Published in final edited form as:

J Insect Physiol. 2007 March ; 53(3): 246–253.

***Drosophila melanogaster* Methoprene-tolerant (Met) gene homologs from three mosquito species: members of PAS transcriptional factor family**

Shaoli Wang, Aaron Baumann, and Thomas G. Wilson *

Department of Entomology, Ohio State University, Columbus, OH 43210, USA

Abstract

The *Methoprene-tolerant* (*Met*) gene in *Drosophila melanogaster* has been shown to function in juvenile hormone (JH) action. *Met* homologs were isolated from three mosquito species, *Culex pipiens*, *Aedes aegypti* and *Anopheles gambiae*. Sequence similarity was found to be high in bHLH and PAS conserved domains, and the majority of the 7–9 introns in *AaMet* and *AgMet* are located in either identical or similar positions, indicating evolutionary relatedness. Sequence comparison with *Met* and the similar *germ-cell expressed* (*gce*) gene in *D. melanogaster* showed that the mosquito genes are more similar to *gce* than to *Met*. Moreover, the multiple introns in *AgMet* and *AaMet* are more similar in number with the 7 introns in *Dmgce* than to the single intron in *DmMet*; in fact, six intron positions in *AaMet* and *AgMet* are similar to those in *Dmgce*. Efforts to identify a second homologous gene in mosquitoes were unsuccessful, suggesting a single gene in lower Diptera, consistent with the single gene uncovered in genomic sequencing of *Ae. aegypti* and *An. gambiae*. These results suggest that a gene duplication occurred during the evolution of higher Diptera, resulting in *Met* and *gce*.

Keywords

Methoprene-tolerant gene; mosquitoes; intron locations; phylogenetic analysis

1. Introduction

Juvenile hormone (JH) is involved in a variety of important functions in insects, including development, reproduction, caste determination, and behavior (Riddiford, 1985,1994;Nijhout and Wheeler, 1982). Although JH is a necessary molecule at certain time in insect development, it is toxic when present during early metamorphosis in certain insects, especially Diptera (Wilson, 2004). Several chemical companies have taken advantage of this characteristic and synthesized chemical analogs of JH (JHA) having insecticidal activity for certain insects (Staal, 1975;Mulla, 1995;Ritchie and Broadsmith, 1997). Perhaps the most successful JH analog is methoprene (isopropyl 11-methoxy-3, 7, 11-trimethyl-2, 4-dodecadienoate) (Henrick et al., 1973), which is very effective against dipteran insects (Staal, 1975;O'Donnell and Klowden, 1997). However, the wide use of methoprene for mosquito control in recent years is perhaps exacting a toll: resistance evolution. Methoprene resistance has been shown in two Florida

*: Author for correspondence. E-mail: wilson.1457@osu.edu, Phone: 614-247-6036, Fax: 614-292-2180

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

populations of *Aedes taeniorhynchus* (Dame et al., 1998) and in California populations of *Ochlerotatus nigromaculis* (Cornel et al., 2000;2002).

The *Methoprene-tolerant (Met)* gene was identified due to its involvement in methoprene resistance in *Drosophila melanogaster*. Examination of *Met* allele phenotypes and rescue by *Met*⁺ transgenes clearly showed *Met* to be directly responsible for JHA resistance as well as its involvement in a JH-mediated physiology, that of oogenesis (Wilson and Fabian, 1986; Wilson et al., 2006). Isolation and sequence determination of *Met* showed this gene to be a member of the bHLH-PAS family of transcriptional regulators (Ashok et al., 1998). *Met* is proposed to be a component of the elusive JH receptor, based on the mutant phenotype (Wilson and Fabian, 1986), the binding of JH III by MET, its ability for transcriptional activation (Miura et al., 2005), and its genetic interaction with the 20-hydroxyecdysone primary response gene, *Broad-Complex* (Wilson et al., 2006). Another *D. melanogaster* bHLH-PAS gene, *germ cell-expressed (gce)*, shows more than 50% homology to *Met* (Moore et al., 2000); however, a function for *gce* has not been reported. Both genes are also present in the recently sequenced and annotated *D. pseudoobscura* genome and have high homology with those genes in *D. melanogaster*.

We wanted to determine if *Met* and/or *gce* homologs exist in mosquitoes, which might allow identification of the responsible gene (s) for resistance to methoprene in mosquitoes. Using degenerate primers from known *Met* genes combined with RACE-PCR, we isolated and sequenced *Met* cDNAs in three different mosquito species. The mosquito *Met* genes have higher similarity to *gce* than to *DmMet*. We propose an evolutionary scheme for these genes in Diptera.

2. Materials and Methods

2.1. Mosquito species

Mosquitoes used in this study were *Cx. pipiens* (provided by Dr. Rebecca M Robich), *Ae. aegypti*, and *An. gambiae* (both provided by Dr. Woodbridge A Foster). The rearing conditions for *Cx. pipiens*, *Ae. aegypti*, and *An. gambiae* can be found from Robich and Denlinger (2005), Gary and Foster (2001), and Mostowy and Foster (2004), respectively. A fourth mosquito *Oc. nigromaculis* (field-caught specimens provided by Dr. Anthony J Cornel) was included in the phylogenetic analysis. The 4th-instar larva of each mosquito species was used for PCR.

