acetylcholine transporter (VACHT) have been well known as marker proteins because their molecular functions are essential for maintaining cholinergic phenotypes (Ichikawa et al., 1997). ChAT works as a catalytic enzyme of the ACh from precursor chemicals. VACHT functions to transport and store ACh into the synaptic vesicles (Song et al., 1997). Interestingly, their functions are completely different, yet they are expressed from the same cholinergic gene locus. In all animal species so far examined, *ChAT* and *VACHT* genes share exons that usually encode 5'-UTR. This suggests that the gene regulatory mechanism for the cholinergic gene locus is conserved among animal species. However, it is still not conclusive how cholinergic neurons express their own essential genes. To address this point, the elucidating the transcriptional factors that works for specifying neuronal identity is pivotal.

The gene regulation of the cholinergic gene locus has been well-studied in *Drosophila melanogaster* as a model system. Immunohistochemical assays and transgenic fly lines containing the cholinergic promoter fused to a GFP determined the distribution of cholinergic neurons in *Drosophila* nervous systems (Yasuyama and Salvaterra, 1999; Salvaterra and Kitamoto, 2001). The reporter gene analysis tested various upstream regions of the cholinergic gene locus and it indicated that the different sets of *cis*-regulatory elements worked for specific expression patterns in the nervous system (Kitamoto et al., 1992). The proximal region (-0.8 kb to -0.5 kb of the cholinergic upstream sequence) contained the necessary elements for *ChAT* expression in the central nervous system (CNS). The same region is sufficient to rescue the lethal phenotype of the *Drosophila Cha* mutant (Kitamoto and Salvaterra, 1993). POU-19/pdm-1 has been identified as a potential transcriptional factor to bind to a 22 bp sequence that is located in the proximal region (Kitamoto and Salvaterra, 1995). In contrast to the proximal regulatory region for the expression in CNS, the distal part of the upstream

sequence is required for the expression in the peripheral nervous system (Kitamoto et al., 1992). Abnormal chemosensory jump 6 (*Acj6*), has been identified as a transcriptional factor that positively controls the *ChAT* mRNA expression in primary olfactory neurons (Lee and Salvaterra, 2002). Transcriptional regulatory mechanisms of the cholinergic gene locus have been previously studied. However, what transcriptional factor works for the CNS specific gene expression in insect species is still not completely understood. Identification of distinct *cis*-regulatory elements is a major step in understanding the transcriptional mechanism regulating target genes. Classically, *in vivo* reporter assays using deletions or mutations in the promoter of a target gene is utilized to identify *cis*-regulatory elements. But such experiments have been inapplicable for non-model organisms like the silkworm due to the difficulty of making and maintaining transgenic lines. The baculovirus-introduction method was developed and adapted to study *cis*-regulatory elements in arthropods (Lu et al., 2005, Shiomi et al, 2003, Chen et al., 2020). This technology allows us to investigate the *in vivo* reporter assay in a non-model organism such as silkworm. Our previous report characterized the cholinergic gene locus in the silkworm, *Bombyx mori* (Banzai et al., 2015). We identified that the 2 kb upstream region of the cholinergic gene locus contained essential regulatory elements for expression in the CNS. In this paper, by the promoter assay based on the baculovirus-introduction method, we demonstrate that multiple *cis*-regulatory regions control the *ChAT* gene expression. Our investigation provides new insight into the regulatory mechanism underlying cholinergic neuron-specific gene machinery in the lepidopteran insect.

RESULTS

Promoter assay using sequential truncations illuminates the *cis*-regulatory regions

As reported in our previous paper (Banzai et al., 2015), the reporter protein, EGFP under the control of 2kb upstream region of *Bombyx* cholinergic gene locus expresses only in the CNS (Fig. 1B and 1B'), but not other tissues. Based on previous results, we constructed recombinant baculoviruses containing sequentially truncated promoters to investigate the *cis*-regulatory sequences (Fig. 1A). The EGFP signals were detected from the -2000 promoter to the -200 promoter (Fig. 1B-G), but not the -100, the -50, and the 0 promoters (Fig. 1H-J). The reporter signals from the sequentially truncated constructs decreased gradually depending on the length of the promoter region. Despite reducing the EGFP signals, the expression patterns seemed to be restricted to a specific subset of neurons. These results suggest that the upstream region between -2000 promoter and -200 promoter functions to promoting expression level and does not regulate the expression specificity. We also quantified the activities of each promoter construct (Fig. 1K). The -2000 promoter to the -761 promoter showed a similar activity level. In contrast, the promoter activity decreased by approximately 40% in the -500 and the -300 promoters. In the -200 promoter, the promoter activities declined by about 65% compared to the -2000 promoter. The -100 and the -50 promoters exhibit similar levels to the 0 promoter, showing almost no promoter activity. These results from the sequentially truncated promoters indicate that three regions between -761 and -500, -300 and -200, -200 and -100 are involved in the gene expression regulation of the *Bombyx* cholinergic gene locus.

Conserved motifs among lepidopteran species

To find the DNA motifs that enhance the cholinergic gene expression, the identification of conserved sequences among different species is informative (ref). To do this among *B. mori* and other moths, we characterized the cholinergic gene loci from two lepidopteran species: the monarch butterfly, *Danaus plexippus*, and the postman butterfly, *Heliconius melpomene*, because their whole genomic sequence data have been published (Zhan et al. 2011, Heliconius Genome Consortium 2012). In the case of *D. plexippus*, putative *ChAT* and *VACHT* genes have been deposited in NCBI as OWR51509.1 and OWR51511.11, which are coding sequences. However, those two deposited genes don't contain the shared

regions, in which mRNAs of *ChAT* and *VACHT* have shared sequences in all species so far examined. This fact indicates that deposited sequences in *D. plexippus* didn't cover the complete sequence of *ChAT* and *VACHT* including 5' and 3' untranslated region (UTR). To determine the transcription initiation point of the cholinergic gene locus from *D. plexippus*, we identified the cholinergic gene exons of *D. plexippus* using the BLAST search using *Bombyx* exons as the query. We found the high similarity sequence (73.1%, 38/52) in the upstream region of *D. plexippus* OWR51509.1 against the first exon of *Bombyx* cholinergic gene locus (Fig. 2A and 2B). In the case of *H. melpomene*, we characterized its whole cholinergic gene locus by a similar procedure. The cholinergic gene locus of *H. melpomene* consists of 14 exons. As well as the *Bombyx* cholinergic gene locus, the first and second exons of *H. melpomene* encoded only 5'-UTR. The third exon encodes *VACHT*. Exons after the fourth exon are specific for the *ChAT* (Fig. 2A). The putative first exon has a high similarity (62.3%, 33/53) against the first exon of the *Bombyx* cholinergic gene locus (Fig. 2B).

