

Research Article

Gene Expression Divergence and Evolutionary Analysis of the Drosomycin Gene Family in *Drosophila melanogaster*

Xiao-Juan Deng,^{1,2} Wan-Ying Yang,² Ya-Dong Huang,³ Yang Cao,² Shuo-Yang Wen,⁴ Qing-You Xia,⁵ and Peilin Xu¹

¹The Key Laboratory of Gene Engineering of Ministry of Education, Sun Yat-sen University, Guangzhou 510275, China

²Department of Sericulture Science, College of Animal Science, South China Agricultural University, Guangzhou 510642, China

³Biopharmaceutical Research and Development Center, Jinan University, Guangzhou 510632, China

⁴Department of Entomology, College of Natural Resource and Environment, South China Agricultural University, Guangzhou 510642, China

⁵Key Sericultural Laboratory of Agriculture Ministry, College of Sericulture and Biotechnology, Southwest University, Chongqing 400716, China

Correspondence should be addressed to Shuo-Yang Wen, shywen@scau.edu.cn

Received 17 March 2009; Revised 6 June 2009; Accepted 7 August 2009

Recommended by Mouldy Sioud

Drosomycin (*Drs*) encoding an inducible 44-residue antifungal peptide is clustered with six additional genes, *Dro1*, *Dro2*, *Dro3*, *Dro4*, *Dro5*, and *Dro6*, forming a multigene family on the 3L chromosome arm in *Drosophila melanogaster*. To get further insight into the regulation of each member of the drosomycin gene family, here we investigated gene expression patterns of this family by either microbe-free injury or microbial challenges using real time RT-PCR. The results indicated that among the seven *drosomycin* genes, *Drs*, *Dro2*, *Dro3*, *Dro4*, and *Dro5* showed constitutive expressions. Three out of five, *Dro2*, *Dro3*, and *Dro5*, were able to be upregulated by simple injury. Interestingly, *Drs* is an only gene strongly upregulated when *Drosophila* was infected with microbes. In contrast to these five genes, *Dro1* and *Dro6* were not transcribed at all in either noninfected or infected flies. Furthermore, by 5' rapid amplification of cDNA ends, two transcription start sites were identified in *Drs* and *Dro2*, and one in *Dro3*, *Dro4*, and *Dro5*. In addition, NF- κ B binding sites were found in promoter regions of *Drs*, *Dro2*, *Dro3*, and *Dro5*, indicating the importance of NF- κ B binding sites for the inducibility of *drosomycin* genes. Based on the analyses of flanking sequences of each gene in *D. melanogaster* and phylogenetic relationship of *drosomycins* in *D. melanogaster* species-group, we concluded that gene duplications were involved in the formation of the drosomycin gene family. The possible evolutionary fates of *drosomycin* genes were discussed according to the combining analysis of gene expression pattern, gene structure, and functional divergence of these genes.

Copyright © 2009 Xiao-Juan Deng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. Introduction

Antimicrobial peptides are the major humoral immune factors in the insect innate immune system. *Drosophila melanogaster* has emerged as a powerful model system for the study of innate immunity, especially for the antimicrobial peptides studies, because of its flexible genetics [1]. To date, seven classes of the inducible antimicrobial peptides (or peptides families), drosomycins, metchnikowin, defensin, attacins, cecropins, drosocin, and dipterocins have been characterized in the *D. melanogaster*, each with a specific spectrum of activity [2]. Among them, *defensin* [3], *drosocin*

[4, 5], and *metchnikowin* [6] have only one single copy in the genome of *D. melanogaster*. Others are encoded by multigene families.

Drosomycin (*Drs*) is an inducible insect antifungal peptide from *D. melanogaster*. It exhibits potent activity against filamentous fungi but no obvious activity to bacteria and yeast [7]. The mature peptide of *Drs* is consisting of 44 amino acid residues with 8 cysteine residues to form 4 intramolecular disulfide bridges [8]. It involves an α -helix and a twisted three-stranded β -sheet, that is, Cysteine-stabilized $\alpha\beta$ motif [9]. In the genome of *D. melanogaster*, there are six additional mRNAs with high sequence homology to *Drs* gene, *Dro1*,

Dro2, *Dro3*, *Dro4*, *Dro5*, and *Dro6* (resp., equivalent to *Drs-IC*, *Drs-ID*, *Drs-IE*, *Drs-IF*, *Drs-IG*, and *Drs-II*). These seven genes are clustered into a drosomycin multigene family on 3L chromosome arm [10, 11]. Similar to *Drs*, the predicted amino acid sequences of *Dro1*, *Dro2*, *Dro3*, *Dro4*, *Dro5*, and *Dro6* also contain eight cysteines and four conservative residues (Ser⁴, Gly⁹, Glu²⁶, and Gly³¹), which have been reported to be involved in protein structure stabilization or in the protein folding pathway [9]. To experimentally determine the antifungal function of the six isoforms, Yang et al. [11] cloned all the seven genes of the drosomycin gene family into pET 3c and expressed in *Escherichia coli*. All purified expression products showed different antifungal activity against tested fungal strains with the exception of *Dro6* which showed no activity against all the tested fungal strains [11]. These experiments provided the evidence of the significant functional divergence in the drosomycin family.

Genes coding for antibacterial and antifungal peptides were differentially induced after injection of various classes of microorganisms [12]. In the drosomycin family, although the *Drs* was constitutively expressed in the larvae and adult of *D. melanogaster*, it was strongly upregulated by infection of fungi and Gram-positive bacteria and weakly upregulated by Gram-negative bacteria [7, 12–14]. Previous results from the genomewide analysis of the *Drosophila* immune response genes indicated that *Dro5* (*Drosomycin B*), in addition to *Drs*, was also significantly upregulated by the immunization of fungi or mixed bacteria [15]. Recent studies have demonstrated that *Dro2* and *Drs* were expressed in larvae, pupae, and adult, *Dro3*, *Dro4*, and *Dro5* expressed in larvae and adult, whereas the transcripts of *Dro1* and *Dro6* were not detected in all the stages of insect development [16]. However, the expression and regulation of each member of drosomycin family in response to microbial infection and injury stimulation has not yet been fully elucidated. In the present study, we investigated the expression patterns of seven members of the drosomycin family by simple injury and the various microbial challenges in the adults of *D. melanogaster*. In addition, we identified the transcription start sites and predicted the cis-regulatory elements at 5'-flanking region of individual gene in the drosomycin family. Moreover, the evolutionary history of this gene family was predicted. These studies will shed light on the gene expression divergence and the functional evolution of antimicrobial peptides multigene family.

