eNeuro

Research Article: Methods/New Tools | Novel Tools and Methods

Immediate-Early Promoter-Driven Transgenic Reporter System for Neuroethological Researches in a Hemimetabolous Insect

Takayuki Watanabe[a], Atsushi Ugajin[b] and Hitoshi Aonuma[c,d]

[a]*Department of Biological Sciences, Faculty of Science, Hokkaido University, Kita 10 Jo Nishi 6 Chome, Kita-Ku, Sapporo 060-0810, Japan*

[b]*Laboratory of Evolution of Insect Feeding Habit, JT Biohistory Research Hall, Takatsuki 569-1125, Japan*

[c]*Research Center of Mathematics for Social Creativity, Research Institute for Electronic Science, Hokkaido University, Sapporo 060-0811, Japan*

[d]*CREST, Japan Science and Technology Agency, Kawaguchi 332-0012, Japan*

DOI: 10.1523/ENEURO.0061-18.2018

Received: 6 February 2018

Revised: 11 July 2018

Accepted: 20 July 2018

Published: 7 August 2018

Author contributions: T.W. designed the study and wrote the manuscript with the help of A.U.. T.W. performed all experiments and analyzed the data. All the authors reviewed the manuscript.

Funding: http://doi.org/10.13039/501100001691Japan Society for the Promotion of Science (JSPS)
24•3065
16K20875

Funding: http://doi.org/10.13039/100007591Akiyama Life Science Foundation

Conflict of Interest: The authors declare that they have no conflicts of interest.

This work was supported, in part, by the Grant-in-Aid for Japanese Society for the Promotion of Science (JSPS) Research Fellow (24•3065) to TW; the JSPS KAKENHI (16K20875) to TW; and Akiyama Life Science Foundation research grant 2014 to TW. The white-eye mutant of ***G. bimaculatus*** (***gwhite*** mutant) was kindly gifted by Prof. Sumihare Noji in the University of Tokushima. The pXL-BacII, pBSII-IFP2-orf, and pBSII-ITR1.1k-EYFP plasmids were kindly gifted by Prof. Malcolm J. Fraser, Jr. in the University of Notre Dame.

Correspondence should be addressed to Author for correspondence Takayuki Watanabe, wtaka226@frontier.hokudai.ac.jp

Cite as: eNeuro 2018; 10.1523/ENEURO.0061-18.2018

Immediate-early promoter-driven transgenic reporter system for neuroethological researches in a hemimetabolous insect

Takayuki Watanabe[a]*, Atsushi Ugajin[b] and Hitoshi Aonuma[c,d]

[a]Department of Biological Sciences, Faculty of Science, Hokkaido University, Kita 10 Jo Nishi 6 Chome, Kita-ku, Sapporo 060-0810, Japan

[b]Laboratory of Evolution of Insect Feeding Habit, JT Biohistory Research Hall, Takatsuki 569-1125, Japan

[c]Research Center of Mathematics for Social Creativity, Research Institute for Electronic Science, Hokkaido University, Sapporo 060-0811, Japan

[d]CREST, Japan Science and Technology Agency, Kawaguchi 332-0012, Japan

*Author for correspondence (wtaka226@frontier.hokudai.ac.jp)

Running title: IEG reporter system in the cricket

Key words: immediate-early gene | activity mapping | transgenesis | *Gryllus bimaculatus*

ABSTRACT

Genes expressed in response to increased neuronal activity are widely used as activity markers in recent behavioral neuroscience. In the present study, we established transgenic reporter system for whole-brain activity mapping in the two-spotted cricket *Gryllus bimaculatus*, a hemimetabolous insect used in neuroethology and behavioral ecology. In the cricket brain, a homolog of *early growth response-1* (*Gryllus egr-B*) was rapidly induced as an immediate-early gene (IEG) in response to neuronal hyperexcitability. The upstream genomic fragment of *Gryllus egr-B* contains potential binding sites for transcription factors regulated by various intracellular signaling pathways, as well as core promoter elements conserved across insect/crustacean *egr-B* homologs. Using the upstream genomic fragment of *Gryllus egr-B,* we established an IEG promoter-driven transgenic reporter system in the cricket. In the brain of transgenic crickets, the reporter gene (a nuclear-targeted destabilized EYFP) was induced in response to neuronal hyperexcitability. Inducible expression of reporter protein was detected in almost all neurons after neuronal hyperexcitability. Using our novel reporter system, we successfully detected neuronal activation evoked by feeding in the cricket brain. Our IEG promoter-driven activity reporting system allows us to visualize behaviorally relevant neural circuits at cellular resolution in the cricket brain.

SIGNIFICANCE STATEMENT

Insects are the largest and most diverse group of organisms, and show a wide variety of behaviors. Despite the importance of a comparative approach, recent insect neuroethology mostly relies on the fruit fly *Drosophila melanogaster*, a versatile model insect with particularly powerful genetic tools, and a relatively small number of researchers use non-*Drosophila* insects. In the present study, a novel genetic reporting system for whole-brain activity mapping was established in the two-spotted cricket *Gryllus bimaculatus* by using the newly identified and evolutionary conserved gene regulatory region of an immediate-early gene, *early growth response*. This reporting system allowed us to analyze the behaviorally-evoked neural activity patterns at cellular resolution in the cricket brain

INTRODUCTION

Insects are the largest and most diverse group of organisms on Earth. They originated ~500 million years ago, at almost the same time as the first terrestrial plants (Misof et al., 2014). Now, they represent ~80% of all living organisms and are found in almost all terrestrial and freshwater environments. As a result of their evolution and diversification, insects show a wide variety of behaviors, including locomotion, feeding, molting, diapause, and social behaviors (e.g., agonistic interaction, courtship and mating behavior). Therefore, insects have been widely used for neuroethological studies over the past decades.