Next, we compared upstream DNA sequences of these three lepidopteran species to find conserved sequences that had the potential to regulate gene expression. For this comparison, we used the Mussa program that visualizes conserved regions across multiple species. Our promoter assay indicated that the 1 kb upstream region of the *Bombyx* cholinergic gene locus is sufficient to control gene expression specifically in the CNS. We therefore selected the 1 kb upstream sequences of three lepidopteran species for this comparison. We found three conserved motifs located within -200 bp upstream of cholinergic gene loci (E1, E2, and E3, Fig. 2C). To predict possible transcriptional factors that bind to the conserved sequence, we investigated three conserved sequences by the STAMP programs. As shown in Fig. 2D, the most distal element E1 contains a consensus sequence (YTAWWWWTAR) of Myocyte-specific enhancer factor 2A (Mef2A). The center element E2 contains a consensus sequence (CATTWAANTT) of *zerknüllt* (*zen*). The most proximal E3 contains a consensus sequence (TGACGTCA) of cAMP response element-binding protein (CREB).

Promoter assay using small deletions reveals *cis*-regulatory elements

To validate whether the conserved sequences are required for the gene expression of *Bombyx* cholinergic gene locus, we tested promoter activities by introducing small deletions into the promoter sequence. As shown in Fig. 3A, the deletion between -200 to -150, including the E1 element causes a 65% reduction

of the GFP signal compared to the -2000 promoter. Moreover, the deletion between -200 to -100, including E1 and E2 elements, causes a more reduction than a single deletion of the E1 element. These results suggest that E1 and E2 elements work together to regulate gene expression.

Next, we analyzed the promoter activity of the E3 element which is located most proximally. A deletion promoter, del 100-50, showed a 90% reduction of promoter activity (Fig. 3B). To narrow down the critical sequence for the gene expression regulation, we deleted DNA sequences in every 10 bp between 100 bp and 50 bp of the promoter region. As shown in Fig. 3B, the del 90-80 promoter excluding the E3 element shows a 45% reduction of promoter activity. This result supports the idea that the consensus sequence (TGACGTCA) of CREB has an important role in the *ChAT* gene expression. Additionally, two deletion promoters, del 70-60 and del 60-50 caused drastic reductions of promoter activity, 65% and 90%, respectively (Fig. 3B). However, between -60 to -50, we could not observe any preserved motifs among lepidopteran species.

cAMP modulates the expression level of *Bombyx ChAT* gene

Our promoter assay using small deletion mutants revealed that the promoter region -100 to -50 consists of three different elements. We predict that the identified regions should be conserved not only moth but also other insect species if they are critical for gene expression. Thus, we expanded the bioinformatic comparison of the promoter region among multiple insect species. As shown in Fig. 4A, the CRE sequence (TGACGTCA) in -90 to -80 and a five-nucleotide sequence (CCAAT) in -70 to -60 were conserved among seven insect species. Interestingly, the distance between the CRE sequence and the CCAAT was within 21 nucleotides of each other in these six insect species (Fig. 4A). This indicates us that the combination of two elements work synergistically. It has been reported that CREB forms heterodimers with C/EBP (CCAAT-enhancer-binding protein) *in vitro* (Vinson et al., 1993). We could not find any conserved motif in the most proximal part (-60 to -50).

We decided to focus on CREB based on our promoter assay and the bioinformatic analysis. Cytoplasmic cAMP works as a second messenger, which activates CREB via the protein kinase A (PKA) signaling pathway (Alberini, 2009). A previous study demonstrated that the addition of analog of cAMP, dibutyryl cAMP (dbcAMP), in mammalian PC12 cells (naturally expresses both ChAT and VAcHT) increases the amount of ChAT activity (Shimojo et al., 1998). Thus, we

examined the effect of dbcAMP on the *Bombyx ChAT* gene expression in the *ex vivo* brain. As shown in Fig. 4B, the different expression level of the *ChAT* gene between brains cultured with or without dbcAMP was observed 24 hours after starting culture, but not other time points. This result indicates that cAMP works negatively for *Bombyx ChAT* gene expression. However, this result is not consistent with the promoter assay that we predicted that CREB, a downstream transcriptional factor of cAMP response, works as a positive enhancer (Fig. 3B). To test the function of CREB in *ChAT* gene expression more directly, we designed the experiment that we express active and repressive forms of CREB in larvae through recombinant baculoviruses and measure the *ChAT* gene expression by quantitative PCR. It is well known that serin 133 (S133) in human CREB gets phosphorylated by cAMP stimulation and the phosphorylation is required for the interaction with co-transcriptional factors (Gonzalez and Montminy, 1989, Suzuki et al., 2011). The mutation into the amino acid sequences including S133, RRPSYR to DIEDML, makes CREB proteins constitutively active (Cardinaux et al., 2000). In *Drosophila*, an artificially modified CREB that lacks the leucine zipper domain functions as dominant-negative (Yin et al., 1994). We expressed active and repressive forms of CREB ubiquitously by Actin promoter or ChAT-positive neurons by the ChAT promoter. As shown in Fig. 4C, *ChAT* gene expressions in control, active or repressive forms of CREB were comparable level. This data suggests that CREB is not a regulatory protein for this locus, and other transcriptional factors are involved in the expression of cholinergic gene locus.

Gene expression profile of candidate transcriptional factors

We narrowed down the candidates for activation of *ChAT* expression using the STAMP program to uncover which transcriptional factors may bind the upstream region of the *Bombyx* cholinergic gene locus. We selected three regions (-761 to -500, -300 to -200, and -200 to -100) identified in our promoter activity assays. As summarized in Fig. 5A, the STMAP program showed sixteen genes for -761_-500, six genes for -300_-200, and two genes for -200_-100. We tested the expression pattern of those predicted transcriptional factors in six tissues. Fifteen genes out of the candidate twenty genes were expressed in the brain (Fig. 5B). Moreover, four genes, *deformed* (*Dfd*), *even-skipped* (*eve*) *tailless* (*tll*), and *twin of eyeless* (*toy*), were expressed in the brain and not in any other tissues.