2. Materials and Methods

2.1. *Drosophila* Strain. Wild-type strain *D. melanogaster* Oregon-R was obtained from *Drosophila* genetic resource center, Kyoto Institute of Technology, Japan. Flies were maintained on the cornmeal-malt medium at 25°C under continuous light.

2.2. Collection of Bacteria or Fungal Spores. Bacteria strains, including Gram-negative *Escherichia coli* K₁₂D₃₁ and *Pseudomonas aeruginosa* and Gram-positive *Staphylococcus aureus* and *Bacillus subtilis*, were grown in Luria broth (LB)

medium at 37°C to an OD₆₀₀ of approximate 0.6. The cells were harvested by centrifugation at 5000 g for 5 minutes and resuspended in sterile Ringer's solution at volume of 1/10.

Fungi strains including *Fusarium oxysporum*, *Fusarium culmorum* (W. G. Smith), *Beauveria bassiana* (vuill), *Botrytis cinerea*, *Penicillium digitatum*, and *Aspergillus niger* were grown on potato dextrose agar (PDA) medium. Spores from agar plates were resuspended in sterile Ringer's solution. After filtering through 8 layers of sterile cheese-cloth, the number of spores was counted under a light microscopy, and the final concentration of the spore solution was adjusted into 10⁹ cells per milliliter.

2.3. Immunization of Flies. Two-day-old adult flies were anesthetized with ethyl ether, then individually pricked into the thorax with a thin needle which was previously dipped into either bacterial culture (OD₆₀₀ ≈ 6) or fungal spore suspension (10⁹ cells/mL). The immunized flies were kept on the cornmeal-malt medium at 28°C. The survival flies at the different times of infection were immediately frozen in liquid Nitrogen and stored at –70°C until RNA extraction.

2.4. RNA Preparation, Regular RT-PCR, and Quantitative Real Time RT-PCR. The immunized flies were ground to a fine powder under liquid nitrogen. Total RNA was extracted with the Trizol reagents (Invitrogen, Carlsbad, USA) according to the manufacturer's instructions, and contaminating genomic DNA was removed by incubation with DNase I (Takara, Dalian, China). RNA (1 µg) was reverse-transcribed using the ReverTraAce reverse transcription kit (Toyobo, Tokyo, Japan), and the resulting first strand cDNA was used as a template for RT-PCR with primers specific to *drosomycin* genes (Table 1). PCR conditions were 94°C, 1 minute; 94°C, 30 seconds; 55°C, 30 seconds and 72°C, 30 seconds for 35 cycles; and 72°C for 7 minutes. RT-PCR fragments were cloned to pTA2 vector with a TA cloning kit (Toyobo, Tokyo, Japan) for sequencing to check the specificity of the primer pairs.

To quantify the expression levels of *drosomycin* genes, fluorescence real-time PCR was performed using an SYBR Green methodology (Applied Biosystems) under the following conditions: 25 µL reaction mixture contained 12.5 µL SYBR Green Real-time PCR Master Mix, 0.5 µmol/L of forward primer, 0.5 µmol/L of reverse primer, and 0.5 µL of cDNA (corresponding to 25 ng of total RNA) template. The real time RT-PCR was run with the 96-well plate on ABI Prism 7300 Real Time PCR system (Applied Biosystems, Foster City, USA) with the program of 95°C for 1 minute, 40 cycles of 95°C for 15 seconds, 55°C for 15 seconds, and 72°C for 45 seconds. The ribosome protein *rp49* gene was used as an internal control. All samples were analyzed in triplicate and normalized against *rp49* gene.

2.5. Rapid Amplification of 5' cDNA Ends (5' RACE). Total RNA was isolated from adult *Drosophila* infected with *S. aureus*. The transcription start site was determined using the 5' RACE kit (Invitrogen, Carlsbad, USA) according to the manufacturer's instructions. Briefly, first strand cDNA

TABLE 1: Primers for RT-PCR and real-time PCR.

Genes	Primer sequence	Product size (bp)
<i>Drs</i>	forward: 5' CCCTCTTCGCTGCCTGA 3'	137
	reverse: 5' GCGTCCCTCCTCCTTGC 3'	
<i>Dro1</i>	forward: 5' TGTCGCTGTCTTGATG 3'	135
	reverse: 5' TTCGCCCTTCCCTCT 3'	
<i>Dro2</i>	forward: 5' TCAAATTCCTTTTCGTCTT 3'	130
	reverse: 5' CGTCGGCACATCTCGT 3'	
<i>Dro3</i>	forward: 5' GCACACTGTTTTGGCAGC 3'	80
	reverse: 5' GCGGCACTTTTCTCC 3'	
<i>Dro4</i>	forward: 5' ATGGCTCAAATTAAGGATT3'	148
	reverse: 5' AGAGGCGACGGCACT 3'	
<i>Dro5</i>	forward: 5' ACCTCTTCCTGGCTGT 3'	87
	reverse: 5' AGGGTCCTCCGTATCT 3'	
<i>Dro6</i>	forward: 5' TGTTACCTTCCTCGCTCTG 3'	156
	reverse: 5' CACTCACTCGTCTCGTCCC 3'	
<i>rp49</i>	forward: 5' CGTTTACTGCGGCGAGAT 3'	102
	reverse: 5' CCGTTGGGGTTGGTGAG 3'	

TABLE 2: Primers for 5' RACE reactions. The GSP1 sequence was corresponded to positions 190~208 in the ORFs of *Dro1* and *Dro5*, 193~211 in *Dro2* and *Drs*, 196~214 in *Dro3* and *Dro4*, and 199~208 in *Dro6*.

Primers	5' to 3' sequence
GSP1	AGCATCCTTCGCACCAGCA
<i>Dro2</i> -GSP2	AAATACGTCGGCACATCTCG
<i>Dro2</i> -GSP3	CGGCATCAGCCATATTGGCGG
<i>Dro3</i> -GSP2	GACGGCGCACTTTTCTCC
<i>Dro3</i> -GSP3	CAATCACGTGCCAAAACAGT
<i>Dro4</i> -GSP2	CCTCCCTGCAGAGGCGACG
<i>Dro4</i> -GSP3	CCAGATGGGCAATCCACGGC
<i>Dro5</i> -GSP2	CGCAGGGTCCTCCGTATCT
<i>Drs</i> -GSP2	CAGCATCAGGACAGCGAAGA

was generated from total RNAs with GSP1, a common primer annealed to all the 7 *drosomycin* mRNAs. After reverse transcription, the cDNAs were purified using columns supplied in 5' RACE kit. The resulting purified cDNAs were then oligo-dC-tailed at its 3' end by TdT (Terminal deoxynucleotidyl transferase). PCR products were amplified from dC-tailed cDNAs using an abridged anchor primer (AAP) and a nested gene-specific primer GSP2 (Table 2). To obtain specific RACE products of *Dro2*, *Dro3*, and *Dro4*, diluted primary PCR products were reamplified by using an abridged universal amplification primer (AUAP) and a nested gene-specific primer GSP3 (Table 2). The PCR products were purified by E. N. Z. A. Cycle-pure kit (Omega, USA) and subsequently cloned into a pTA2 vector (Toyobo, Tokyo, Japan). About 20 to 30 positive clones were selected for sequencing using T7 or M13R primers.