2.2. Degenerate primer design and RT-PCR

The amino acid sequences of *Met* (GenBank protein accession no. T09462), CG1705-PA (NP_511126), *gce* CG6211-PA (NP_511160), and *An. gambiae* homolog ENSANGP00000010697 (XP_316059)) were compared to identify their conserved regions. The degenerate primers were designed using Consensus-Degenerate Hybrid Oligonucleotide (CODEHOP) program (Rose et al., 1998). The forward primer 5'-CGC GAT AAG CTG AACGGC WSN ATH CAR GA -3' and the reverse primer 5'-CGC ACC TCA TCC GTC ATG TAN CCN GCN AC -3' were chosen to amplify the *Met* homologs from each mosquito species.

Total RNA was extracted from fourth instar larvae using Trizol Reagent (Invitrogen). Five micrograms of total RNA were used for the first strand cDNA synthesis at 50°C in a 20 µl volume containing SuperScript III reverse transcriptase and oligo(dT)₁₂₋₁₈ (Invitrogen). RT-PCR amplification was performed using 1 µl first-strand cDNA obtained from the RT reaction as template, 1 µl (10 µM) of each primer, 0.1 µM of each dNTP, and 1.25 U High Fidelity Taq DNA polymerase (Invitrogen) in a final volume of 25 µl. PCR amplification was carried out in an Eppendorf cyclotherm with an initial denaturation step at 94°C for 3 min. Amplifications were

achieved through 35 cycles at 94°C for 30 sec, 56°C for 1 min, and 72°C for 1 min. A final extension reaction was carried out for 10 min at 72°C. PCR fragments were purified, ligated into the pCR 2.1 TOPO TA Cloning vector (Invitrogen), and transformed into TOP 10 cells. Inserts from positive clones were sequenced, and the sequencing results were used to design gene-specific primers for 5' and 3' rapid amplification of cDNA ends (RACE).

2.3. RACE

Full-length *Met* gene cDNAs were synthesized by 5' and 3' RACE (Invitrogen). One gene-specific primer for 3'RACE-PCR and 3 gene-specific primers for 5' RACE-PCR were designed in each mosquito species (Table 1) using the *Gene Fisher* software (<http://bibiserv.techfak.uni-bielefeld.de/genefisher/>).

After separation on a 0.8% agarose gel, the RACE-PCR products were cut from the gels, purified, and sequenced.

2.4. Intron positions in genomic sequences of *Ae. aegypti* and *An. gambiae*

Genomic DNA was extracted using the method of Qiao and Raymond (1995). Genomic DNA amplifications of ORFs were carried out using primer design based on the cDNA sequences obtained from RACE-PCR. cDNA sequences from *Ae. aegypti* and *An. gambiae* (*AaMet* and *AgMet* respectively) were arbitrarily divided into two fragments and two pairs of gene-specific primers were designed to amplify the separate fragments. We obtained the latter half of sequence containing four introns for *AaMet* and *AgMet* and the remaining intron sequences were obtained when genomic sequence data of *Ae. aegypti* (http://www.broad.mit.edu/annotation/disease_vector/aedes_aegypti/) and *An. gambiae* (<http://www.ncbi.nlm.nih.gov>) became available.

2.5. Phylogenetic tree construction

Multiple alignments were performed using Clustal X 1.83 (Thompson et al., 1997). Amino acids at the 3' and 5' termini were trimmed, resulting in a sequence alignment spanning the HLH domain through the majority of the PASB domain. Portions of the alignment containing multiple gaps were manually removed. Phylogenetic tree construction was done with PAUP* 4.0 for Windows (Swofford, 2002) using the parsimony method. The eight in-group sequences include *Met* homologs of four mosquito species and *Met* and *gce* sequences from *D. melanogaster* and *D. pseudoobscura*. Three bHLH-PAS genes obtained via BLAST searches, *Dmcycle* (bHLH-PAS gene from *D. melanogaster*), *XlARNT2*, and *DrARNT2* (bHLH-PAS genes from *Xenopus laevis* and *Danio rerio*, respectively), were assigned as outgroups. An exhaustive search produced a single tree, which was opened with the TreeView program (Page, 1996).

3. Results

3.1. Isolation of *Met* full-length cDNAs in mosquitoes

Based on the conserved amino acid sequences among *DmMet*, *Dmgce*, and a partial homologous sequence from *An. gambiae*, a pair of degenerate primers were designed and used for RT-PCR amplification of a fragment including the bHLH through PASB domains. A single gel band of ~900 bp was obtained from each mosquito species, whose translated sequence showed similarity to those of *DmMet* and *Dmgce*.

We used each of these sequences to design primers for 3'RACE and 5'RACE amplifications. After sequence alignment of the products, the full-length cDNA of each species was constructed. The *Cx. pipiens* (*CpMet*) cDNA (AY895165) is 3,109 bp in length and contains an open reading frame encoding a protein of 854 amino acid residues with a predicted MW of

93.3 kDa. Before the start codon, there are 94 nucleotides of 5' untranslated region (5'UTR) upstream. A TAG stop codon is followed by 453 nucleotides of 3' untranslated sequence, which includes the canonical AATAAA polyadenylation signal upstream from an 18 bp poly (A) tail.