DISCUSSION

We discovered four *cis*-regulatory elements on the *ChAT* gene promoter region. The most distal region spans from -761 to -500. We could not find any significant binding site based on the bioinformatic analysis. Further promoter assays using different sets of truncations and mutations of promoter sequences is required to discover regulatory elements located in this region. We found consensus sequences for the other three regulatory elements (named as E1, E2, and E3, Fig. 2C). The E1 element has a Mef2 consensus sequence. Mef2 expresses strongly in the mushroom body from embryo to adult in the fruit fly (Crittenden et al., 2018). Besides, *Mef2* is annotated as one of the enriched genes in insulin-producing cells (IPCs) in *Drosophila* (Cao et al., 2014). Previously, we identified that *ChAT* was expressed in the majority of the CNS including mushroom body and IPCs (Banzai et al., 2015). The overlapped expression pattern of *Mef2* and *ChAT* in CNS supports that Mef2 has a strong potential to regulate cholinergic gene expression. The E2 element has zen motif, but our gene expression analysis of several tissue samples didn't show the expression of *zen* in the brain. The E3 element possesses a CRE motif where CREB transcriptional factor binds to. We tested the effect of CREB on *ChAT* gene expression by overexpression of activated or inactivated proteins. However, we did not obtain evidence that CREB influences *ChAT* gene expression. There are several methods to isolate sequence-specific binding proteins (e.g., the yeast one-hybrid assay, the electrophoretic mobility shift assays (EMSAs), and the reverse ChIP). In the future, we will apply those approaches to elucidate the regulatory proteins.

Previous reports in *Drosophila* have uncovered that -0.8 kb to -0.5 kb upstream sequence of ChAT start codon is an important regulatory region for *ChAT* gene expression in CNS (Kitamoto and Salvaterra, 1995). In addition, their DNase-I foot-printing analysis of the identified 0.3 kb revealed a 22 bp sequence that is conserved among *Drosophila* species and contains an octamer-like motif (ATTCAAAT). We analyzed but could not find the *Drosophila* identified sequence in *B. mori* and other insect species. Conversely, in *Drosophila*, we could not see the combination element of the Cre (TGACGTCA) and the CCAAT sequences observed in our study. Further advanced and optimized searching tools for transcriptional factor binding sites would help to resolve this issue.

ChAT and *VACHT* genes are essential in cholinergic neurons. However, there are

other specific genes to maintain cholinergic neuronal identity. For example, acetylcholine esterase (AChE) breaks down acetylcholine to choline and acetic acid at the synaptic cleft to tightly control the signal transduction. The choline transporter (ChT) works to re-uptake the decomposed choline to presynaptic neurons, which is critical to resynthesize acetylcholine. Acetylcholine receptors (AChRs) are responsible for the autoregulation of neurotransmitter release. A previous report using *C. elegans* has identified a transcriptional factor, Unc-3 as a terminal selector for cholinergic motor neurons by analyzing promoters from several cholinergic neuron identity genes (Kratsios et al., 2012). Similar approach in *B. mori* would be crucial to elucidate the specific regulatory machinery in the future.

Experimental procedures

Insects and cell lines

Eggs of the multivoltine strain, N4, of *B. mori* were obtained from Kyushu University (Fukuoka, Japan). Larvae were reared on an artificial diet, Silkmate II (Nihon Nosan Kogyo, Yokohama, Japan) at 25 °C under a 14-h light/10-h dark photoperiod cycle. The cell line of *Spodoptera frugiperda* Sf9 was maintained in SF900-III medium (Invitrogen) at 27 °C in a cell culture incubator.

Construction of the reporter plasmids

Different DNA fragments of the cholinergic gene locus promoters and *EGFP* ORF sequence were amplified using specific primers (Supplementary Table 1) and genomic DNA or pEGFP-1 (Clontech) as template DNA. The PCR products were inserted into a pFastBac-HTa vector (Invitrogen). The deletion plasmids were amplified by inverse PCR using specific primers (Supplementary Table 1) and circularized by self-ligation. We used KOD-FX neo (Toyobo) as DNA polymerase in our PCR reactions.

Preparation of recombinant AcNPVs and infection into larvae

Recombinant AcNPVs were prepared according to the instruction manual of the Bac-to-Bac baculovirus expression system (Invitrogen). The recombinant baculoviruses were amplified by infecting sf9 cells 2-3 times. The virus titer was determined by plaque assay (Shirata et al., 1999). The virus solution was diluted to $1-2 \times 10^6$ pfu/ml with sf900-III medium before use. Virus injection and observation of samples are followed from a previous report (Banzai et al., 2015). We used multiple animals (n = 5) for measuring the promoter activity.

Determination of putative cholinergic gene loci from insect species

We utilized two moths (*Danaus plexippus* and *Heliconius melpomene*), three mosquitoes (*Aedes aegypti*, *Anopheles albimanus* and *Anopheles arabiensis*), and one beetle (*Tribolium castaneum*). Genome DNA sequences of insect species were obtained from online databases (Supplementary Table 2). We searched the cholinergic gene loci by BLAST of each database using data of *B. mori* cholinergic locus as the query.

Computer-assisted retrieval of the conserved DNA regions

The conserved nucleotide sequences among three lepidopteran species were

predicted by comparing their upstream regions of the cholinergic loci using the Mussa program (<http://mussa.caltech.edu/mussa>) with changing parameters of the threshold (nt) and window (nt). The consensus sequences were searched from the conserved sequences by using the STAMP (Mahony and Benos, 2007, <http://www.benoslab.pitt.edu/stamp/>).

Isolation of total RNA, Reverse transcription, and PCR

We prepared total RNAs of 6 tissues: fat body, midgut, salivary gland, malpighian tubule, epidermal cells, and brain from newly ecdysed and non-fed fifth instar larvae. We collected each tissue from at least three individual animals. Total RNA was isolated following the acid guanidinium thiocyanate-phenol-chloroform extraction method (Chomczynski and Sacchi, 1987). The total RNA (200 ng) was converted to cDNAs by the PrimeScript™ RT reagent kit with gDNA Eraser (Takara Bio) in 10-μl reaction volume. The synthesized cDNAs were diluted with 90 μl of MilliQ water (final concentration of total RNAs is 2 ng/μl).

For the PCRs, the reactions were carried out in 20-μl reaction volumes containing 0.5 μl of template cDNAs (equivalent to 1 ng of total RNA), 1×GoTaq Green Master Mix (Promega), and primers at final 0.5 μM concentration. The PCR products were separated using a 2.0% agarose gel electrophoresis and stained with ethidium bromide. Information about the primers, annealing temperatures and the number of cycles for PCR reactions is summarized in Supplementary Table 4. The primer sequence for *pho* gene was obtained from Li et al. (2012).