2.6. Prediction of Promoters and Regulatory Elements. The 5' flanking sequences of the 7 *drosomycin* genes were loaded

to predict the core promoter sequences by NNPP version 2.2 [17] (http://www.fruitfly.org/seq_tools/promoter.html). Binding sites of regulatory factors were analyzed by pDRAW32 software (<http://www.acaclone.com/>) within the 1 kb 5'-flanking region of *drosomycin* genes. The putative regulatory elements involved in the immune response were NF- κ B/Rel sites with consensus sequence of GGGRAYYYYY in *Drosophila* [18], GATA sites with WGATAR [19], IL6-RE (interleukin-6 response element) sites with TKNNGNAAK [20], and ICRE (interferon consensus element) sites with GGAAANN [21].

2.7. Evolutionary Analysis of the *Drosomycin* Gene Family. The 5' and 3'-flanking sequences of *drosomycin* genes in *D. melanogaster* were obtained from the Flybase (<http://www.flybase.org>). A Blast search of the open reading frames (ORFs) corresponding to *drosomycins* in *D. melanogaster* was applied to identify the *drosomycin* genes in 11 *Drosophila* species other than *D. melanogaster* at Flybase and UCSC genome websites (<http://genome.ucsc.edu/cgi-bin/hgBlat>).

Phylogenetic analysis of ORFs sequences of *drosomycins* in *Drosophila* species was performed using the Neighbor-Joining method with 1000 bootstrap replicates, as implemented by the Mega 4 programme [22]. Gaps were pairwise deleted and the Maximum Composite Likelihood model was applied to estimate the branch length.

To predict the gene duplication events of *drosomycin* gene family, repetitive sequences of the flanking regions were identified by RepeatMasker (<http://www.repeatmasker.org/>). Transposable elements were recognized by BLAST in *D. melanogaster* transposable element database at Flybase.

3. Results

3.1. Transcriptional Profile of the *Drosomycin* Gene Family. To detect the primary transcriptional profiles of 7 members of the *drosomycin* gene family, we used the regular RT-PCR with specific primers (Table 1) to identify the transcriptional products from adult flies after challenging to either bacteria or fungi as indicated in Section 2. The RT-PCR bands were amplified with primers specific to *Dro2*, *Dro3*, *Dro4*, *Dro5*, and *Drs* when templates of both the immunized and the native control flies were used. In contrast, no visible RT-PCR bands were observed with primers specific to *Dro1* and *Dro6* for all the tested templates (data not shown). These results suggested that the expressions of *Dro2*, *Dro3*, *Dro4*, *Dro5*, and *Drs* were constitutive, whereas *Dro1* and *Dro6* were not transcribed at all even in the adult flies that challenged to microbial infections. We cloned and sequenced the RT-PCR fragments including PCR products of *Dro1* and *Dro6* using the genomic DNA as template, confirming that all primers were gene-specific.

We next examined the time-course expression of *Dro2*, *Dro3*, *Dro4*, *Dro5*, and *Drs* in adult flies infected by fungi (*F. oxysporum* or *B. bassiana*) or bacteria (*E. coli* K12D31 or *S. aureus*) for 3 to 48 hours. The native flies and microbial-free pricked flies were used as controls. The survival rate of microbial-free pricked flies and infected flies with various

microorganisms was shown in Supplementary Figure 1. The results of the quantitative real time RT-PCR revealed that (i) in native flies, the transcription activity of *Dro2* was the weakest among five genes, only 1/1000 of control *rp49*. *Dro3* and *Dro5* expressions were 2 ~ 3-folds higher than that of *Dro2*. The expression of *Dro4* showed even higher, about 1/4 of that of *rp49*. The transcription activity of *Drs* was the highest, reaching an mRNA levels similar to *rp49* (Figure 1(a)), equivalent to about 80% of the total transcription of all the *drosomycin* genes. (ii) In the simple injured flies, with the microbe-free pricking treatment, the transcription activity of *Dro2*, *Dro3*, and *Dro5*, but not *Dro4* and *Drs*, was significantly upregulated (Figure 1(a)). (iii) In the microorganism-pricked flies, the transcription activity of *Drs* was significantly upregulated under stimulation by Gram-positive bacteria and fungi (Figures 1(c)–1(e)) but only increased about 1.5-folds after infection of Gram-negative bacteria (Figure 1(b)), compared to the control flies, suggesting that Gram-negative bacterium was a weak inducer. Taking all these results together, we concluded that of the 7 members, only *Drs* was greatly triggered by microbial infection. In other words, the microorganisms did not induce the expression of *Dro2*, *Dro3*, and *Dro5* above the level of simple injury and did not induce *Dro4* above the constitutional expression level.

3.2. Analysis of the Promoter Regions of *Drosomycin* Genes. Firstly, we analyzed the promoter sequences of the *drosomycin* genes for the prediction of core promoter sequence by NNPP software. The results showed that almost all of the *drosomycin* genes contained one or two typical core promoter sequences within 150 bp upstream of the ORFs with the exception of *Dro1*. The absence of a core promoter sequence in the *Dro1* may explain why this gene has no transcription activity.

Secondly, we determined the position of transcription start sites of *drosomycin* genes by 5' RACE. We initially synthesized the cDNA by primer GSP1 (Table 2) and then amplified these specific 5' RACE products with the primers GSP2 and AAP in PCR reaction, respectively. A single 5' RACE band about 200 bp for *Dro5* and 120 bp for *Drs* was obtained by one turn of PCR reaction (Figures 2 (d), 2 (e)). In the second round of PCR reaction, a single specific 5' RACE band for *Dro2*, *Dro3*, and *Dro4* was obtained with the primers GSP3 and AUAP, respectively (Figures 2 (a)–2 (c)). The 5' RACE products were subsequently purified and cloned into a T-vector for sequencing. DNA sequence analysis revealed that *Dro3*, *Dro4*, or *Dro5* contain one transcription start site, which located at ⁻⁵⁵A (1 bp downstream of the predicted TSS by NNPP), ⁻³⁶A (same to the predicted site), and ⁻⁵⁴A (1bp upstream of the predicted TSS) of the genome of *D. melanogaster*, respectively. Two transcription start sites were identified from the 5' RACE sequences of *Drs* and *Dro2*. The transcription of *Dro2* initiated at ⁻⁵⁵A or ⁻⁵⁸A, and *Drs* initiated at ⁻⁵⁹A or ⁻⁶¹A (Figure 3).