The *AaMet* cDNA (AY902310) is 3,135 bp in length and contained an open reading frame of 2,727 bp encoding a protein of 909 amino acids with MW of 98.4 kDa. There are 123 bp nucleotides of 5'UTR and 285 bp of 3'UTR sequences containing 20 bp of poly (A) tail. The *AgMet* cDNA (DQ303468) is 4,782 bp including an open reading frame of 3,348 bp encoding a protein of 1,116 amino acids of MW 119.4 kDa. There are 189 bp nucleotides of 5'UTR and 1,245 bp of 3'UTR sequences including 28 bp of poly (A) tail.

3.2. Sequence analysis of *Met* genes in mosquitoes

All three mosquito *Met* genes showed (1) primary domains characteristic of bHLH-PAS proteins: bHLH, PASA, and PASB and (2) sequence homology with each other, being highest (64%) between *CpMet* and *AaMet* (Fig. 1). Sequence identity between each mosquito gene and *DmMet* is 43–45%, and *Dmgce* is 48–57%. The bHLH conserved domains show much higher identities (70–97%), favoring *gce* (84–97%) more than *Met* (70–83%) (Table 2).

3.3. A single *Met* homologous gene found in *Ae. aegypti*

PCR amplification with degenerate primers in each mosquito species consistently gave one band as described and shown in Fig. 2. To determine if a second *Met* homolog with higher similarity to *Met* than to *gce* is present, two pairs of *Met*-specific primers were designed to specifically target a *Met* homolog from *Ae. aegypti*. Two forward primers were FP1 (5' CCC AGT CTS CAT CTR ACG GAC 3') and FP2 (5' ATM GAG ACS CTG TTC TAT CAR CA 3') were paired with one reverse primer RP (5' CAG CTG RTA SGG TTC YGG CTG 3') and used for amplification. However, no band was amplified from the mosquito species (Fig. 2, lanes 6 and 7), although the expected *Met* sequence was amplified from *D. melanogaster* (Fig. 2, lanes 4 and 5).

3.4. Comparison of intron positions among *AaMet*, *AgMet*, and *gce*

We identified and examined introns of *AeMet* and *AgMet*. In *DmMet*, there is only one intron located in the PASB domain (Ashok et al., 1998), but in *Dmgce*, 7 introns (none in the PASB domain) are present. The intron numbers and locations in *AaMet* and *AgMet* are more similar to *gce* than to *Met* in *D. melanogaster*: there are 9 and 7 introns for *AaMet* and *AgMet*, respectively. Four introns are located in the same positions in *AaMet* and *AgMet*, and 6 introns are located in similar positions in *Dmgce*, *AaMet*, and *AgMet* (Fig. 1). One intron located in the PASA region of *AaMet* and *AgMet* is not present in *gce*.

3.5. Phylogenetic analysis

Sequences used in this analysis spanned the bHLH region through the PASB domain, a total of 286 amino acids. The six genes examined (*OnMet*, *AaMet*, *CpMet*, *AgMet*, *Dmgce*, *Dpgce*) were found to form a sister clade to the *Met* group from *Drosophila* (*DmMet*, *DpMet*), a result consistent with initial observations of amino acid identities among these sequences. An exhaustive search using the parsimony criterion produced a single tree (Fig. 3) indicating that *gce* and the mosquito *Met* homologs diverge from a distinct node, which shares a common ancestor with *Drosophila Met*. An identical tree topology was obtained using nucleotide characters (data not shown).

4. Discussion

Mosquitoes transmit diseases affecting millions of humans worldwide. Vector control by insecticides is still important for disease control. Methoprene has been widely used for control of mosquitoes in the past 25 years (Schaefer et al., 1975). Resistance to methoprene has appeared in *Ae. taeniorhynchus* (Dame et al., 1998) and *Oc. nigromaculis* (Cornel et al., 2002), but the resistance mechanism is unknown. Identifying the mechanisms of resistance to methoprene in mosquitoes will aid in mosquito control.

In order to identify the potential methoprene resistance genes in mosquito species, *Met* homologs have been isolated from *Ae. aegypti*, *An. gambiae*, and *Cx. pipiens* using CODEHOP degenerate primers (Rose et al., 1998). The degenerate primers in this study were designed based on bHLH and PASB sequences to allow amplification of both *DmMet* and *Dmgce* homologs simultaneously (Fig. 2, lane 2). However, in each mosquito species, only one band amplified (Fig. 2, lane 3), suggesting only one *Met* homologous gene in each mosquito species, a result corroborated after searches of the genomic sequencing data of *Ae. aegypti* and *An. gambiae*.

High identity between *Met* and *gce* suggests that a gene duplication event occurred in higher Diptera, producing *Met* from a *gce* template. The similarities in the intron positions among *gce* and the putative mosquito *Met* homologs lend support to this hypothesis. It is also possible (though less likely) that the duplication occurred before the divergence of higher and lower Diptera, and one copy was subsequently lost in lower Diptera. Future study of homologous sequences from other dipteran species will serve to further elucidate the evolutionary history of these genes.