Effect of cAMP on the *ChAT* gene expression

We dissected out the brains from newly-ecdysed and non-fed fifth instar larvae. Those brains were washed four times with PBS (20 mM NaPB, 150 mM NaCl, pH 7.0) and two times with TC100-insect medium (Sigma Aldrich). Then, the dissected brains were stored for one hour in a TC100-insect medium containing 10% fetal bovine serum and penicillinG/streptomycin cocktail (Roche). We added the dbcAMP (Sigma Aldrich) with a final 1% concentration into the medium. The tissues were incubated at 27 °C in a cell culture incubator. We collected five cultured brains at 24, 48, 72, 96, and 120 hours after the addition of dbcAMP. For each time point, we prepared three replicates. We operated all procedures in the culture hood to prevent contamination of bacterium except the brain dissection from larvae. Total RNA isolation and cDNA synthesis were performed by the same protocol as described above.

For quantitative PCRs (qPCRs), we used THUNDERBIRD™ SYBR® qPCR Mix (Toyobo), following the manufacturer's instruction. The 20-μl reaction system contained six components: (1) 5 μl of cDNA (equivalent to 2 ng of total RNA), (2) 10 μl of THUNDERBIRD SYBR qPCR Mix, (3) 0.4 μl of 50× ROX reference dye, (4) 3.4 μl of SDW, (5) 0.6 μl of forward primer (10 μM), and (6) 0.6 μl of reverse primer (10 μM). The real-time PCR program was performed using 7300 Real-Time PCR System (Applied Biosystems) following the instrument manual. It consisted of an initial step at 95°C for 60 s, followed by 40 cycles of 95°C for 10 s and 60°C for 50 s. qRT-PCR quantitative assays of all samples were repeated three times. Primers used in PCR reactions were listed in Supplementary Table 1.

Effect of active or repressive CREB overexpression on the *ChAT* gene expression

Total RNAs were extracted from larval brains by using Isogen (Nippon). cDNA was synthesized from the total RNAs by using the PrimeScrip RT reagent Kit with gDNA Eraser (Takara bio). DNA fragments of the CREB ORF and the repressive CREB (repCREB) were amplified using specific primers (Supplementary Table 1) and the brain cDNA and cloned into pFastBac-HTa vector. The constructed plasmids were named as CREB vector and repCREB vector, respectively. Next, we constructed a plasmid for the active CREB (actCREB) by performing the inverse PCR using specific primers (Supplementary Table 1) and the CREB vector. This vector is named as actCREB vector. The DNA fragment of the Actin A3 gene promoter (-127_+542 bp, GeneBank: U49854.1) and cholinergic gene locus promoter (-2000_+30 bp) are amplified by using specific primers (Supplementary Table 1) and genomic DNA. The PCR products were inserted into both of actCREB or repCREB vectors.

We prepared and injected the recombinant virus into animals with the same protocol described above. Five days after injection, we dissected out brains from the virus-infected larvae. We collected three brains from each experimental condition. We prepared three replicates for measuring *ChAT* gene expression by qPCR. Total RNA isolation, cDNA synthesis and the qPCRs were performed by same protocol as described above.

Selection of candidate transcriptional factors

We identified candidates of transcriptional factors that may be involved in expression regulation of cholinergic gene locus by using the STAMP (Mahony

and Benos, 2007, <http://www.benoslab.pitt.edu/stamp/>). We individually characterized the four enhancer elements (-761_-500, -300_-200, -200_-100 and -100_-50) by using two datasets: flyreg (Bergman/Polland) and Fly (curated by Bergman).

Statistical analysis

All data in graphs were shown as mean $\pm$ standard deviation (SD). Significant differences ($p < 0.05$) among multiple sample groups or conditions were analyzed using one-way ANOVA followed by the Tukey HSD test. Same letter indicates no significant difference. Different letters indicate significant differences. If a non-significant difference was found between two groups, over-lapping letters are used to indicate the statistical significance.

Figure Legends

Figure 1. Promoter activity assays of sequential truncation promoters. (A) Schematic image of promoter constructs. Different length of ChAT promoter region is fused to EGFP protein. (B-J, B'-J') EGFP expression patterns of different promoters in the larval brain. Pictures with alphabet were merged views with bright field and fluorescence. Pictures with alphabet plus (') were fluorescent views. Scale bars are 100 μ m. (K) Quantitative fluorescent intensity. Data of each promoter activity are collected from multiple animals ($n = 5$). Small letters indicate significant differences ($p < 0.05$, ANOVA followed by the Tukey HSD test) among groups. If a non-significant difference was found between two groups, overlapping letters are used to indicate the statistical significance. Each bar represents the mean $\pm$ standard deviation (SD) of assays. (L) The drawing diagram shows three regions (yellow boxes) identified via promoter assays.

Figure 2. The cholinergic gene loci of three moth species. (A) Schematic representation of the cholinergic loci in *B. mori*, *D. plexippus* and *H. melpomene*. The genomic DNA is shown as a straight line. The exons common to *ChAT* and *VACHT*, specific for *VACHT*, and specific for *ChAT*, are indicated by black, striped, and white boxes, respectively. Approximate gene locus sizes (kbp) are shown on the right-hand side. (B) The DNA sequences of putative first exons among moth species. Abbreviations indicate each species; Bm: *B. mori*, Dp: *D. plexippus*, Hm: *H. melpomene*. Nucleotides conserved among three species and between two species are annotated by (:) and (.). The identity of nucleotide against *B. mori* sequence is shown on the right-hand side. (C) Schematic diagram of mussagil analysis. Three yellow boxes indicate the identified region via promoter assay in Fig. 1. Red bars represent calculated sequences with high similarity among moth species (E1, E2, and E3). (D) Nucleotide fragments are identified in (C). Each fragment possesses a potential consensus sequence, Mef2 (YTAWWWWTAR), zen (CATTWAANTT), and CREB (TGACGTCA), respectively.

Figure 3. Promoter activity assays of deletion promotes. Quantitative fluorescent intensity using deletions spanning -200 to -100 (A) and -100 to -50 (B). Data of each promoter activity are collected from multiple animals ($n = 5$). Small letters indicate significant differences ($p < 0.05$, ANOVA followed by the Tukey HSD test) among groups. Each bar represents the mean $\pm$ standard

deviation (SD) of assays.