Thirdly, we scanned the potential binding sites for the regulatory factors, such as NF- κ B/Rel, GATA, IL-6RE, and ICRE, by pDRAW32 software, as these DNA motifs were known to be important for regulation of antimicrobial

gene expression. As indicated in Figure 4, GATA sites were found in the promoter regions of all the *drosomycin* genes. In addition, IL6-RE sites were identified in most of the *drosomycin* genes with the exception of *Dro3*. The putative ICRE sites were also present in most of the *drosomycin* genes except *Dro1* and *Dro4*. However, several ICRE sites would be identified at the 5'-flanking region of *Dro1* and *Dro4* when sequence GAAANN [23] was used instead of GGAAANN [21] for the prediction. It was noteworthy to point out that there was one NF- κ B site in the promoter region of *Dro2* and *Dro3*. *Dro5* contained two overlapping sequences of NF- κ B sites and *Drs* contained three NF- κ B sites. Interestingly, *Dro2*, *Dro3*, *Dro5*, and *Drs* contained at least one NF- κ B site near the transcription start site(s). However, no NF- κ B site was found in *Dro4* and *Dro1*. A putative NF- κ B site was identified at around 1 kb upstream from the translation start site of *Dro6*, much farther than the general distance of functional NF- κ B sites as previously reported [24]. The presence of the potential cis-regulatory modules in the *drosomycin* genes suggested that they might be involved in the regulation of the differential expression pattern of *drosomycin* genes. NF- κ B site, in particular, was important for the inducibility of *drosomycin* genes.

3.3. The Evolutionary History of *Drosomycin* Gene Family. We searched the *drosomycin*s homologs from the 12 sequenced *Drosophila* genomes by reciprocal Blast with seven members of *drosomycin*s in *D. melanogaster*. From the Blast results, we found that homologs were present only in the *Drosophila melanogaster* species-group of *Sophophora* subgenus: 4 homologs in *D. ananassae* of *ananassae* species-subgroup, 6 homologs in the *D. sechellia* and *D. yakuba*, and 7 homologs in *D. simulans* and *D. erecta* of *melanogaster* species-subgroup. In *melanogaster* species-subgroup, *Dro3* was lost in *D. yakuba* and *Dro4* was lost in *D. sechellia*. To predict the possibility of gene duplication involving in the formation of the *drosomycin* family, we investigated the TEs and repetitive sequences in the flanking regions of each *drosomycin* gene in *D. melanogaster*. As shown in Figure 5, there are TEs in flanking regions of each gene, except the 387 bp region between *Dro3* and *Dro4* and the 3'-flanking region of *Drs*. There is a repetitive sequence in the 3'-flanking region of *Drs* (Figure 5). These results suggested that gene duplications took place in the formation of *drosomycin* family. On the basis of Nei's birth-and-death model of evolution that some duplicate genes are maintained in the genome for a long time but others are deleted or become nonfunctional by deleterious mutations [25], the loss of *Dro3* and *Dro4* and the existence of other genes in *D. yakuba* and *D. sechellia* suggested that the *drosomycin* family underwent the birth-and-death model of evolution in *melanogaster* species-subgroup.

The phylogenetic relationship as described in the Figure 6 indicated that *D. ananassae* species-subgroup might share the common *drosomycin* ancestor with *D. melanogaster* before the speciation event which led to two separate species-subgroups. After splitting of two species-subgroups, the gene duplications happened and from the *drosomycin* common ancestor to 4 genes in *D. ananassae*; from the