Although the function of *gce* in *D. melanogaster* is unknown, GCE can dimerize with MET in the absence of JH or JH agonist (Godlewski et al., 2006), indicating some function in common with *Met*. Since methoprene applied to *Ae. aegypti* results in toxicity and morphogenetic defects (Braga et al., 2005) similar to those seen for *D. melanogaster* (Postlethwait 1974), we believe that these *Met* homologs function similarly in JH action and possibly resistance in mosquitoes, and may have *Dmgce* function as well. Elucidation of the function of *Dmgce* will help our understanding of mosquito *Met* functions. Neither *Met* nor *gce* appears to be a homolog of the vertebrate bHLH-PAS *aryl hydrocarbon receptor (ahr)* gene, however. AHR binds various xenobiotic compounds, such as dioxin, and subsequently transcriptionally activates certain genes, especially cytochrome P450s, involved in the degradation of these compounds (Hahn, 1998). Although the property of ligand binding by MET might suggest an evolutionary relationship of *Met* with *ahr*, the resistance studies have shown *Met* interaction with JHAs, not with a variety of chemically dissimilar insecticides (Wilson and Fabian, 1986) as might be expected for an *ahr* homolog. Based on gene homology, the *spineless (ss)* gene of *Drosophila* has been proposed as an invertebrate *ahr* (Duncan et al., 1998). Future ligand-binding studies with SS (and GCE) may help our understanding of the evolution of these genes.

In conclusion, the *Met* homologous genes isolated from these mosquito species should facilitate studies of the mechanism of methoprene resistance seen in field populations as well as roles for JH in mosquitoes.

Acknowledgements

We thank Dr. Woodbridge A Foster of The Ohio State University for generous help to provide *An. gambiae* and Dr. Rebecca M Robich for *Cx. pipiens* in the work. This work was supported by National Institutes of Health grant AI052290 to T.G.W.

References

- Ashok M, Turner C, Wilson TG. Insect juvenile hormone resistance gene homology with the bHLH-PAS family of transcriptional regulators. *Proceedings of the National Academy of Sciences, USA* 1998;95:2761–2766.
- Braga IA, Mello CB, Peixoto AA, Valle D. Evaluation of methoprene effect on *Aedes aegypti* (Diptera: Culicidae) development in laboratory conditions. *Memórias do Instituto Oswaldo Cruz* 2005;100:435–440.
- Cornel AJ, Stanich MA, Farley D, Mulligan FS, Byde G. Methoprene tolerance in *Aedes nigromaculis* in Fresno County, California. *Journal of American Mosquito Control Association* 2000;16:223–228.
- Cornel AJ, Stanich MA, McAbee RD, Mulligan FS. High level methoprene resistance in the mosquito *Ochlerotatus nigromaculis* (Ludlow) in Central California. *Pest Management science* 2002;58:791–798. [PubMed: 12192903]
- Dame DA, Wichterman GJ, Hornby JA. Mosquito (*Aedes taeniorhynchus*) resistance to methoprene in an isolated habitat. *Journal of the American Mosquito Control Association* 1998;14:200–203. [PubMed: 9673923]
- Duncan DM, Burgess EA, Duncan I. Control of distal antennal identity and tarsal development in *Drosophila* by *spineless-aristopedia*, a homolog of the mammalian dioxin receptor. *Genes & Development* 1998;12:1290–1303. [PubMed: 9573046]
- Gary RE Jr, Foster WA. Effect of available sugar on the reproductive fitness and vectorial capacity of the malaria vector *Anopheles gambiae* (Diptera: Culicidae). *Journal of Medical Entomology* 2001;38:22–28.
- Godlewski J, Wang SL, Wilson TG. Interaction of bHLH-PAS proteins involved in juvenile hormone reception in *Drosophila*. *Biochemical and Biophysical Research Communications* 2006;342:1305–1311. [PubMed: 16516852]
- Hahn ME. The aryl hydrocarbon receptor: A comparative perspective. *Comparative Biochemistry and Physiology Part C* 1998;121:23–53.
- Henrick CA, Staal GB, Siddall JB. Alkyl 3, 7, 11-trimethyl-2, 4-dodecadienoates, a new class of potent insect growth regulators with juvenile hormone activity. *Journal of Agricultural and Food Chemistry* 1973;21:354–359. [PubMed: 4708794]
- Klowden MJ. Endocrine aspects of mosquito reproduction. *Archives of Insect Biochemistry and Physiology* 1997;35:491–512.
- Miura K, Oda M, Makita S, Chinzei Y. Characterization of the *Drosophila* *Methoprene-tolerant* gene product. *FEBS Journal* 2005;272:1169–1178. [PubMed: 15720391]
- Moore AW, Barbel S, Jan LY, Jan YN. A genomewide survey of basic helix-loop-helix factors in *Drosophila*. *Proceedings of National Academy of Sciences, USA* 2000;97:10436–10441.
- Mulla MS. The future of insect growth regulators in vector control. *Journal of American Mosquito Control Association* 1995;11:269–273.
- Nijhout HF, Wheeler DE. Juvenile hormone and the physiological basis of insect polymorphism. *Quarterly Review of Biology* 1982;57:109–133.
- O'Donnell PP, Klowden MJ. Methoprene affects the rotation of the male terminalia of *Aedes aegypti* mosquitoes. *Journal of American Mosquito Control Association* 1997;13:1–4.
- Page RDM. TREEVIEW: An application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences* 1996;12:357–358. [PubMed: 8902363]
- Paskewitz SM, Gorman MJ. Mosquito immunity and malaria parasites. *American Entomologist* 1999;45:80–94.
- Postlethwait JH. Juvenile hormone and the adult development of *Drosophila*. *Biological Bulletin* 1974;147:119–135. [PubMed: 4210845]
- Qiao CL, Raymond M. The same esterase B1 haplotype is amplified in insecticide resistant mosquitoes of the *Culex pipiens* complex from the Americas and China. *Heredity* 1995;74:339–345. [PubMed: 7538988]
- Riddiford, LM. Hormone action at the cellular level. In: Kerkut, G.; Gilbert, LI., editors. *Comprehensive insect biochemistry, physiology, and pharmacology*. 8. Pergamon; New York: 1985. p. 37–84.