Figure 4. dbcAMP influences to the ChAT gene expression. (A) Comparison of the cholinergic gene locus promoter among insect species. Nucleotide sequences are collected through the Blast using the genome database of each insect species. Two conserved sequences, TGACGTCA and CCAAT are surrounded by boxes. The *ChAT* expression levels in (B) and (C) were stabilized by a housekeeping gene, *rp49*. Data of each condition are collected from multiple animals ($n = 3$). Each bar represents the mean $\pm$ standard deviation (SD) of assays. (B) qPCR analysis in *ex-vivo* cultured larval brain between dbcAMP (-) and (+). Asterisk (*) indicates significant differences ($p < 0.01$, ANOVA followed by the Tukey HSD test). (C) qPCR analysis in larval brain injected actCREB or repCREB virus vectors. Small letters indicate significant differences ($p < 0.05$, ANOVA followed by the Tukey HSD test). If a non-significant difference was found between two groups, over-lapping letters are used to indicate the statistical significance.

Figure 5. Gene expression patterns of candidate transcriptional factors (A) The schematic structure of the cholinergic gene promoter region. Yellow boxes indicate that three identified regions in Fig. 1. Candidate transcriptional factors predicted by STAMP are described. Detail data of E-values against each TF consensus sequence is listed in Supplementary Table 3. (B) Gene expression profile of each TF gene in different tissues; FB: Fat body, MG: Malpighian tube, SG: Silk gland, M: Midgut, E: Epidermal cells, Br: Brain. Primer information of target genes is listed in Supplementary Table 4.

Reference

- Alberini, C.M. (2009) Transcription factors in long-term memory and synaptic plasticity. *Physiological Reviews*, **89**, 121-145. <https://doi.org/10.1152/physrev.00017.2008>.
- Banzai, K., Adachi, T., Izumi, S. (2015) Comparative analyses of the cholinergic locus of ChAT and VAcHT and its expression in the silkworm *Bombyx mori*. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, **185**, 1-9. <https://doi.org/10.1016/j.cbpb.2015.03.001>.
- Cao, J., Ni, J., Ma, W., Shiu, V., Milla, L.A., Park, S., et al. (2014) Insight into insulin secretion from transcriptome and genetic analysis of insulin-producing cells of *Drosophila*. *Genetics*. **197**(1), 175-192. <https://doi.org/10.1534/genetics.113.160663>.
- Cardinaux, J.R., Notis, J.C., Zhang, Q., Vo, N., Craig, J.C., Fass, D.M. et al. (2000) Recruitment of CREB binding protein is sufficient for CREB-mediated gene activation. *Molecular and Cellular Biology*, **20**(5), 1546-1552. <https://doi.org/10.1128/mcb.20.5.1546-1552.2000>
- Chen, X., Chereddy, S.C.R.R., Gurusamy, D., Palli S.R. (2020) Identification and characterization of highly active promoters from the fall armyworm, *Spodoptera frugiperda*. *Insect Biochemistry and Molecular Biology*. **126**, 103455. <https://doi.org/10.1016/j.ibmb.2020.103455>
- Chomczynski, P, Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidium thiocyanate-phenol-chloroform extraction. *Analytical Biochemistry*, **162**, 156-159. <https://doi.org/10.1006/abio.1987.9999>.
- Crittenden, J.R., Skoulakis, E.M.C., Goldstein, E.S., Davis, R.L. (2018) *Drosophila mef2* is essential for normal mushroom body and wing development. *Biology Open*. **7**(9), bio035618. <https://doi.org/10.1242/bio.035618>.
- El Hassani, A.K., Dacher, M., Garry, V., Lambin, M., Gauthier, M., Armengaud, C. (2008) Effects of sublethal doses of acetamiprid and thiamethoxam on the behavior of the honeybee (*Apis mellifera*). *Archives of Environmental*

Contamination and Toxicology, **54**, 653-661. <https://doi.org/10.1007/s00244-007-9071-8>.

Gauglitz, S., Pfluger, H.J. (2001) Cholinergic transmission via central synapses in the locust nervous system. *Journal of Comparative Physiology A*, **187**, 825-836. <https://doi.org/10.1007/s00359-001-0253-y>.

Gonzalez, G.A, Montminy M.R. (1989) Cyclic AMP stimulates somatostatin gene transcription by phosphorylation of CREB at serine 133. *Cell*. **59**(4), 675-680. [https://doi.org/10.1016/0092-8674\(89\)90013-5](https://doi.org/10.1016/0092-8674(89)90013-5).

Gupta, A.P. (1987) *Arthropod Brain - Its Evolution, Development, Structure, and Functions*. Wiley-Interscience. New York

Heliconius Genome Consortium (2012) Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature*. **487**(7405), 94-98. <https://doi.org/10.1038/nature11041>.

Ichikawa, T., Ajiki, K., Matsuura, J., Misawa, H. (1997) Localization of two cholinergic markers, choline acetyltransferase and vesicular acetylcholine transporter in the central nervous system of the rat: in situ hybridization histochemistry and immunohistochemistry. *Journal of Chemical Neuroanatomy*, **13**(1), 23-39. [https://doi.org/10.1016/S0891-0618\(97\)00021-5](https://doi.org/10.1016/S0891-0618(97)00021-5)

Kitamoto, T., Ikeda, K., Salvaterra, P.M. (1992) Analysis of *cis*-regulatory elements in the 5' flanking region of the *Drosophila melanogaster* choline acetyltransferase gene. *The Journal of Neuroscience*, **12**(5), 1628-1639. <https://doi.org/10.1523/JNEUROSCI.12-05-01628.1992>.

Kitamoto, T., Salvaterra, P.M. (1993) Developmental regulatory elements in the 5' flanking DNA of the *Drosophila* choline acetyltransferase gene. *Roux's archives of developmental biology*, **202**(3), 159-169. <https://doi.org/10.1007/BF00365306>.

Kitamoto, T., Salvaterra P.M. (1995) A POU homeo domain protein related to dPOU-19/pdm-1 binds to the regulatory DNA necessary for vital expression of the *Drosophila* choline acetyltransferase gene. *The Journal of Neuroscience*, **15**(5),

3509-3518. <https://doi.org/10.1523/JNEUROSCI.15-05-03509.1995>

Kratsios, P., Stolfi, A., Levine, M., Hobert O. (2011) Coordinated regulation of cholinergic motor neuron traits through a conserved terminal selector gene. *Nature Neuroscience*, **15**(2), 205-214. <https://doi.org/10.1038/nn.2989>.