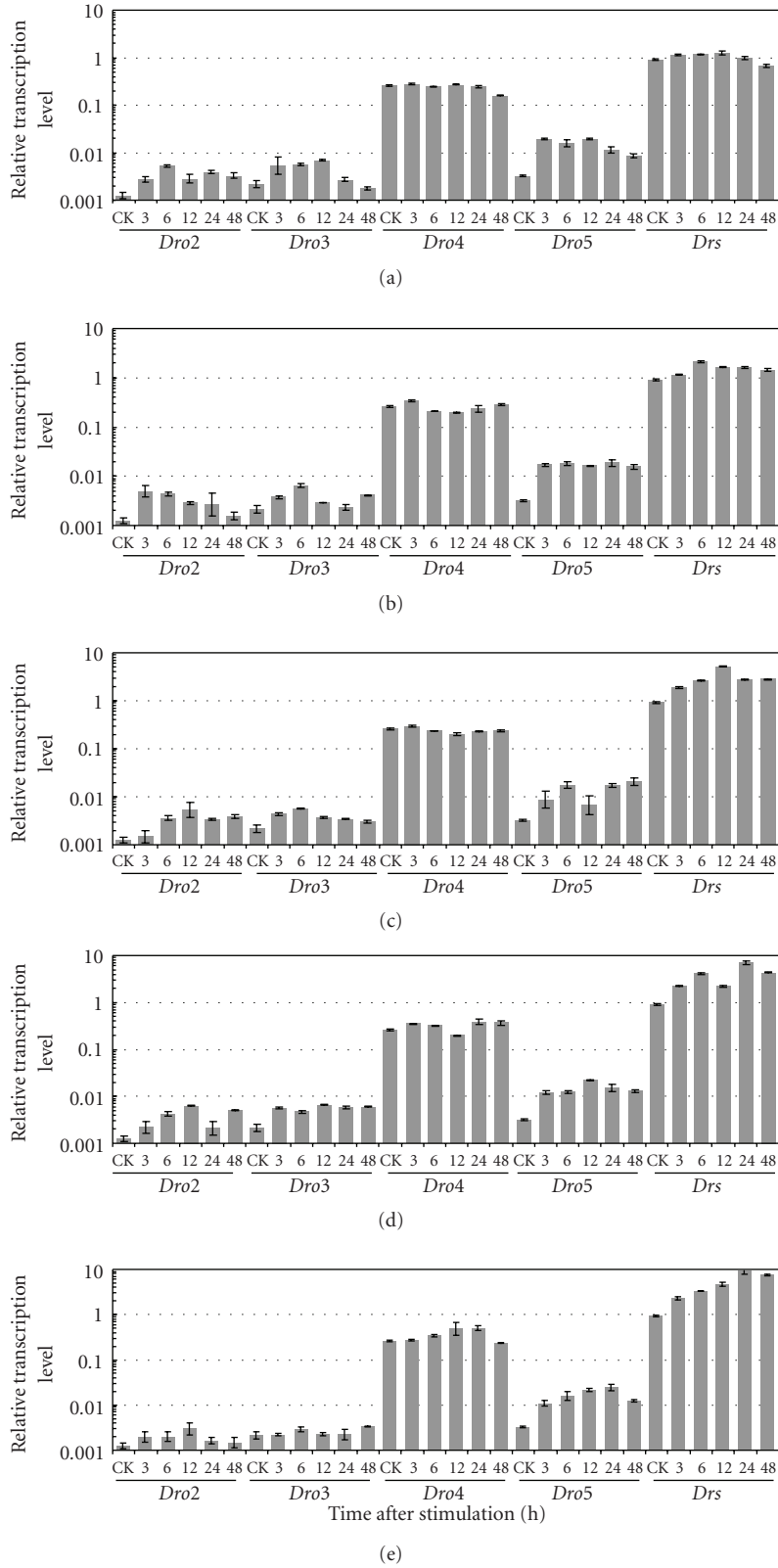


FIGURE 1: The transcription activity of drosomycin gene family in *D. melanogaster* stimulated by microbe-free injury and various microorganisms. Quantitative real time RT-PCR was performed for the detection of the expression levels of drosomycin genes in wild-type Oregon-R flies with SYBR Green. The expression levels of drosomycin genes were normalized to *rp49* in the samples. The experiments were done in triplicate and the error bars represented standard deviation. CK: native flies. (a) The 2-day-old adult flies were pricked by a sterile needle (microbe-free injury). (b) to (e) The 2-day-old adult flies were infected with *E. coli* K₁₂D₃₁, *Staphylococcus aureus*, *Fusarium oxysporum*, and *Beauveria bassiana*, respectively.

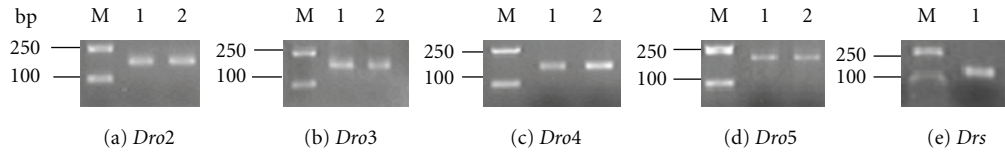


FIGURE 2: Detection of 5' RACE products of *drosomyacin* genes. The fragments were visualized by staining with ethidium bromide after separation in 2% agarose gel. M. DNA marker. Both lanes 1 and 2 were the 5' RACE PCR products.

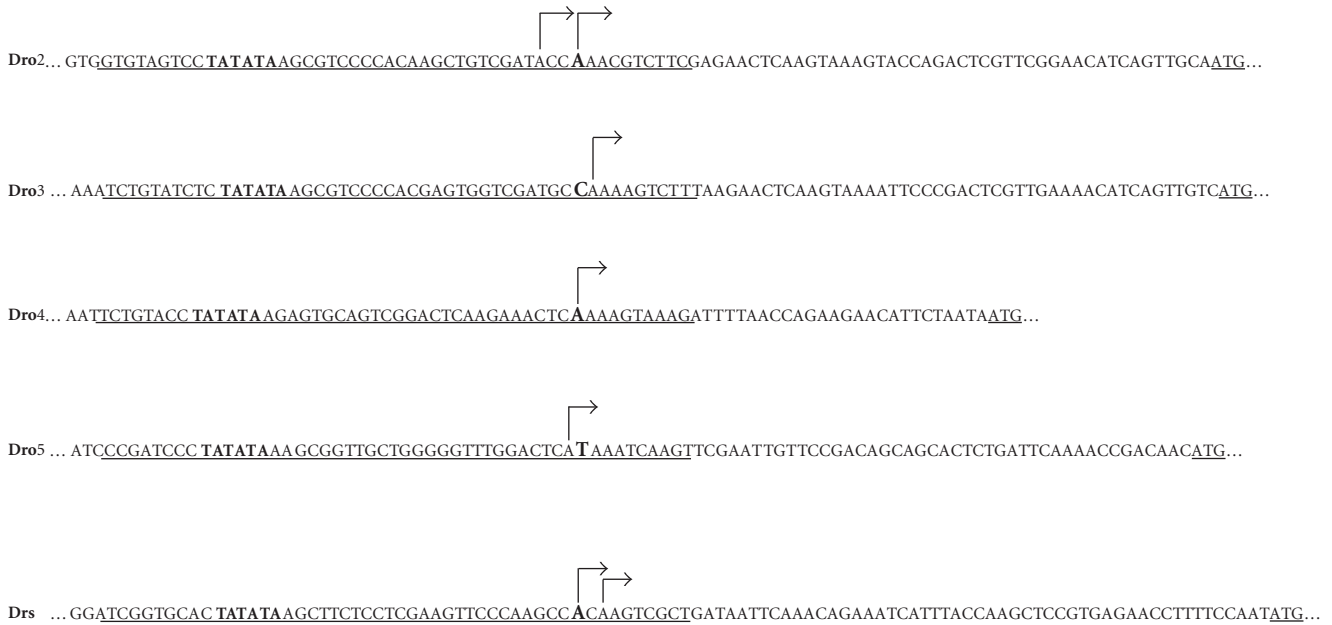


FIGURE 3: Identification of transcription start sites of *drosomyacin* genes by 5' RACE. The underlined ATG indicated the start of open reading frames (ORFs). Arrows indicated the transcription start sites and transcriptional directions. The putative core promoters predicted by NNPP software were underlined. And the predicated transcription start sites were indicated in larger fonts. The TATA boxes in the promoter were shown in bold.

drosomyacin common ancestor to 6-7 paralogs in each species of *melanogaster* species-subgroup, and the gene duplications took place before the speciation events of the *melanogaster* species-subgroup.

4. Discussion

The *drosomyacin* gene family in *D. melanogaster* comprises seven genes. For such a multigene family, expression patterns of individual genes are extremely difficult to be analyzed because of the high degree of sequence identity. In the present study, we used the real time RT-PCR with designed specific primers to uncover the transcriptional pattern of each gene after microbial challenges and physical stimulation. We provide evidence that the *drosomyacin* genes may play a critical role in response not only to microbial infections but also to physical injury in the adult stage. Of particular interest, *Drs* was the only gene that could be significantly upregulated by microbial challenges, especially by fungi and Gram-positive bacteria (Figures 1(b)–1(e)). Further promoter analysis revealed that there were two transcription start sites and three putative NF- κ B sites in 5'-flanking region of the ORF of *Drs* gene. Such motifs should be important

for response following microbial infections. In contrast to *Drs*, the expressions of *Dro2*, *Dro3*, and *Dro5* appeared to be upregulated by simple injury but not by microbial infection (Figure 1). As a matter of fact, the NF- κ B site was also found in 5'-flanking region of *Dro2*, *Dro3*, and *Dro5*, and we propose that the physical injury may also activate a signal transduction pathway that is related to the Rel family. Lacking of NF- κ B site in *Dro4* may explain why this gene with high level of constitutive expression is noninducible by both microbe-free injury and microbial infection. Moreover, the gene expression of *Dro1* and *Dro6* could not be detected in either control flies or microbe-free injured or microbial infected flies. The analysis of promoter sequence indicated that NF- κ B site and transcription start sites were not found at the upstream of *Dro1* and *Dro6*.