Recent insect neuroethology largely relies on the fruit fly *Drosophila melanogaster*, a versatile model organism with particularly powerful neurogenetic tools (e.g., an abundant collection of transgenic lines (Jenett et al., 2012; Manning et al., 2012), recombination-based genetic systems for targeted gene expression in specific neural circuits (Lee and Luo, 2001; Lee, 2014; Griffin et al, 2014), optogenetic/thermogenic tools and genetically encoded calcium indicators (Owald et al., 2015; Riemensperger et al., 2016)), which allow us to manipulate/monitor the activity of specific neural circuits during behavior. In contrast, a relatively small number of researchers use non-*Drosophila* insects despite the importance of a comparative approach to understanding the general principles and evolution of insect behavior and its underlying neural mechanisms. Without sophisticated genetic tools, researchers have to depend on limited experimental techniques such as electrophysiology, neuroimaging, and behavioral pharmacology. Although these classical techniques are well established, they have some inevitable weaknesses and limitations: For example, neurophysiological techniques are vulnerable to mechanical disturbances caused by animal movements, and pharmacology provides us limited information on the involvement of receptors/signaling cascades in a certain behavior. To facilitate neuroethological studies in non-*Drosophila* insects, a technical breakthrough is necessary to compensate/overcome these problems. In recent behavioral neuroscience, histological detection of

activity-regulated genes has been widely employed for retroactive labeling of behaviorally relevant neural circuits (Clayton, 2000; Guzowski et al., 2005; Maruska et al., 2013; Kawashima et al., 2014). In this study, aiming to introduce such a powerful technique to non-*Drosophila* insects, we established a novel transgenic reporter system for whole-brain activity mapping in a hemimetabolous insect.

In neurons, the expression of a wide variety of genes is regulated in response to increased neuronal excitation and activation of cellular signaling pathways (Flavell and Greenberg, 2008), which are involved in various cellular functions, including transcription regulation, signaling pathways, metabolism, and synaptic function (Loebrich and Nedivi, 2009; Okuno, 2011). In terms of expression time-course and regulatory mechanisms, activity-regulated genes can be classified into two categories, immediate-early gene (IEGs) and delayed-early genes (DEGs). The neuronal IEGs are rapidly and transiently induced in response to neuronal activation and whose expression does not require *de novo* protein synthesis (Clayton, 2000). Some IEGs encode transcription factors whose protein products, in turn, regulate expression of DEGs. Therefore, transcription factor IEGs are considered as the first wave of genomic response to shape the cellular response toward various physiological/environmental stimuli. In the nervous system, transcription factor IEGs, such as *c-fos* and *egr-1/zif268/NGFI-A*, are expressed in response to a wide variety of stimuli and in multiple cell types, therefore they are widely used for activity mapping in the vertebrate system. In the early studies, researchers directly detected the activity-regulated expression of gene products of IEGs to map behaviorally-evoked neuronal activity. In the recent studies, the promoter regions of activity-regulated genes were used in combination with optogenetic tools to address the functional importance of behaviorally relevant circuits (Ramirez et al., 2014; Minatohara et al., 2016).

So far, activity-regulated genes-based activity mapping has been conducted on limited insect species, such as the honeybee, *Drosophila*, and silk moth. In the brain of the honeybee, the non-coding RNA *kakusei* and a homolog of transcription factor c-*jun, jun-related antigen* (*jra*), were used to map a

91 neural activity associated with the forging behavior and waggle dance (Kiya et al,, 2007; Kiya and
92 Kubo, 2011) and defensive behaviors (Alaux and Robinson, 2007; Ugajin et al., 2012). In the brain of
93 *Drosophila* and the silkmoth, a nuclear receptor *hr38* (*NGFI-B/nur77/NR4A1* homolog) was used for
94 activity mapping (Fujita et al., 2013). In contrast, there has been no report on the activity-regulated
95 genes or neuronal IEGs in basal hemimetabolous insects. Moreover, the activity-regulated promoter has
96 not been characterization in any insect species.

97 Crickets have been widely used for neuroethology and behavioral ecology for decades (Huber
98 et al., 1989; Horch et al., 2017) with a particular focus on their prominent social behaviors including
99 courtship and aggressive behaviors (Adamo et al., 1995; Pollack, 2000; Hedwig, 2006; Stevenson and
100 Schildberger, 2013). Moreover, modern genetic techniques (e.g., transposon-mediated transgenesis,
101 genome editing techniques) have been introduced to the two-spotted cricket *Gryllus bimaculatus* over
102 the last decade (Nakamura et al., 2010; Watanabe et al., 2012; Horch et al., 2017). For this reason, we
103 chose *G. bimaculatus* as a model system to establish a transgenic activity reporting system using an
104 activity-regulated promoter. In the present study, we first identified neuronal IEGs in the cricket to
105 isolate and characterize the activity-regulated promoter in basal insects. We then established a
106 transgenic reporter line for a whole-brain, single cell-resolution activity mapping, which is required for
107 further functional studies of behaviorally relevant neural circuits.