Krnjevic, K., Miledi, R. (1958) Acetylcholine in mammalian neuromuscular transmission. *Nature*. **182**(4638), 805-806. <https://doi.org/10.1038/182805b0>.

Lee, M.H., Salvaterra, P.M. (2002) Abnormal chemosensory jump 6 is a positive transcriptional regulator of the cholinergic gene locus in *Drosophila* olfactory neurons. *The Journal of Neuroscience*, **22**(13), 5291-5299. <https://doi.org/10.1523/JNEUROSCI.22-13-05291.2002>.

Li, Z., Tatsuke, T., Sakashita, K., Zhu, L., Xu, J., Mon, H., et al. (2012) Identification and characterization of Polycomb group genes in the silkworm, *Bombyx mori*. *Molecular Biology Reports*. **39**(5), 5575-5588. <https://doi.org/10.1007/s11033-011-1362-5>

Lu, L., Wang, H., Manopo, I., Yu, L., Kwang, J. (2005) Baculovirus-mediated promoter assay and transcriptional analysis of white spot syndrome virus *orf427* gene. *Virology Journal*, **2**, 71. <https://doi.org/10.1186/1743-422X-2-71>.

Mahony, S., Benos, P.V. (2007) STAMP: a web tool for exploring DNA-binding motif similarities. *Nucleic Acids Research*, **35**, W253-258. <https://doi.org/10.1093/nar/gkm272>.

Salvaterra, P.M., Kitamoto, T. (2001) *Drosophila* cholinergic neurons and processes visualized with Gal4/UAS-GFP. *Gene Expression Patterns*, **1**(1), 73-82. [https://doi.org/10.1016/s1567-133x\(01\)00011-4](https://doi.org/10.1016/s1567-133x(01)00011-4).

Shimojo, M., Wu, D., Hersh, L.B. (1998) The cholinergic gene locus is coordinately regulated by protein kinase A II in PC12 cells. *Journal of Neurochemistry*, **71**(3), 1118-1126. <https://doi.org/10.1046/j.1471-4159.1998.71031118.x>.

- Shiomi, K., Kajiura, Z., Nakagaki, M., Yamashita, O. (2003) Baculovirus-Mediated Efficient Gene Transfer into the Central Nervous System of the Silkworm, *Bombyx mori*. *Journal of Insect Biotechnology and Sericulture*, **72**, 149-155. <https://doi.org/10.11416/jibs.72.149>.
- Shirata, N. Ikeda, M., Kamiya, K. Kawamura, S., Kunimi, Y., Kobayashi, M. (1999) Replication of nucleopolyhedroviruses of *Autographa californica* (Lepidoptera: Noctuidae), *Bombyx mori* (Lepidoptera: Bombycidae), *Hyphantria cunea* (Lepidoptera: Arctiidae), and *Spodoptera exigua* (Lepidoptera: Noctuidae) in four lepidopteran cell lines. *Applied Entomology and Zoology*, **34**, 507-516. <https://doi.org/10.1303/aez.34.507>.
- Song., H., Ming, G., Fon, E., Bellocchio, E., Edwards, R.H., Poo, M. (1997) Expression of a putative vesicular acetylcholine transporter facilitates quantal transmitter packaging. *Neuron*, **18**(5), 815-826. [https://doi.org/10.1016/s0896-6273\(00\)80320-7](https://doi.org/10.1016/s0896-6273(00)80320-7).
- Suzuki, A., Fukushima, H., Mukawa, T., Toyoda, H., Wu, L.J., Zhao, M.G. et al. (2011) Upregulation of CREB-mediated transcription enhances both short- and long-term memory. *The Journal of Neuroscience*, **31**(24), 8786-802. <https://doi.org/10.1523/JNEUROSCI.3257-10.2011>.
- Vinson, C.R., Hai, T., Boyd, S.M. (1993) Dimerization specificity of the leucine zipper-containing bZIP motif on DNA binding: prediction and rational design. *Genes and Development*, **7**(6), 1047-1058. <https://doi.org/10.1101/gad.7.6.1047>.
- Yasuyama, K., Salvaterra, P.M. (1999) Localization of choline acetyltransferase-expressing neurons in *Drosophila* nervous system. *Microscopy Research and Technique*, **45**(2), 65-79. [https://doi.org/10.1002/\(SICI\)1097-0029\(19990415\)45:2<65::AID-JEMT2>3.0.CO;2-0](https://doi.org/10.1002/(SICI)1097-0029(19990415)45:2<65::AID-JEMT2>3.0.CO;2-0).
- Yin, J.C., Wallach, J.S., Vecchio, M.D., Wilder, E.L., Zhou, H., Quinn, W.G. et al. (1994) Induction of a dominant negative CREB transgene specifically blocks long-term memory in *Drosophila*. *Cell*. **79**(1), 49-58. [https://doi.org/10.1016/0092-8674\(94\)90399-9](https://doi.org/10.1016/0092-8674(94)90399-9).

634 Zhan, S., Merlin, C., Boore, J.L., Reppert S.M. (2011) The monarch butterfly
635 genome yields insights into long-distance migration. *Cell*. **147**(5), 1171-1185.
636 <https://doi.org/10.1016/j.cell.2011.09.052>.
637
638

Fig. 1. Promoter activity assays of sequential truncation promoters

K

L

Fig. 2. The cholinergic gene loci of three moth species

B The DNA sequences of putative first exons among moth species

Bm: AGTCCGGGCAAGGTGTTCCCAAGTGTGCCAAGCTCGCCCCGCCACAAGAACGAG (Identity vs B. mori)
Dp: AGTCTGGGCTCG-CGGGACCAGGTGTGCCAAGCTCGCCCCGATACTACA-CAAG (73.1%, 38/52)
Hm: AGTCTGGGCCCTCGGTGCTCAGGTGTGCCGAGCTCGCGATCACTTTACA CAAG (62.3%, 33/53)
.... :..... :.....