Genomic analyses of model organisms have shown that over one-third of all protein-coding genes belong to multi-gene family [26, 27] originating from the gene duplication [28]. In *Drosophila* genomes, genes can be duplicated via retroposition and DNA-based duplication [28, 29]. In *D. melanogaster*, most of genes coding for AMPs, for example, *cecropins* [30], *attacins* [31], *dipterocins* [32], *lysozymes* [33], and *drosomyacins* [11], were known to be clustered into

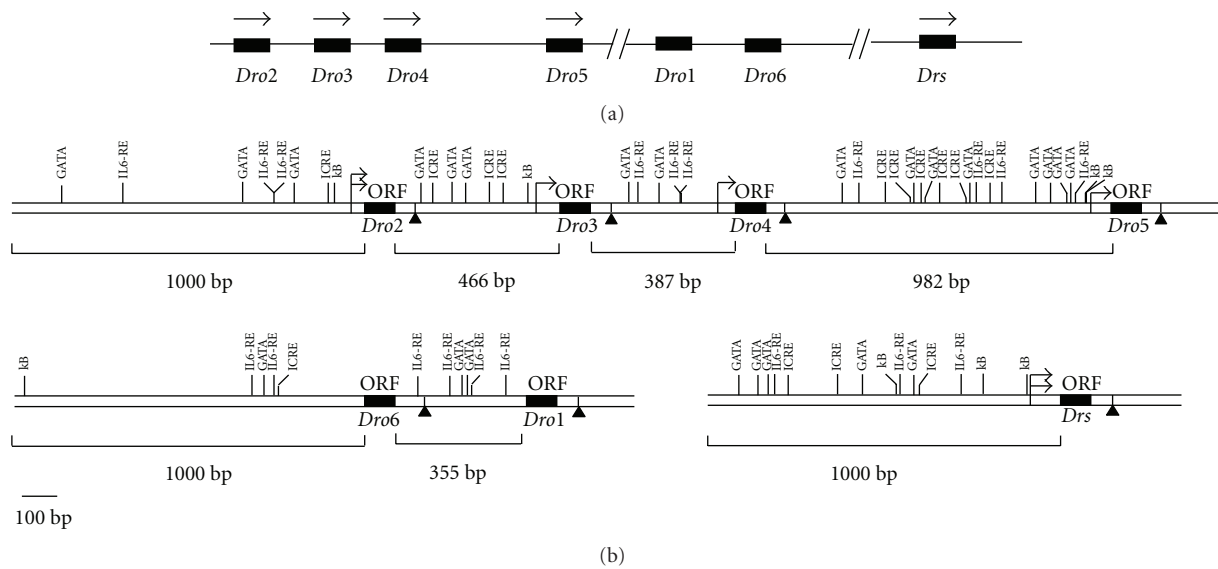


FIGURE 4: The schematic organization of the transcription factor binding sites upstream *drosomycin* genes in *D. melanogaster*. (a) General position of *drosomycin* genes on the 3L chromosome. (b) Organization of binding sites for putative transcription factors at the upstream region of *drosomycin* genes. κB: NF-κB site, GGGRAYYYYY of consensus sequence (there was no sites with perfectly matched sequence of GGGRAYYYYY. Sites with one nucleotide mismatch were given in the figure); GATA: binding site of the GATA factor, WGATAR of consensus sequence; IL-6 RE: sites for interleukin-6 response element with TKNNNGNAAK of consensus sequence; ICRE: sites for interferon consensus element with GGAAANN of consensus sequence. Black boxes indicated ORFs of *drosomycin* genes. Arrows indicated transcription start sites (TSSs) and arrows with double heads represented double TSSs. Bolded triangle (▲) indicated polyadenylation signal.

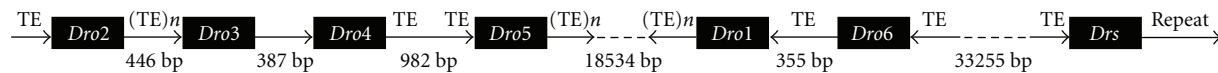


FIGURE 5: The distribution of TEs and the repetitive sequence in the flanking regions of *drosomycin* family in *D. melanogaster*. ORFs of *drosomycins* were indicated by black squares and the transcription directions were indicated by arrows. Number below the arrows indicated the intergenic length. The location of TEs and repetitive sequence was marked above the arrows. (TE)*n* meant that there was more than one TE in the region.

multigene families. In the *drosomycin* gene family of *D. melanogaster*, one repetitive sequence at the 3'-flanking region of *Drs* and TEs in flanking regions of most genes (Figure 5) suggested that gene duplications were involved in the formation of *drosomycin* multigene family. The gene duplication might occur during 12.8–44.2 million years ago, because species within the *melanogaster* species-subgroup diverged 12.8 million years ago, and the *melanogaster* species-subgroup and the *ananassae* species-subgroup separated from each other 44.2 million years ago [34]. At this time, the ancestor gene evolved into the *Drs* at the 63D2 of 3L in the *D. melanogaster*, and then gene duplication led to form the cluster 2 and cluster 3 in the 63D1 of 3L. The gene arrangement of cluster 2 indicated that the intergenic length of these genes were very small that from 387 bp to 982 bp, the tandem gene duplication might give rise to these genes.

The evolutionary fates of these duplication genes were very different. In the ORF of *Dro1*, 10% of DNA sequences in the populations of *D. melanogaster* mutated resulting in the loss of a disulphide bridge; 5% of them had an internal stop codon [10]. The expression product generated from

Dro6 gene with no antifungal activity could be due to the insert of two amino acid residues [11]. From the above evidences, we propose that *Dro1* and *Dro6* of cluster 2 have evolved as the pseudogenes. This is one more example of nonfunctionalization of duplicate immune genes. The similar event of nonfunctionalization of duplicative genes also happened to cecropin and attacin gene family in *D. melanogaster* [29, 32]. For the *Dro2*, *Dro3*, *Dro4*, and *Dro5* of cluster 3, their expression product displayed antifungal activity against 3–5 strains of tested fungi [11]. Therefore, we proposed that the fate of *Dro2*, *Dro3*, *Dro4*, and *Dro5* was of subfunctionalization.