MATERIALS AND METHODS

1. Animals

A wild-type strain of two-spotted crickets *Gryllus bimaculatus* DeGeer was inbred for decades in our laboratory (Hokudai WT strain). The Hokudai WT strain was used for identification and initial expression analysis of the candidate IEGs, and 5' RACE experiment. A congenic white-eye strain (Hokudai *gwhite* strain) was established by five-time-backcrossing of *gwhite* mutant (kindly provided by Prof. Noji in Tokushima Univ., Japan) to the Hokudai WT strain. The Hokudai *gwhite* strain was used for transgenesis. Crickets were reared in a group on a 14-h:10-h light/dark cycle at 28 °C. They were fed a diet of insect food pellet (Oriental Yeast Co., Tokyo, Japan) and water *ad libitum*. Specimens of the European honeybee *Apis mellifera L.* were collected from colonies maintained at Tamagawa University. Specimens of the house cricket *Acheta domesticus* were purchased at a local pet store. Specimens of the katydid *Gampsocleis buergeri* and migratory locust *Locusta migratoria* were collected in Tokyo (Japan).

2. Identification of candidate neuronal IEGs

We amplified cDNAs encoding a full-length protein coding sequence of three candidate IEGs and a partial cDNA of *Gryllus hr38* from cDNA libraries derived from the brains of adult crickets prepared in Watanabe et al. (2011).

Partial cDNA cloning of *Gryllus fra*: Partial cDNA of *Gryllus fra* was amplified using gene-specific primers (GSPs) designed on the basis of *G. bimaculatus* expressed sequence tag (EST) clones corresponding to *fra* (GenBank IDs; AK282142 and AK254978).

Partial cDNA cloning of *Gryllus jra*: Partial cDNA of *Gryllus jra* was amplified using GSPs designed on the basis of a *G. bimaculatus* EST clone corresponding to *Gryllus jra* (GenBank ID; AK263019).

Partial cDNA cloning of *Gryllus egr*: Partial cDNA of *Gryllus egr* was amplified using degenerate primers. Degenerate primers were designed on the basis of conserved amino acid sequences (GVQLAEY, TSKGHEI, FQCRICMR and HAKVHLK) among insect Egr homologs. Sequences of the degenerate primers are listed in Table 1.

Partial cDNA cloning of *Gryllus hr38*: Partial cDNA of *Gryllus hr38* was amplified using degenerate primers. Degenerate primers were designed on the basis of conserved amino acid sequences (NRCQFC and RDDQELL) among insect *hr38* homologs. Sequences of the degenerate primers are listed in Table 1.

Full-length cDNA cloning of *Gryllus fra*, *jra,* and *egr*: First, nucleotide sequences flanking the translational initiation and termination sites of each gene were obtained by 5' and 3' RACEs. RACE PCR was carried out using GSPs designed on the basis of the nucleotide sequences of partial cDNAs of each gene. Complementary DNAs containing full-length ORF of the genes were amplified using GSPs designed at the 5′ and 3′ untranslated region of the genes. Primers used to amplify cDNAs containing full-length ORF are listed in Table 1. All PCRs were performed using the Q5 High-Fidelity DNA polymerase (New England Biolabs, Tokyo, Japan). Amplified cDNA fragments were cloned into the pGEM-T Easy vector (Promega, WI, USA) and their nucleotide sequences were determined.

3. Sequence comparison and structural analysis of the proteins

The deduced amino acid sequences of *Gryllus fra*, *jra, egr*, and *hr38* were aligned with those of the corresponding parts of homologous genes of other species using the MAFFT (Katoh and Standley, 2013; RRID: SCR_011811) or MUSCLE algorithms (Edgar, 2004; RRID: SCR_011812) and refined by manual inspection on the Geneious program (ver. 9) created by Biomatters (available from http://www.geneious.com/; RRID: SCR_010519). The leucine zipper domain of *Gryllus* Fra and Jra

proteins, and the zinc finger motifs of *Gryllus* Egr protein were predicted by the SMART program (Schultz et al., 1998; Letunic et al., 2012; available from http://smart.embl-heidelberg.de; RRID: SCR_005026).

4. Pharmacology

Induction of neuronal hyperexcitability: Adult male crickets one week after the imaginal molt were used for pharmacological experiment. Crickets were individually isolated in a 100 ml beaker (ø4.5 cm) for three days without food and water before pharmacological treatment. To induce neuronal hyperexcitability, 5 mM picrotoxin (PTX; Sigma-Aldrich; CAS: 124-87-8) resolved in the cricket physiological saline (140 mM NaCl, 10 mM KCl, 1.6 mM $CaCl_2$, 2 mM $MgCl_2$, 44 mM glucose, 2 mM TES, pH 7.2) containing 5% dimethyl sulfoxide (DMSO) was injected. The cricket physiological saline containing 5% DMSO was injected as a vehicle control. To block *de novo* protein synthesis, 20 mM cycloheximide (Sigma-Aldrich; CAS: 66-81-9) dissolved in a cricket physiological saline was injected 1 hr before PTX injection. 3 µl of each solution was injected into the head capsule using a 27-gauge needle attached to a 10 µl microsyringe (Hamilton 701 LT Syringe; Sigma-Aldrich, Tokyo, Japan).