D

E1

B:-166, TCTATTTTTAGTT, -153
 D:-166, TCTATATTTAGCG, -153
 H:-162, CCTATTTTAGCG, -149
 Mef2: YTAWWWTAR

E2

B:-132, CCTACATGAAAGTTTCAGG, -114
 D:-131, CCTACATGAAAATTTCAGG, -113
 H:-129, CCTACGTGAAA--TTCAGG, -113
 Zen: CATTWAANTT

E3

B:-92, CGGTGACGTCAGG, -79
 D:-91, CGGTGACGTCACC, -78
 H:-91, GAATGACGTCACA, -78
 CREB: TGACGTCA

W:A or T; R:A or G; Y:T or C; N:A,T,G or C

Fig. 3. Promoter activity assays of deletion promoters

Fig. 4. dbcAMP influences to the ChAT gene expression

A Comparison of the cholinergic gene locus promoter among species

B. mori	-100:	tgggtcgtcgg	TGACGTCA	gggcaagccggggc	-----	CCAAT	cacggggcttcg	:-50
D. plexippus	-99:	cgaggctccgg	TGACGTCA	ccaccggctgggct	-----	CCAAT	gagggccggagg	
H. melpomene	-99:	tccactccgaa	TGACGTCA	cactccactgtcag	-----	CCAAT	ggtggagggagg	
A. aegyptis	-217:	tgccgatactg	TGACGTCA	cgatggcaagga	-----	CCAAT	caggctgccatgga	
A. albimanus	-352:	cgaatggtctg	TGACGTCA	ctcgaggtacgggtcaaaagca	-----	CCAAT	cagcg	
A. arabiensis	-431:	acgcacggctg	TGACGTCA	ctcgagct--ggccgga	----	CCAAT	caacgctaacc	
T. casraneum	-183:	gcgtcgagatg	TGACGTCT	ggcgctcgaaa	-----	CCAAT	cgtgtggcggcagtaa	
Consensus			TGACGTCA			CCAAT		

cAMP Response Element (CRE) CCAAT box

B

C

Fig. 5. Gene expression patterns of candidate transcriptional factors

A Screening candidate TFs from database

B Tissue-specific expression analysis (RT-PCR)

Table 1. Oligonucleotide sequences of primer used for this study.

Primer name	Sequence (5'-3')	Applications
EGFP-F	AACCATGGTGAGCAAGGGCGAG	EGFP ORF
EGFP-R	AATCTAGAATTACTTGTACAGCTCGTCCATG	EGFP ORF
ChAT-2000-F	GGGAATTCGGTGCGATGGTAAAAACGAG	Truncation promoter
ChAT-1000-F	GGGAATTCAGTTACAACGGCTGCCCAAC	Truncation promoter
ChAT-761-F	GGGAATTCCTAGCTGTCCAATATACATAG	Truncation promoter
ChAT-500-F	GGGAATTCCTTATTTGTAATATGGTAGC	Truncation promoter
ChAT-300-F	GGGAATTCACAATGTCTGGCTCGTCAAAACG	Truncation promoter
ChAT-200-F	GGGAATTCATCGACCAATACGGATCCCAAATG	Truncation promoter
ChAT-100-F	GGGAATTCCTGGGTCGTGGTGACGTCAG	Truncation promoter
ChAT-50-F	GGGAATTCGTCCGCGCATGTCAATACAG	Truncation promoter
ChAT+30-R	GGCCATGGTGGCACACTTGGGAACACCTTG	Truncation promoter
Del200-150-invF	ATTTGTGTACGTGTGTGCAC	Deletion mutant
Del200-150-invR	GACCTATCTCGTCTCCTGCC	Deletion mutant
Del200-100-invF	TGGGTCGTCCGTGACGTCAG	Deletion mutant
Del200-100-invR	GACCTATCTCGTCTCCTGCC	Deletion mutant
Del100-50-invF	GTCCGCGCATGTCAATACAG	Deletion mutant
Del100-50-invR	CCAACCACGGCGCCCTGAAAC	Deletion mutant
Del100-90-F	GTGACGTCAGGGCAAGC	Deletion mutant
Del100-90-R	CCAACCACGGCGCCCTGAAAC	Deletion mutant
Del90-80-F	GGCAAGCCGGGGCCCAATCA	Deletion mutant
Del90-80-R	CGACGACCCACCAACCACG	Deletion mutant
Del80-70-F	GGCCCAATCACGGGGCTTCG	Deletion mutant
Del80-70-R	CTGACGTCACCGACGACCCAC	Deletion mutant
Del70-60-F	CGGGGCTTCGGTCCGCGC	Deletion mutant
Del70-60-R	CCGGCTTGCCCTGACGTCAC	Deletion mutant
Del60-50-F	GTCCGCGCATGTCAATACAG	Deletion mutant
Del60-50-R	TGATTGGGCCCGGCTTG	Deletion mutant
CREB-F	GGGACCATGGAAGGGATGGTTGAAGA	CREB overexpression
CREB-R	GGTCTAGACTCATTCCGTTTTCTGTTGGC	CREB overexpression
CREB(act)-F	TTCGATGTCAGTCAACATTTCTCTATATTTACG	CREB overexpression
CREB(act)-R	GACATGCTGAAAATTTTAAATGACCTTGGAG	CREB overexpression
CREB(rep)-F	GGGACCATGGAAGGGATGGTTGAAGA	CREB overexpression
CREB(rep)-R	AATCTAGATCGGTACATAGAATTGTCCATC	CREB overexpression
Actin-Promoter-F	GGGGTACCGCGTTACCATATATGGTGAC	CREB overexpression
Actin-Promoter-R	GGAAGCTTATCTTGAATTAGTCTGCAAG	CREB overexpression
ChAT-Promoter-F	GGGAATTCGGTGCGATGGTAAAAACGAG	CREB overexpression
ChAT-Promoter-R	GGCCATGGTGGCACACTTGGGAACACCTTG	CREB overexpression
ChAT-F	GGAATCACTGCTGGATAGGCA	gene expression analysis, qPCR
ChAT-R	CAACCTCCTCCATCTTGACG	gene expression analysis, qPCR
rp49-F	GGCGTAAACCTAGAGGTATTGAC	gene expression analysis, qPCR
rp49-R	CATTGTGAAGTACGACCTTACGG	gene expression analysis, qPCR