- Riddiford LM. Cellular and molecular actions of juvenile hormone I. General considerations and premetamorphic actions. *Advances in Insect Physiology* 1994;24:213–274.
- Ritchie SA, Broadsmith G. Efficacy of ALTOSID pellets and granules against *Aedes aegypti* in ornamental bromeliads. *Journal of the American Mosquito Control Association* 1997;13:201–201. [PubMed: 9249661]
- Robich RM, Denlinger DL. Diapause in the mosquito *Culex pipiens* evokes a metabolic switch from blood feeding to sugar gluttony. *Proceedings of National Academy of Sciences, USA* 2005;102:15912–15917.
- Schaefer CH, Miura T, Wilder WH, Mulligan FS III. Evaluation of new chemicals as mosquito control agent. *Proceedings and Papers of the California Mosquito and Vector Control Association* 1975;43:75–77.
- Staal GB. Insect growth regulators with juvenile hormone activity. *Annual Review of Entomology* 1975;20:417–460.
- Swofford, DL. PAUP* 4.0: Phylogenetic Analysis Using Parsimony (*and Other Methods). Sinauer Associates; Sunderland, MA.: 2002.
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* 1997;24:4876–4882. [PubMed: 9396791]
- Rose TR, Schultz ER, Henikoff JG, Pietrokovski S, McCallum CM, Henikoff S. Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequence. *Nucleic Acids Research* 1998;26:1628–1635. [PubMed: 9512532]
- Wilson TG. The molecular site of action of juvenile hormone and juvenile hormone insecticides during metamorphosis: how these compounds kill insects. *Journal of Insect Physiology* 2004;50:111–121. [PubMed: 15019512]
- Wilson TG, Fabian J. A *Drosophila melanogaster* mutant resistant to a chemical analog of juvenile hormone. *Developmental Biology* 1986;118:190–201. [PubMed: 3095161]
- Wilson TG, Wang S, Beno M, Farkas R. Wide mutational spectrum of a gene involved in hormone action and insecticide resistance in *Drosophila melanogaster*. *Molecular Genetics and Genomics*. 2006in press
- Wilson TG, Yerushalmi Y, Donnell DM, Restifo LL. Interaction between hormonal signaling pathways in *Drosophila melanogaster* as revealed by genetic interaction between *Methoprene-tolerant* and *Broad-Complex*. *Genetics* 2006;172:253–64. [PubMed: 16204218]

AaMet -----MQRSGPSSVATTTATEPGEK----- 20
AgMet MNISTEYAPTDMRSRSSYTRSSYSAYYGGPPEIGGTGAGALGSQQHPPYGGGVETIGF 60
Dmgce -----
CpMet -----MINHAGVSYGHQMPWYGGPP----- 20
DmMet -----MAAPETGNTGSTGAGSTGSG----- 21

AaMet -----LDASTSAG----- 28
AgMet ADGGSHSRHHDRSA SMAMGGSFEGDCNEKPFYAP SPLGSDSYASDEARHSGGYLANPYY 120
Dmgce -----
CpMet -----RAYPIPPG----- 28
DmMet -----SGSGSGSG----- 29

AaMet -----
AgMet GATAGGGGGYGHQGSMMRAMPPELGMVGGGCGYSPVPVWYQLPQQQPS SYHQQHPGH 180
Dmgce -----
CpMet -----
DmMet -----

I

AaMet --NNS--A SKYDCEDSKAACK-----ILNGREARNRAEKNRRDKLNG 66
AgMet SQHHHHHHHHHPHSQQQHSASPRCYMPPEHMYNMFNFNRNGREARNRAEKNRRDKLNG 240
Dmgce --MQTAVQ-----NNGREARNRAEKNRRDKLNG 25
CpMet --HMYNMYHFNRDYEDSKAACK-----ILNGREARNRAEKNRRDKLNG 70
DmMet --SSSDP-----ANGREARNLAEKQRDKLNA 54
 ***** ***:*****.

II

AaMet SIQELS GMVPHVAES PRVDKTAVLRFSAHALRLKVVFDTQEQTKEPSEN----- 118
AgMet SIQELS AMVPHVAES PRVDKTAVLRFSAHGLRVDYVFGKSKPEE----- 285
Dmgce SIQELS TMVPHVAES PRVDKTAVLRFAAHALRLKHAFGN SLMQQ----- 70
CpMet SIQELS GMVPHVAES PRMDKTAVLRSAHALRVDYVFGKTTPKPETKPP----- 121
DmMet SIQELATMVPHAAES SRRLDKTAVLRFATHGLRLQYVFGKSASRRRKTGLKGTGMSASP 114
 ***** ***:***.***:*****:;*.**:.*.