Activation of intracellular signaling pathways: Adult male crickets one week after the imaginal molt were used for the pharmacological experiment. Crickets were individually isolated in a 100 ml beaker (ø4.5 cm) for three days without food and water before pharmacological treatment. To activate specific intracellular signaling pathways, 200 µM 12-O-tetradecanoylphorbol-13-acetate (TPA; Cayman Chemical; CAS: 16561-29-8), 1 mg/ml anisomycin (Cayman Chemical; CAS: 22862-76-6), 200 µM forskolin (Millipore; CAS: 66575-29-9), 10 mM S-nitroso-N-acetyl-DL-penicillamine (SNAP; Sigma-Aldrich; CAS: 67776-06-1), 200 µM A23187 (Sigma-Aldrich; CAS: 52665-69-7), which dissolved in the cricket physiological saline containing 5% DMSO were injected. 3 µl of each solution was injected

177 into the head capsule using a 27-gauge needle attached to a 10 µl microsyringe (Hamilton 701 LT
178 Syringe; Sigma-Aldrich, Tokyo, Japan).

179

180 **5. Quantification of transcripts of the candidate IEGs**

181 **RNA extraction and reverse transcription for the initial expression analysis:** Crickets were
182 anesthetized on ice, and the brain (the supraesophageal and suboesophageal ganglia without the optic
183 lobes) was dissected in ice-cold saline, immediately chilled in liquid nitrogen, and stored at -80 ºC until
184 use. Total RNA was extracted from a single cricket brain using the PureLink RNA Mini kit (Life
185 Technologies, CA, USA) according to the manufacturer's instruction. RNA samples were reverse
186 transcribed using the High-Capacity cDNA Reverse Transcription Kit (Life Technologies, CA, USA).
187 0.5 µg of total RNA was reverse transcribed in a 20 µl reaction according to the manufacturer's
188 instruction.

189 **RNA extraction and reverse transcription for the other RT-qPCR expression analyses:** Total
190 RNA was extracted from a single cricket brain using the TRIzol reagent (Life Technologies). Genomic
191 DNA contamination was digested with DNase I (1 U in 10 µl reaction; TaKaRa, Shiga, Japan) at 37 ºC
192 for 1 hr, followed by incubation at 75 ºC for 10 min to inactivate the enzyme. RNA samples were
193 reverse transcribed using the High-Capacity cDNA Reverse Transcription Kit (Life Technologies). 1.5
194 µl of total RNA was reverse transcribed in a 10 µl reaction according to the manufacturer's instruction.

195 **Quantitive PCR:** Quantitive PCR analysis was performed using the KAPA SYBR FAST qPCR Kit
196 (Kapa Biosystems, MA, USA) and the Eco Real-Time PCR System (Illumina, CA, USA). 0.5 µl of
197 each reverse transcription product was added to a 10 µl qPCR reaction. Quantitative PCR reaction was
198 performed at 50 ºC for 2 min and 95 °C for 5 min followed by 40 cycles of 95 °C for 10 sec and 60 °C
199 for 30 sec each. Sequences of the primers used for qPCR are listed in Table 1.

Gene expression levels were measured by the standard curve method using the EcoStudy Software (ver. 5.0) (Illumina). Ten-fold serial dilutions of plasmids (pGEM-T Easy backbone) containing cDNA fragment(s) of each target gene(s) were used to plot standard curves. The template plasmid for *Gryllus fra* contained the isoform-specific sequence of *fra-A* (base 1 to 371 of LC215243) and the full-length cDNA of *Gryllus fra-B* (LC215244), which were tandemly inserted into the pGEM-T Easy. The template plasmid for *Gryllus jra* contained the full-length cDNA of *Gryllus jra* (LC215245). The template plasmid for *Gryllus egr* used for the initial RT-qPCR experiment (see Figure 1) contained the partial cDNA of *Gryllus egr-B* (base 1 to 429 of LC215246). The template plasmid for *Gryllus egr-B* and *EYFPnls:PEST* contained the partial cDNA of *Gryllus egr-B* (base 1 to 929 of LC215246) and the full-length coding sequence of *EYFP*, which were tandemly inserted into the pGEM-T Easy. The template plasmid for *Gryllus hr38* contained the partial cDNA sequence of *hr38* (base 1 to 374 of LC341255). The template plasmid for the pre-mRNAs of *Gryllus egr-B* and *hr38* contained the genomic fragment corresponding to the intron/exon boundary of *Gryllus egr* and *hr38* genes, respectively (base 1 to 437 of LC341288; base 1 to 260 of LC341256).

Expression levels of each target gene was normalized with that of Gryllus *ef1α* gene, which was selected from three housekeeping genes (*Gryllus β-actin*, *ef1α*, and *rpl32*) according to expression stability in the brain (see Figure 1-1). The template plasmids for *Gryllus β-actin*, *ef1α*, and *rpl32* contained the full-length cDNA of the genes (AB626808.1, AB583232.1, and AB626807.1), respectively.