Table 2. Genomic DNA databases of insect species

Species	Database name	URL
<i>Danaus plexippus</i>	Monarchbase	http://monarchbase.umassmed.edu/
<i>Heliconius melpomene</i>	Butterfly genome database	http://www.butterflygenome.org/
<i>Aedes aegypti</i>	VectorBase	https://vectorbase.org/vectorbase/app
<i>Anopheles albimanus</i>	VectorBase	https://vectorbase.org/vectorbase/app
<i>Anopheles arabiensis</i>	VectorBase	https://vectorbase.org/vectorbase/app
<i>Tribolium castaneum</i>	BeetleBase	http://www.bioinformatics.ksu.edu/BeetleBase/

dataset	region	gene	E-value	Consensus	sequence on cholinergic gene locus promoter	Orientation
flyreg (Bergman/Pollard)	761-500	til	6.32E-07	TRACTTT	TGACTTT	(+)
		Dfd	6.86E-07	TTAATTRY	TTAATTGT	(-)
		toy	1.08E-06	NYSNRTGAGNGG	GTGGGTGAGGGG	(+)
		p120	8.74E-08	AACATTCAATAG	AACAATTAATAG	(+)
		hb	2.68E-06	NYAAAAAA	TTAAAAAA	(+)
		exd	8.70E-06	NARNWYNAAA	AAAATCAAAA	(-)
		ara	2.57E-05	AATYAA	AATCAA	(-)
		Abd-B	2.15E-06	TTTTATGR	TTTTATGT	(+)
		hb	2.68E-06	NYAAAAAA	ATAAAAAA	(-)
		Dfd	9.07E-06	TTAATTRY	TTAATGAT	(-)
		ara	2.57E-05	AATYAA	AATTAA	(-)
		HLHm5	5.74E-05	TGNCNCGTGC	TCTCGCGTGC	(+)
		200-100				
		br-Z1	3.43E-08	AAWTANAA	AAATAGAA	(-)
Fly (curated by Bergman)	761-500	Aef1	8.90E-07	TTGTTG	TTGTTG	(-)
		gcm	3.04E-05	ATGCGGGT	ATGTGGGT	(+)
		til	4.71E-05	TTAATTT	TTAATTG	(-)
		tin	7.39E-07	TYAAGTG	TTAAGTG	(-)
		br-Z1	2.02E-06	TAAWWRACAARW	TAAAAATCAAAA	(-)
		pan	3.29E-05	AAGATCAAAGG	AAAATCAAAAT	(-)
		hb	4.61E-05	TTTTTTATKS	TTTTTTAAGT	(-)
		hb	6.78E-09	TTTTTTATKS	TTTTTTATGT	(+)
		Aef1	8.90E-07	TTGTTG	TTGTTG	(+)
		cad	8.90E-07	TTTATG	TTTATG	(+)
		croc	3.41E-06	TTTTAT	TTTTAT	(+)
		300-200				
		pho	8.90E-07	AATGGC	AATGGC	(+)
		til	4.71E-05	TTAATTT	TTAATTC	(+)
		eve	7.48E-05	AKTTAAWTGA	ATTTAATTCC	(+)
		zen	7.48E-05	AKTTAAWTGA	ATTTAATTCC	(+)
		200-100				
		br-Z2	7.48E-05	AAATAGWWAA	AAATAGAACC	(-)

#	Gene	#GeneBank		Sequence (5'-3')	Tm (°C)	size (bp)	annealing on PCR	number of cylce
1	Abd-B	NM_001173331.1	F	TCTACGAAATTACGCACCAG	53.64	260	55	35
			R	GATTCAGATTTCTCGCGAGC	56.15			
2	Aef1	XM_004922137.3	F	CATTAGAATACATACTGCGG	50.1	455	50	35
			R	AACATAAGGTTTCTCTGCAG	51.15			
3	ara	XM_004929763.2	F	AGATGACGCTGACACAAGTC	54.91	442	55	35
			R	AAGTAGGGCAGTTCGTCAGA	55.45			
4	br-Z1	NM_001111334.1	F	GCGAGCGTTTTTGATCGACT	56.84	586	55	35
			R	GTGTACGCCACGTCTTGAAG	55.86			
5	br-Z2	NM_001111333.1	F	CACGATTAAATCCAGTGTGTAGTAC	54.5	371	55	35
			R	ACGATCACTGGATGCTTGCA	57.04			
6	cad	NM_001044049.1	F	GGTCGCCATACGAGTGGATG	57.07	204	55	35
			R	TACCTGTGTTCCGAGAGAC	55.39			
7	croc	NM_001252674.1	F	GGACATTAGACCCAGATTCTG	52.48	406	55	35
			R	GGTGCTGTATACCATCCCCGA	54.74			
8	Dfd	D83534.1	F	AATATGCCGCCGGTGCAGAT	57.38	440	55	35
			R	GTGCGCTGACGTTTAGGCT	58.12			
9	eve	NM_001043862.1	F	GCTGGAGAAGGAGTTCATGA	54.37	509	55	40
			R	CATAAAACATGGCCAGCACATC	55.41			
10	exd	AB841084.1	F	CAGCTGATGAGGCTTGACAA	56.3	516	55	35
			R	TACGGTAATGCCGCACTTGC	57.88			
11	gcm	XM_004929318.3	F	GTCACGTGCGAAGAGTCTAC	55.58	579	55	35
			R	CGCATGTGCATGTTAGTTGA	54.49			
12	hb	XM_004922967.2	F	CCACCGACAGTCAAACTGA	54.12	528	55	35
			R	GTGGAAGTCGCTCAAGAC	54.95			
13	HLHm5	NM_001170995.1	F	AACTCAAGGAGCTGATGGTC	54.15	367	55	35
			R	CTCACGTAACCAGACGATGA	54.49			
14	p120	XM_012690954.2	F	ATCTGCAACAGAGGTACGAC	54.24	388	55	35
			R	TCCATATAGGTCAGGTGCTG	55.08			
15	pan	XM_012690912.2	F	AGATTCTAAACTAGCGGCTG	54.44	461	55	35
			R	TCTCCTTCATGTAGAGCATG	52.95			
16	pho	NM_001201433.1	F	ATAGCATGTCCACACAAAGG	49.79	339	50	35
			R	GCGTTAGTATGTGCGACTTC	52.2			
17	tin	XM_004933067.2	F	ATGAATACGACACGGGTACG	54.12	341	55	35
			R	TTGGGTAGGCGACAGGTTCA	56.44			
18	tll	XM_004932789.3	F	ACACCGAAATCAGTGTGCGAG	54.88	344	55	35
			R	TGGAGATAACAAGGCAGCTG	55.6			
19	toy	NM_001202531.1	F	CAAAAGATCGTGGAAGTAGC	52.45	261	55	35
			R	ATTGCACACGTTTTTCGCTGA	56.15			
20	zen	XM_021350827.1	F	ACGACATTCTTGGATTGCATC	53.91	496	55	35
			R	TCGTTCTCTAGTTCGACTAG	53.07			