AaMet ---AAGOKTEVHDALFRMLNGFLITVTCRGQIVLVSVASVEHFLGHCOTDLYGQNLFNLIH 175
AgMet ---TVKPEAQDSLFRMLNGFLITVTCRGQIVLVSPSVSQFLGHCOTDLYGQNLFTLTH 340
Dmgce ---RPQITDTLMDMLD SFFLT LTCHGHILLI SASIE QHLGHCQSDLYGQS IMQITH 123
CpMet ---VVVQQPEVQDTLFRMLNGFLITVTCRGQIVLVSVASVEQFLGHCOTDLYGQSLYGMVH 178
DmMet VGDLPNPSLHLDTLMLQLDCCFLT LTCSGQIVLVSTSVQQLGHCQSDLYGQNLQITH 174
 . *:.*:.*: .***:**:*.***: *****:*****. : *

AaMet PDDHNLKQQLVPMNLVNLFD SAVSAPSTSR-----TSSGTETS AEE----- 217
AgMet PDDHALKQQLIPSNLVNMF DNGSPSTSCGAPPAL LGS GGGSGTHGTG DGGGE----- 393
Dmgce PEDQNM LKQQLIPTELENLFDAHGDS-----DAEGE----- 154
CpMet PDDLALKQQLVPTNLANLFESSAPASSSS-----SGSRSAESNED----- 219
DmMet PDDQLLRQQLIPRD IETLFYQHQQHQQGHNPQGHSTTSASASGSDLEEEMETEEHR 234
 :. :*:***:* :. :* :

AaMet ---QQRKSQDEEDE-----IDRKL RQDRRKFTLR IARAGPRSEPT-AY 256
AgMet ---FOR-SYDDEEE-----IDRKL RQDRRRFTIR MARAGPRSEAT-TY 431
Dmgce ---PRQRSKAEEDA-----IDRKL REDRRSFRVRLARAGPRSEPT-AY 193
CpMet ---LQKRSRDEEAE-----IDRKL RQDRRRFNLRIARAGPRSEPT-AY 258
DmMet LGRQQGEADDDEDHP YNRRTPSPRMAHLATIDDR LRMWRCFTVRLARASTRAEATRHY 294
 : : :* ** :** * * * :*:***.***: *

AaMet ELVTIDGFRRADAA PRGERPS-----GPSG-LQLLRARG----- 291
AgMet ELVTIDGCFRRADSA PRGSSSSSSATKAGAAGAGGTAGSMQIRRVRG----- 480
Dmgce EVVKIDGCFRRSDEA PRGVRSN-----HFSSNLQLIRRTRG----- 229
CpMet EMVTVDGCFRRADAA PRGLKSS-----GPLG-LQLLRGRGG----- 294
DmMet ERVKIDGCFRRSDSSLTGAAA-----NYPIVSQILRRSRNNMMLAAAAVA 341
 * * :* ** :* : * :* :* :

III

|

AaMet -----RDDGIT--LQSIGNDIIVLVAVARVQK-----VPTICDRLEIACR 329
AgMet -----RDDAIP--LHSINGNDIVLIALARVMK-----VPSICDRLEIACK 518
Dmgce -----RDDVIP--LHTISGNDIILTGCARIIR-----PPKIASRLIDANT 267
CpMet -----REDGIS--LHSINGNDIVLVAMVRVMK-----VPTICDRLEIACK 332
DmMet AEAATVPPOHDAIAQAALHGISGNDIVLVAMARVLRERPEETE GTVGLTIYRQPEPYQ 401
 :.* *. :.*.***:* .*. : *

IV

|

AaMet YEYKTRHLIDGRIVQCDHRI SVVAGYLTTEVSGLS PFTFMHKDDVRWVIVALRQMYDYSQ 389
AgMet YEYKTRHLIDGRIVQCDQRI SIVAGYLTDEVSGLS PFTFMHRDDVRWVIVALRQMYDYNQ 578
Dmgce LEYKTRHLIDGRIIDCDQRI GIVAGYMTDEVNLS PFTFMHMDVRWVIVALRQMYDCNS 327
CpMet YEYKTRHLIDGRIVQCDHRI SVVAGYMTNEVSGLS PFTFMHKDDVRWVIVALRQMYDYNQ 392
DmMet LEYHTRHLIDGSIIDCDQRI GLVAGYMKDEVNLS PFCFMHLDDVRWVIVALRQMYDCNS 461
 :*** **:***:***: . ** .**** ** *

AaMet NYGESCYRLMTRTGD FVYLKTRGYLEVDDSSKKVQSFVCINTLVSDDEEGRRLVREMKHKF 449
AgMet -NGESCYRLMTRTGD FIYLKTRGYLEVDDSTKVQSFVCINTLVSEEEGRRLVREMKRKF 637
Dmgce SYGESTYRLFTNRNGNIYQLQSKGYLEIDKETNKVHSFVCVNTLLGEEGKRRVQEMKKF 387
CpMet PNGESCYRLMTRTGD FIYLKTRGYLEVDDSTKVQSFVCINTLVSDDEGRRLVQEMKRF 452
DmMet DYGESCYRLLSRNGRFIYLHTKGFLEVDRGSKNVHSFLCVNTLLDEEAGRQKQVQEMKEKF 521
 *** **:*. :*:***:***: :*:***:***: :*:***:***: :*:***:***