6. Determination of the transcription start sites (TSSs) of *Gryllus* and *Apis egr-B* homologs

We performed the 5' and 3' rapid amplification of cDNA ends (RACE) using the FirstChoice RLM-RACE kit (Ambion, Austin, TX, USA) to clone the 5' ends of *Gryllus* and *Apis egr-B* homologs. RNA processing reactions were carried out according to the manufacturer's instruction. Reverse transcription

was carried out using the Superscript III reverse transcriptase (Life Technologies) according to the manufacturer's instruction. 5' RACE PCRs were carried out using the Q5 High-Fidelity DNA polymerase (New England Biolabs). Complementary DNAs were inserted into the pGEM-T Easy vector (Promega) by standard TA cloning procedure using the T4 DNA ligase or the Gibson assembly technique using the Gibson Assembly Master Mix (New England Biolabs). Complementary DNA fragments for the Gibson assembly cloning were amplified with primers with 5' extensions (5'-GCC GCG GGA ATT CGA TT-3' was attached to the 5' end of the 5' RACE inner primer; 5'-CCG CGA ATT CAC TAG TGA TT-3' was attached to the 5' end of the gene-specific reverse primer). Over twenty cDNA clones were sequenced to determine the TSSs. All PCRs were carried out using the Q5 High-Fidelity DNA Polymerase (New England Biolabs). PCR products were cloned into the pGEM-T Easy vector (Promega) and their sequences were determined.

7. Isolation of the upstream regions of orthopteran *egr-B* homologs

To isolate the putative core promoter region, genomic DNA of three orthopteran insects were extracted using the Wizard Genomic DNA Purification Kit (Promega). Based on the high-level sequence similarity in the core promoter region, we designed primers to amplify the putative core promoter region of orthopteran *egr-B* homologs. PCRs were performed using a pair of the forward and reverse primers listed in Table 1. Then, the upstream regions of orthopteran *egr-B* homologs were obtained by using inverse PCR. The procedure for inverse PCR was modified from Watanabe and Aonuma (2014). Genomic DNAs of three orthopteran insects were digested with *Bfa*I, *Bgl*II, *Dpn*II, *Eae*I, *Hha*I, *Hind*III, *Mse*I, *Msp*I, *Nla*III, *Taq*αI, *Xap*I or *Xce*I. 10 ng of restriction fragments were circularized using the T4 DNA ligase (Thermo Fisher Scientific, MA, USA) in 10 µl reaction. Primers were designed on the basis of the nucleotide sequence of the putative core promoter of *egr-B* homologs. Inverse PCRs were repeatedly performed to obtain >1.5 kbp genomic DNA fragments flanking the first exon of *egr-B*

248 homologs. All PCRs were carried out using the Q5 High-Fidelity DNA Polymerase (New England
249 Biolabs). PCR products were cloned into the pGEM-T Easy vector (Promega) and their sequences were
250 determined.

251

252 **8. Database search and sequence comparison of stimulus-regulated transcription factors among**
253 **vertebrates and insects**

254 To understand the structural features of the upstream sequence of insect/crustacean *egr-B* homologs,
255 we searched potential binding sites for stimulus-regulated transcription factors (TFs) conserved among
256 vertebrates, insects and crustaceans. There have been few reports on the DNA-binding specificities of
257 stimulus-regulated TFs in insects, and crustaceans so far, we examined sequence conservation in the
258 DNA-binding domains of stimulus-regulated TFs between mouse and insects, in order to assess
259 whether each candidate TF in insects might show have DNA-binding properties to vertebrate
260 homolog(s). We analyzed components of the stimulus-regulated transcription factor complexes which
261 are known to involved in regulating stimulus-regulated gene expression (Dolmetsch et al., 1997; Hu et
262 al., 2001; Roy et al., 2007; Tullai et al., 2007) as follows: AP-1, CREB, C/EBP, Egr, NFAT, MEF2,
263 and SRF, which include Fos-family proteins, Jun-family proteins, ATF/CREB family proteins, Maf-
264 family proteins, C/EBPs, Egr-family proteins, NFATs, MEF2, and SRF. We first retrieved mouse
265 homologs of stimulus-regulated TFs from the GenBank protein sequence database. Then, BLASTp
266 searches were conducted in search for homologs of stimulus-regulated TFs in *Drosophila*
267 *melanogaster, Tribolium castaneum, Apis mellifera, Acyrthosiphon pisum,* and *Zootermopsis*
268 *nevadensis*. We also performed tBLASTn searches on the *Gryllus firmus* transcriptome shotgun
269 assembly (TSA) database to find cDNA sequence of the cricket homologs of the target genes. The
270 amino acid sequences of the DNA binding domains of stimulus-regulated TFs of mouse and insects

were aligned using the MAFFT or MUSCLE algorithms on the Geneious program (ver. 9). GenBank IDs of stimulus-regulated TFs are listed in Table 3-1.

9. Structural analysis of the potential gene regulatory region of *egr* genes

Potential transcriptional regulatory sequences in the upstream sequences of insect/crustacean *egr-B* homologs were predicted using the LASAGNA-Search 2.0 program (Lee and Huang, 2014; available from http://biogrid-lasagna.engr.uconn.edu/lasagna_search/; RRID: SCR_010883). Potential binding sites for stimulus-regulated transcription factors common to vertebrates and insects were searched. Phylogenetic footprinting analysis of the upstream regions of the polyneopteran *egr-B* homologs was conducted using the mVISTA program (Frazer et al., 2004; available from http://genome.lbl.gov/vista/mvista/submit.shtml) with default setting. GenBank IDs of genomic sequences are listed in Table 2.