Neofunctionalization is another fate of duplication gene [33]. The *Bmg1v2-4*, which duplicated from the ancestor gene *Bmg1v1* in the gloverin gene family of silkworm (*Bombyx mori*), gained a new function in the embryonic stage due to an intron loss [34]. Does the high level of constitutive transcription of *Dro4* identified in present study suggest a new function beyond the immune defense? Or are the high expression levels of *Drs* and *Dro4* in the native flies just for battling pathogens? How other genes in the *drosomycin* gene family perform the synergistic effect with the predominate

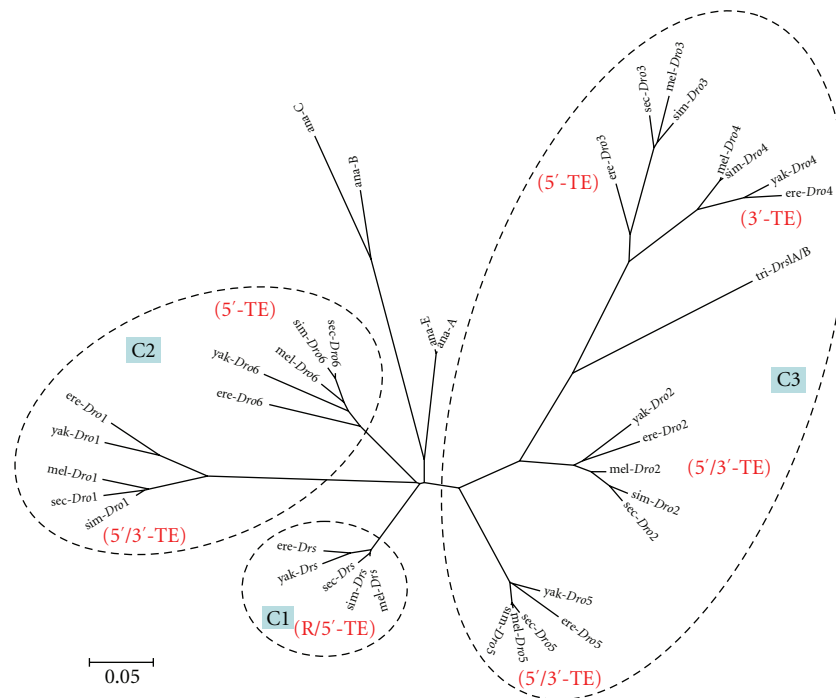


FIGURE 6: Neighbor-Joining tree of drosomycin gene family in *Drosophila* species. The tree was reconstructed by the NJ method implemented in Mega 4 with 1000 bootstrap replicates. Gaps were pairwise deleted, and the Maximum Composite Likelihood model was applied to estimate the branch length. Dot cycles with the light blue green square enclosed genes in the same cluster: the C1 is the cluster 1 at the 63D2 of 3L; C2 and C3 are the cluster 2 and cluster 3 at the 63D1 of 3L. TEs and repetitive sequence in *D. melanogaster* were shown in red beside each branch: R indicated the repetitive sequence, and 5'-TE was the TE at the 5'-flanking region, 3'-TE was the TE at the 3'-flanking region, and 5'/3'-TE was TEs at both flanking regions. Abbreviation for species: mel: *D. melanogaster*; sim: *D. simulans*; sec: *D. sechellia*; yak: *D. yakuba*; ere: *D. erecta*; ana: *D. ananassae*; tri: *D. triauraria*.

gene *Drs* to combat with the invading pathogens or exoteric stimulations? These questions are worth further investigation.

National Important Project of China (no. 2006AA10A119), and the project of transformation fund for achievement of science and technology of agriculture (2007GB2E000240).

Abbreviations

AAP:	Abridged anchor primer
AMPs:	Antimicrobial peptides
AUAP:	Abridged universal amplification primer
Drs:	Drosomycin
GSP:	Gene-specific primer
ICRE:	Interferon consensus element
IL6-RE:	Interleukin-6 response element
LB:	Luria broth
ORF:	Open reading frame
PDA:	Potato dextrose agar
rp49:	Ribosome protein 49
RACE:	Rapid amplification of cDNA ends
RT-PCR:	Reverse transcription-polymerase chain reaction
TdT:	Terminal deoxynucleotidyl transferase
TSS:	Transcription start site.

Acknowledgments

This work was supported by Grants of the 973 National Basic Research Program of China (no. 2005CB121000), the 863

References

- [1] P. Irving, L. Troxler, T. S. Heuer, et al., "A genome-wide analysis of immune responses in *Drosophila*," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 26, pp. 15119–15124, 2001.
- [2] J. A. Hoffmann, "The immune response of *Drosophila*," *Nature*, vol. 426, no. 6962, pp. 33–38, 2003.
- [3] J. L. Dimarcq, D. Hoffmann, M. Meister, et al., "Characterization and transcriptional profiles of a *Drosophila* gene encoding an insect defensin. A study in insect immunity," *European Journal of Biochemistry*, vol. 221, no. 1, pp. 201–209, 1994.
- [4] P. Bulet, J.-L. Dimarcq, C. Hetru, et al., "A novel inducible antibacterial peptide of *Drosophila* carries an O-glycosylated substitution," *The Journal of Biological Chemistry*, vol. 268, no. 20, pp. 14893–14897, 1993.
- [5] M. Charlet, M. Lagueux, J. M. Reichhart, D. Hoffmann, A. Braun, and M. Meister, "Cloning of the gene encoding the antibacterial peptide drosocin involved in *Drosophila* immunity: expression studies during the immune response," *European Journal of Biochemistry*, vol. 241, no. 3, pp. 699–706, 1996.
- [6] E. A. Levashina, S. Ohresser, P. Bulet, J. M. Reichhart, C. Hetru, and J. A. Hoffman, "Metchnikowin, a novel