AaMet SVIVEADELPDES--DEPAVENPTQIEKAVLNLLTNLHSEDEPSEALP SNTS-QETDG 506
AgMet SVIVDKVELPDES--GEPVVENPKQIEEAVMNLITNLQPD SEDKL LNTMPASPASSIKS 695
Dmgce SVIINTQIPQSTI--DVPASEHPAL LEKAVLRLIQNLQKS GENGHDDGDEDDA QDQDD 445
CpMet SVIVDQKDLPDES--DEPAVENQQQIEQAVMNLITNFHPSNEEDVLRALP PAAS PGRSSG 510
DmMet STIIKAEMPTQSSSPDLPASQAPQQLERIVLYLIENLQKS-VDSAETVGGQGMESLMDDG 580
 .: . . *. : :*. *: *: *: . : .

AaMet S-EGSQLAIIAPSSKAVKSAIVKSMNVIAIASKKMSNG----- 543
AgMet YGEGAP LAIVAPEKN SVKSAIVKSINVVSAIAAKTMREGRC SSVSGGDWSP MGGDQQNSNG 755
Dmgce D-----EEDDDDDQDD GARSMEFGD PYGSHHGR----- 474
CpMet P-TGDGLAIIAPRPNAVKLS ILKSMRVITIASKRVSTT----- 547
DmMet Y-----SSPANTLTLEELAPS----- 596
 .:: .

AaMet -----SNSPLRSTGASPRSKDHS GSPSAIITMQRPSVLQKAH----- 581
AgMet SSSSTAS SSVSTGMPSPSRSAFADGIGSPPPPPPPSSLS PGGVPRPTVLQKLD PQLSEE 815
Dmgce -----SHHGSS-----ALSSHGHGNAKTPLALVP----- 499
CpMet -----SFS-----ETDGEP-----NNTSAIT--QRPSVLQKT----- 572
DmMet -----PTPALALVP----- 605
 *.:

AaMet ----SPALESRVPLHGTDHD FIQPFSPAGSTSSSSSSS-GSVFSPAS-----HQRN 628
AgMet AAAGVANGVEQRVPPLPNSQHYFTQPFSPSGGTPVPVSLSPASSHYS-----QRS 866
Dmgce -----PEASVKSAAITKSI SVVNVTAAKHLRGIHASTAVKS----- 535
CpMet -----ESRVPLPSSDHYFTQPYSPAGSTSSSSSNSNS SILSPASAAAATQONLFQRN 624
DmMet -----PAPSSVKSSISKSVSVNVTAARKFQGEHQKQRERD----- 641
 * . . : . . * . *.: .

AaMet SRSPFGGSIPP-TSPPAIAATPSTS VSRIV--DNTASGHSRSISV LKRTHSNCLDSS-- 683
AgMet ARSPYGGCGSGIASPPAAIHG-----GGSRTTITV LKRTY SNSS INSLNS 910
Dmgce -PSSLGSC TCSDSHSPCDFCQG-----APTITD LQAVG SNLKRST-- 574
CpMet SRSPFGGSIPPPLSP PGGSGGPMAS SASYVRGDN TASGHTRITTSV LKRTY SNSS ISSS-- 682
DmMet -----REQLKERTNSTQG----- 654
 : . *.

```

                                V
AaMet  -----DGSIGHTAD FVVQKRPKINTILDLDTCRTRLSNPATDLITMLPQAFHAVDQSL 738
AgMet  LDNYGDGGEAG-TGDSGKRARMHALGNSDLANCFSDLINPGSDLTSLLPNSFEALDQSL 969
Dmgce  -----AHVETEEKLS---KRRFIPSTEIEHVLHTSLDQIGRNL 610
CpMet  -----EGSEL-----KRPKAS-PSDL SACLSELVDPGADLTLLPDAFDAVDQTL 727
DmMet  -----VIRQLSSCLSEAEATASC 671
                                : * .:

                                VI
AaMet  KVDDDTAKLRAQVSGYENLQGKSLQ--LD SIVEASHEQRNVLS IKQELTQFQSNNGSN 796
AgMet  TMETTTTIIRDVSGYSSLHES---QNRLDRIVEEHQEQREMLQSIQEQYQLQMOSNGT- 1025
Dmgce  QQLNVARNLREQSQR YELPHAN---QRFDEIMQEHQKQSELYVNIKSEYEVQLQHKAS- 665
CpMet  TMESTTAIMRGQVGE YEELHGAQVQ QHRLDKIAEEQQEHREMLQSIQEQYTLQIQSNGG- 786
DmMet  I LSPASSLSASEAPDTPDPHSN-----TSPPPSLHTRPSVLHRTLSTLR----- 716
                                : . : . . . : .:

AaMet  DQDGAGGRGPIVVNDLLPGGGAIVDDLDPPSLMMEDFPSDSSNLMADLDNNGSLLEGDSP 856
AgMet  -----GSDAIEADNGSVGGGGGGGGGSDPNMGLCNLSQRVSTITS----- 1066
Dmgce  -----TRKSSDSRNQEQPPPLQEDDQD----- 689
CpMet  ---GAGG---AVALAIGGNVGPIRRIDP-VIMRPEGPS----- 817
DmMet  -----

AaMet  LEFFSSINTPLISPTTATASVITSSRSVGGTGTGGASGTAGSFYQGADLNPS 908
AgMet  --TMAAAATPLMMPSVITTSVGQQQQQRTAATGSAGPVAGTARSEKDAGS- 1115
Dmgce  -----VDVPAAVAVITTSVITPN-----NMSTSSSHFYHRGADLSPS 853
CpMet  -----
DmMet  -----

```

Fig 1.