10. Construction of *piggyBac* donor plasmid

A *piggyBac* donor plasmid pXL-BacII[*3xP3-mCherry*]{*Gbegr-EYFPnls:PEST-2xARE*} was constructed to examine promoter activity of *Gryllus egr-B in vivo*. This plasmid contains the expression cassettes for an IEG reporter (*Gbegr-EYFPnls:PEST-2xARE*) and a visible selection marker (*3xP3-mCherry*). The IEG reporter cassette consists of (1) a 2.2-kb upstream sequence of *Gryllus* e*gr-B* coding sequence; (2) a coding sequence of nuclear-targeted EYFP C-terminally fused to a PEST domain of mouse ornithine decarboxylase (EYFPnls:PEST); (3) two repeats of synthetic AU-rich elements (2xARE); and (4) a SV40 3' untranslated region (UTR). The expression cassettes of the IEG reporter and *3xP3-mCherry* were separated by the *gypsy* insulator sequence. The donor plasmid was constructed through the following procedures.

294 ***Gbegr-EYFPnls:PEST-2xARE* expression cassette**: A ~2.2-kb genomic fragment upstream to *Gryllus*
295 e*gr-B* was amplified by PCR from *Gryllus bimaculatus* genomic DNA purified using the Wizard
296 Genomic DNA Purification Kit (Promega). A nucleotide sequence encoding the nuclear localization
297 signal of the SV40 Large T-antigen (PKKKRKV) was added to the 3' end of the coding sequence of
298 EYFP derived from the pBSII-ITR1.1k-EYFP plasmid by PCR. The coding sequence of the PEST
299 domain of mouse ornithine decarboxylase (residues 416-461 of NP_038642.2) was synthesized by
300 Eurofins Genomics (Tokyo, Japan). Two repeats of synthetic AU-rich elements (5'-TTT ATT TAT
301 TTA TTT ATT TA-3') were added to the 5' end of the SV40 3' UTR derived from the pBSII-ITR1.1k-
302 EYFP plasmid by PCR. The DNA fragments of the promoter, protein coding sequence, and 3' UTR
303 were assembled in order by restriction enzyme digestion and ligation, or by using the Gibson assembly
304 technique.

305 ***3xP3-mCherry* expression cassette**: The *mCherry* coding sequence derived from the pTRE3G-
306 mCherry vector (TaKaRa) was inserted into the *Nco*I/*Xba*I-digested pBSII-ITR1.1k-EYFP plasmid.

307 **Plasmid construction:** First, a *gypsy* insulator sequence derived from the pGreen-Pelican plasmid
308 (DGRC stock number: 1015) was inserted between the *Hind*III and *Sph*I sites of the pXL-BacII
309 plasmid. Then, the *Gbegr-EYFPnls:PEST-2xARE* expression cassette was inserted between the *Sph*I
310 and *Xba*I sites of the vector to construct pXL-BacII{*Gbegr-EYFPnls:PEST-2xARE*}. Finally, the *3xP3-*
311 *mCherry* expression cassette (*3xP3-mCherry-SV40 3'UTR*) was inserted into the *Hind*III-digested pXL-
312 BacII{*Gbegr-EYFPnls:PEST-2xARE*} using the Gibson assembly technique.

313 All PCRs were carried out using the Q5 High-Fidelity DNA Polymerase (New England
314 Biolabs). Gibson assembly was performed using the Gibson Assembly Master Mix (New England
315 Biolabs). After construction, the plasmid was amplified in *Escherichia coli* strain DH5α and purified
316 using the PureLink HiPure Plasmid Kit (Life technologies).

317

11. *piggyBac* transposon-mediated transgenesis

***In vitro* synthesis of *piggyBac* transposase mRNA:** To construct the expression plasmid for *piggyBac* transposase mRNA (pTD1-*piggyBac*), the coding sequence of *piggyBac* transposase derived from the pBSII-IFP2-orf plasmid was inserted into the pTD1 expression vector (Shimazu, Kyoto, Japan). The expression cassette of the *piggyBac* transposase in the pTD1-*piggyBac* was amplified by PCR with the Q5 High-Fidelity DNA Polymerase (New England Biolabs, Tokyo, Japan) with following primers (forward 5'-GCA GAT TGT ACT GAG AGT G-3' and reverse 5'-CAG GAA ACA GCT ATG AC-3'). The PCR fragment was subsequently used as templates for *in vitro* transcription using the T7 mScript Standard mRNA Production System (CELLSCRIPT, Madison, WI, USA). Then, the 5' cap and poly(A) tail were added to the transcribed RNA using the kit. The *piggyBac* mRNA was purified by phenol-chroloform extraction and concentrated by ammonium acetate precipitation.

Microinjection: The *piggyBac* donor plasmid (1 µg/µl) and *piggyBac* mRNA (1 µg/µl) were dissolved in nuclease-free water (Qiagen, Tokyo, Japan) to make an injection solution for germline transformation. The injection solution was injected into fertilized eggs of the Hokudai g*white* strain. Eggs were laid in a wet cotton or paper towel, and collected 30 min within egg laying. They were rinsed with 70% ethanol, aligned on the hand-made injection chamber, shortly air dried and covered with mineral oil (Nacalai Tesque, Kyoto, Japan). Micropipettes were pulled on a micropipette puller (PA-81; Narishige, Tokyo, Japan) using thin-walled filament glass capillary (GD-1.2; Narishige). The capillaries were backfilled with the injection solution using the GELoader tip (Eppendorf, Tokyo, Japan) and connected to the IM-6 microinjector (Narishige) filled with mineral oil. Injection solution was injected into the dorsal posterior part of the eggs. Eggs were transferred onto the wet cotton in the glass petri dish immediately after injection. Petri dishes were incubated at 28 ºC in the moist chamber until hatching.