- immune-inducible proline-rich peptide from *Drosophila* with antibacterial and antifungal properties," *European Journal of Biochemistry*, vol. 233, no. 2, pp. 694–700, 1995.
- [7] P. Fehlbaum, P. Bulet, L. Michaut, et al., "Insect immunity: septic injury of *Drosophila* induces the synthesis of a potent antifungal peptide with sequence homology to plant antifungal peptides," *The Journal of Biological Chemistry*, vol. 269, no. 52, pp. 33159–33163, 1994.
- [8] L. Michaut, P. Fehlbaum, M. Moniatte, A. Van Dorsselaer, J.-M. Reichhart, and P. Bulet, "Determination of the disulfide array of the first inducible antifungal peptide from insects: drosomycin from *Drosophila melanogaster*," *FEBS Letters*, vol. 395, no. 1, pp. 6–10, 1996.
- [9] C. Landon, P. Sodano, C. Hetru, J. Hoffmann, and M. Ptak, "Solution structure of drosomycin, the first inducible antifungal protein from insects," *Protein Science*, vol. 6, no. 9, pp. 1878–1884, 1997.
- [10] F. M. Jiggins and K. W. Kim, "The evolution of antifungal peptides in *Drosophila*," *Genetics*, vol. 171, no. 4, pp. 1847–1859, 2005.
- [11] W. Y. Yang, S. Y. Wen, Y. D. Huang, et al., "Functional divergence of six isoforms of antifungal peptide Drosomycin in *Drosophila melanogaster*," *Gene*, vol. 379, pp. 26–32, 2006.
- [12] B. Lemaître, J. M. Reichhart, and J. A. Hoffmann, "*Drosophila* host defense: differential induction of antimicrobial peptide genes after infection by various classes of microorganisms," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no. 26, pp. 14614–14619, 1997.
- [13] S. Uttenweiler-Joseph, M. Moniatte, M. Lagueux, A. Van Dorsselaer, J. A. Hoffmann, and P. Bulet, "Differential display of peptides induced during the immune response of *Drosophila*: a matrix-assisted laser desorption ionization time-of-flight mass spectrometry study," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 19, pp. 11342–11347, 1998.
- [14] F. Levy, D. Rabel, M. Charlet, P. Bulet, J. A. Hoffmann, and L. Ehret-Sabatier, "Peptidomic and proteomic analyses of the systemic immune response of *Drosophila*," *Biochimie*, vol. 86, no. 9–10, pp. 607–616, 2004.
- [15] E. De Gregorio, P. T. Spellman, G. M. Rubin, and B. Lemaître, "Genome-wide analysis of the *Drosophila* immune response by using oligonucleotide microarrays," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 22, pp. 12590–12595, 2001.
- [16] C. Tian, B. Gao, M. d.C. Rodriguez, H. Lanz-Mendoza, B. Ma, and S. Zhu, "Gene expression, antiparasitic activity, and functional evolution of the drosomycin family," *Molecular Immunology*, vol. 45, no. 15, pp. 3909–3916, 2008.
- [17] M. G. Reese, "Application of a time-delay neural network to promoter annotation in the *Drosophila melanogaster* genome," *Computers and Chemistry*, vol. 26, no. 1, pp. 51–56, 2001.
- [18] D. Hultmark, "Immune reactions in *Drosophila* and other insects: a model for innate immunity," *Trends in Genetics*, vol. 9, no. 5, pp. 178–183, 1993.
- [19] L. Kadalayil, U. M. Petersen, and Y. Engström, "Adjacent GATA and κ B-like motifs regulate the expression of a *Drosophila* immune gene," *Nucleic Acids Research*, vol. 25, no. 6, pp. 1233–1239, 1997.
- [20] X. Zhou, T. Nguyen, and D. A. Kimbrell, "Identification and characterization of the *Cecropin* antibacterial protein gene locus in *Drosophila virilis*," *Journal of Molecular Evolution*, vol. 44, no. 3, pp. 272–281, 1997.
- [21] E. A. Levashina, S. Ohresser, B. Lemaître, and J.-L. Imler, "Two distinct pathways can control expression of the gene encoding the *Drosophila* antimicrobial peptide metchnikowin," *Journal of Molecular Biology*, vol. 278, no. 3, pp. 515–527, 1998.
- [22] K. Tamura, J. Dudley, M. Nei, and S. Kumar, "MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0," *Molecular Biology and Evolution*, vol. 24, no. 8, pp. 1596–1599, 2007.
- [23] M. S. Dushay, J. B. Roethel, J. M. Chaverri, et al., "Two attacin antibacterial genes of *Drosophila melanogaster*," *Gene*, vol. 246, no. 1–2, pp. 49–57, 2000.
- [24] K. Senger, G. W. Armstrong, W. J. Rowell, J. M. Kwan, M. Markstein, and M. Levine, "Immunity regulatory DNAs share common organizational features in *Drosophila*," *Molecular Cell*, vol. 13, no. 1, pp. 19–32, 2004.
- [25] M. S. Nei, X. Gu, and T. Sitnikova, "Evolution by the birth-and-death process in multigene families of the vertebrate immune system," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no. 15, pp. 7799–7806, 1997.
- [26] R. P. Meisel, "Repeat mediated gene duplication in the *Drosophila pseudoobscura* genome," *Gene*, vol. 438, no. 1–2, pp. 1–7, 2009.
- [27] A. G. Clark and L. Wang, "Molecular population genetics of *Drosophila* immune system genes," *Genetics*, vol. 147, no. 2, pp. 713–724, 1997.
- [28] B. P. Lazzaro and A. G. Clark, "Evidence for recurrent paralogous gene conversion and exceptional allelic divergence in the Attacin genes of *Drosophila melanogaster*," *Genetics*, vol. 159, no. 2, pp. 659–671, 2001.
- [29] M. Hedengren, K. Borge, and D. Hultmark, "Expression and evolution of the *Drosophila* attacin/diptericin gene family," *Biochemical and Biophysical Research Communications*, vol. 279, no. 2, pp. 574–581, 2000.
- [30] S. Daffre, P. Kylsten, C. Samakovlis, and D. Hultmark, "The lysozyme locus in *Drosophila melanogaster*: an expanded gene family adapted for expression in the digestive tract," *Molecular and General Genetics*, vol. 242, no. 2, pp. 152–162, 1994.
- [31] K. Tamura, S. Subramanian, and S. Kumar, "Temporal patterns of fruit fly (*Drosophila*) evolution revealed by mutation clocks," *Molecular Biology and Evolution*, vol. 21, no. 1, pp. 36–44, 2004.
- [32] S. Ramos-Onsins and M. Aguadé, "Molecular evolution of the Cecropin multigene family in *Drosophila*: functional genes vs. pseudogenes," *Genetics*, vol. 150, no. 1, pp. 157–171, 1998.
- [33] M. Lynch and J. S. Conery, "The evolutionary fate and consequences of duplicate genes," *Science*, vol. 290, no. 5494, pp. 1151–1155, 2000.
- [34] N. Mrinal and J. Nagaraju, "Intron loss is associated with gain of function in the evolution of the gloverin family of antibacterial genes in *Bombyx mori*," *The Journal of Biological Chemistry*, vol. 283, no. 34, pp. 23376–23387, 2008.