Amino acid sequences comparison among *AaMet*, *AgMet*, *Dmgce*, *DmMet*, and *CpMet* and intron positions among *AaMet*, *AgMet*, and *Dmgce* using Clustal W program (<http://www.ebi.ac.uk/clustalw.html>).

Asterisks and dots denote perfect and similar identities in the alignment positions, respectively; bHLH, PASA, PASB domains are indicated by underline, bold underline, and double underline, respectively, in the *AaMet* gene.

Six conserved intron positions among *AaMet*, *AgMet*, and *Dmgce* are indicated as bold and underlined. The other intron positions are all shown in bold. One intron in *AaMet* is located in the PASA region and is the same position (162Q) as *AgMet*; the second intron location (*AaMet*, 243R) is the same as *Dmgce*, but not in *AgMet*; the third one (34K) is unique in *AaMet*.

Fig 2.

RT-PCR results using *Met* and *gce* degenerate primers or two pairs of *Met*-specific primers. Lane 1: 1 KB ladder (Biolabs).

Lane 2: Amplification using *Met* and *gce* degenerate primers in *D. melanogaster*. Top band: *DmMet*; bottom band: *Dmgce*.

Lane 3: Amplification using *Met* and *gce* degenerate primers in *Ae. aegypti*. The single band shown is *AaMet*.

Lane 4, 5: Amplification in *D. melanogaster* using each of two different pairs of *Met*-specific primers, showing the expected *DmMet* band in each lane.

Lane 6, 7: Amplification in *Ae. aegypti* using the same primer pairs as in lanes 4 and 5, showing failure of amplification.

Fig 3.

Phylogenetic tree of *DmMet* and *DpMet*, *Dmgce* and *Dpgce*, and four mosquito *Met* homologs. An exhaustive search using amino acid sequences spanning the bHLH through PASB domain produced a single most parsimonious tree. Three additional bHLH-PAS family members serve as outgroups. Bootstrap values from 2000 replications are displayed on the appropriate branches. The *Ochlerotatus nigromaculis* (formerly *Aedes nigromaculis*) formed a clade with *Aedes Met*, as expected.

Table 1
Gene-specific primers used in 5' and 3' RACE-PCR in three mosquito species

Name of mosquito species	Name of primers	Primers	5' to 3' sequences
<i>Cx. pipiens</i>	3'RACE	Sense primer	CCTGTACGGGCAGAGTTTGT
	5'RACE	Antisense primer for RT	GTAGCTTTTCGTCAAT
	5'Cx-1	Antisense primer for primary PCR	GGACCATCCCGTACAACTCT
<i>Ae. aegypti</i>	5'Cx-2	Antisense primer for nested PCR	GGTCTTGCCAAACACGTAGTC
	3'RACE	Sense primer	GTCAGACCGATCTCTACGGCCA
	5'RACE	Antisense primer for RT	GACTAGCTCGTATGCT
<i>An. gambiae</i>	5'Ae-1	Antisense primer for primary PCR	GGGTGAATCAGGTTGAACAGA
	5'Ae-2	Antisense primer for nested PCR	CTTGCGATCAATCTCGTCTTC
	3'RACE	Sense primer	TCGATCAGTCTTTGCTGACG
	5'RACE	Antisense primer for RT	GCTACTATCACCATC
	5'An-1	Antisense primer for primary PCR	CGGGAAGCGCGGAACCGAG
	5'An-2	Antisense primer for nested PCR	GATTTGCCAAACACATAGTCCA
	AP ^a	Antisense primer	GGCCACGCGTCGACTAGTACTTTT
	AAP ^a	Sense primer	TTTTTTTTTTTTT
	AUAP ^a	Antisense primer	GGCCACGCGTCGACTAGTACGGGI
			IGGGIIGGGIIG
			GGCCACGCGTCGACTAGTAC

^a: The anchor primers are provided by the RACE Kit (Invitrogen). AP is for adapter primer, AAP for abridged anchor primer and AUAP for abridged universal amplification primers.

Table 2Amino acids identities among *DmMet* and *Dmgce* and *Met* homologs from three mosquito species

	DmMet	Dmgce	CpMet	AaMet	AgMet
<i>DmMet</i>	***	50	44	43	45
<i>Dmgce</i>	83	***	57	48	50
<i>CpMet</i>	81	95	***	64	51
<i>AaMet</i>	70	84	84	***	52
<i>AgMet</i>	81	97	95	84	***

Amino acid identities in the entire ORF and in the bHLH conserved domain are shown above and below the diagonal, respectively, for each species pair. *DmMet*: (GenBank nucleotide accession no. AF034859); *Dmgce*: (AE003498); *CpMet*: (AY895165); *AaMet*: (AY902310); *AgMet*: (DQ303468).