Genetics: Embryos injected with the *piggyBac* donor plasmid and *piggyBac* mRNA were raised to adult and backcrossed to Hokudai g*white* strain. The F_1 embryos exhibiting mCherry fluorescence in the compound eyes were selected. Each transgenic cricket was repeatedly backcrossed to the Hokudai g*white* strain to generate heterozygous lines.

12. Genotyping

To establish IEG promoter-driven reporter lines, heterozygous transgenic lines were selected according to the inducible expression of EYFPnls:PEST protein after PTX treatment. 6 hr after PTX injection, the brains were dissected for EYFP immunohistochemistry according to the experimental procedure described below. Then, the genomic flanking sequence of *piggyBac* insertion was determined by inverse PCR. To establish homozygous lines, each heterozygous line was inbred, and homozygous transgenic crickets were selected by PCR-based genotyping.

Inverse PCR: First, the genomic regions flanking the *piggyBac* insertions were amplified by inverse PCR. Genomic DNA of the heterozygous transgenic crickets was extracted using the Wizard Genomic DNA Purification Kit (Promega). Genomic DNA was digested with *Dpn*II, *Hind*III, or *Msp*I. 10 ng of restriction fragments were circularized using the T4 DNA ligase (Promega) in 10 µl reaction. PCRs were performed to amplify the 5' and 3' flanking regions of the *piggyBac* insertion using the Q5 High-Fidelity DNA Polymerase (New England Biolabs). Following primers were used for amplification of the 5' flanking region of the insertion (primers for 1st PCR: 3'-GCT CCA AGC GGC GAC TGA GAT GTC C-3' and 5'-GCT TGT CAA TGC GGT AAG TGT CAC TG-3'; primers for nested PCR: 5'-GAC GGA TTC GCG CTA TTT AGA AAG AGA G-3' and 5'-CGG TAA GTG TCA CTG ATT TTG AAC TAT AAC G-3'). The 3' flanking region of the insertion was amplified by inverse PCR using primers designed on the basis of the nucleotide sequence of the 5' flanking region of the insertion. PCR products were cloned into the pGEM-T Easy vector (Promega) and their sequences were determined.

PCR-based genotyping: Heterozygous transgenic lines were inbred to generate homozygous transgenic animals. Each adult transgenic cricket in the inbred colonies was isolated after the imaginal molt, and homozygous animals were selected by PCR-based genotyping. Genomic DNA was isolated from the wings of newly emerged transgenic crickets using the modified HotSHOT method (Meeker et al., 2007). The tip of a hind wing (approx. 2 mg) was dissected from crickets, and homogenized in liquid nitrogen. Samples were incubated at 95 ºC for 30-60 min in 50 µl of 50 mM NaOH. 5 µl of 1 M Tris-HCl (pH 7.5) was added to each tube for neutralization. After brief centrifugation, supernatants were diluted 16 times with TE buffer (pH 8.0). PCR was carried out using the Q5 High-Fidelity DNA Polymerase (New England Biolabs) in a 10 µl reaction. 0.5 µl of genomic DNA solutions were added to the reactions. For genotyping, forward and reverse primers were designed at the 5' and 3' genomic regions flanking the *piggyBac* insertion (forward primer [line19_fw primer], 5'-CAC ATT CAC ACA TAT CCG CAG TTC-3'; reverse primer [line19_rv primer], 5'-CGT TCT TCA ATT TCA TTT TTC TCC TC-3'), respectively (see Figure 5C and D). PCR products were run through 1.5% agarose gel in TBE buffer and visualized by ethidium bromide. We selected homozygous transgenic animals as follows: a 358-bp genomic fragment was amplified if the cricket has no transgene or one copy of transgene (heterozygous crickets), whereas no PCR product was amplified from homozygous crickets.

13. Generation of anti-*Gryllus* Tdc2 polyclonal antibody

Complementary DNA cloning for *tyrosine decarboxylase* (*tdc*) genes: The full-length cDNA of two *tdc* genes (*Gryllus tdc1* and *tdc2*) were obtained as follows: A partial cDNA of *Gryllus tdc1* was amplified using GSPs designed on the basis of a *G. bimaculatus* EST clone corresponding to *Gryllus tdc1* (GenBank ID; AK258423). A partial cDNA of *Gryllus tdc2* was amplified using degenerate primers designed on the basis of conserved amino acid sequences (IEYADSFNT and FFVRMVSDP) among the insect Tdc2 proteins. Then, 5′ and 3′ RACEs were performed as described above to

determine translational initiation and termination sites of the genes. Finally, full-length cDNAs of the genes were amplified, and their nucleotide sequences were determined. Sequences of the primers are listed in Table 1.

Recombinant protein expression and antibody production: A cDNA fragment encoding amino acids 470-683 of *Gryllus* Tdc2 protein (GenBank ID: BAO52000.1) was inserted between the *BamH*I and *EcoR*I sites of the pGEX-6P-1 vector (GE Healthcare Japan, Tokyo, Japan). The antigen peptide was produced as a glutathione S-transferase (GST) fusion protein in the Rossetta 2(DE3)pLysS cells (Millipore, Billerica, MA, USA). GST-tagged protein was bound to the Glutathione HiCap Matrix (QIAGEN), and antigen peptide was cleaved using the PreScission protease (GE Healthcare Japan). The anti-*Gryllus* Tdc2 polyclonal antibody was generated in the guinea pig, and IgG fraction was purified from antiserum with Protein A (Frontier Institute; Ishikari-shi, Hokkaido, Japan). Antibody specificity was checked by Western blot. An intensely stained band of approximately ~200 kDa was detected in the lane of the central brain, which matches the calculated molecular mass of *Gryllus* Tdc2 protein (data not shown). Anti-*Gryllus* Tdc2 polyclonal antibody (1.4 mg IgG/ml) was stored in 50% glycerol at -20 ºC.

14. Whole-mount immunohistochemistry

Whole-mount fluorescent immunohistochemistry was performed according to Gonzalez-Bellido and Wardill (2012) with modifications. Crickets were anesthetized on the ice, and the brain (the supraesophageal and suboesophageal ganglia without the optic lobes) were dissected in ice-cold saline. The brains were fixed with 4% paraformaldehyde in phosphate-buffered saline (PBS; P-5493, Sigma-Aldrich, Tokyo, Japan) at 4 ºC overnight. When we conduct anti-octopamine immunohistochemistry, brains were fixed with 1% glutaraldehyde in 0.1 M sodium cacodylate buffer (pH 7.4; Electron Microscopy Sciences, PA, USA) with 1% sodium metabisulfite (SMB) for 2 hr on ice, and with 4%

PFA in PBS at 4 ºC overnight. The samples were treated with 0.5% sodium borohydride in Tris-buffered saline (TBS; 0.05 M Tris-HCl, 77 mM NaCl, pH 7.4) containing 0.45% SMB for 1 hr.

After three 30-min rinses with PBS, the brains were incubated with 500 µg/ml collagenase/dispase (Roche Applied Science, Tokyo, Japan), and 300 µg/ml hyaluronidase (Sigma-Aldrich) in PBS at 37 ºC for 1 hr. After three 15-min rinses with PBS containing 0.5% Triton-X 100 (PBS-Tx), the brains were blocked with 5% normal goat serum in PBS-Tx (blocking solution) at 4 ºC overnight. Then, the brains were incubated with the primary antibodies in the blocking solution at 4 ºC for 3-4 days. The reporter protein was detected by the chicken anti-GFP IgY (1:1000; Aves Labs (OR, USA), Cat# GFP-1020; RRID: AB_10000240) or the rabbit anti-GFP pAb (1:1000; Medical & Biological Laboratories (Nagoya, Japan), Cat# 598S; RRID: AB_591816). The DUM neurons were labeled with the guinea pig anti-*Gryllus* Tdc2 antibody (1:1000) and/or the mouse monoclonal anti-octopamine antibody (1:1000; Jena Bioscience (Jena, Germany), Cat# ABD-029, RRID: AB_2315000). After five-time 30-min-rinses in PBS-Tx, the brains were incubated with secondary antibodies in the blocking solution at 4 ºC for 3 days. The Cy3-conjugated goat anti-chicken IgY (1:400; Jackson ImmunoResearch (PA, USA), Cat# 103-165-155; RRID: AB_2337386) and the Cy2-conjugated goat anti-rabbit IgG (1:400; Jackson ImmunoResearch, Cat# 111-225-144; RRID: AB_2338021) were used to detect the reporter protein. The Cy3-conjugated donkey anti-guinea pig IgG antibody (1:400; Jackson ImmunoResearch, Cat# 706-165-148; RRID: AB_2340460) and the Cy2-conjugated goat anti-mouse IgG antibody (1:100; Jackson ImmunoResearch, Cat# 115-225-146; RRID: AB_2307343) were used to detect *Gryllus* Tdc2 and octopamine, respectively. After five-times 30-min-rinses in PBS-Tx, brains were dehydrated through 70-100% ethanol and cleared with methyl salicylate.

Fluorescent images were captured using the Olympus FLUOVIEW FV1000 confocal laser scanning microscope (Olympus, Tokyo, Japan) or the Zeiss LSM 7 DUO laser scanning microscope (Zeiss, Jena, Germany). Microscopy parameters were adjusted to the brains of naïve crickets to prevent

saturated pixels in the final images. For each brain, we captured a stack of optical sections using a 10x objective lens. Image processing was carried out using the Fiji software (Schindelin et al., 2012; Schneider et al., 2012; RRID: SCR_002285). To visualize intensity of reporter protein expression, the EYFP immunoreactivity was displayed with 'Red Hot' or 'Magenta Hot' look up tables.

15. Behavioral experiments

Adult male crickets one week after the imaginal molt were used for behavioral experiments. Each cricket was individually isolated in a 100 ml beaker (ø4.5 cm) for three days without food and water before behavioral experiment.

Gustatory stimulation: Each cricket was fed a drop of 0.5 M sucrose solution (5 µl) three times with 5 min interval. One hour after feeding, crickets were collected for RT-qPCR expression analysis. Three and six hours after feeding, crickets were collected for immunohistochemistry.

Agonistic behavior: A pair of weight-matched crickets (weight range: 0.6-0.8 g) were introduced into a round glass arena (ø12 cm) for agonistic interaction. We observed their behavior for 5 min, and winner and loser were determined according to following criteria: the winner sings an aggressive song and chases up the loser, while the loser flees from the winner. In some cases, aggressive behavior did not escalated and dominance hierarchy was not established in 5 min. After 5-min-interaction, each cricket was re-isolated in a 100 ml beaker. One hour after the start of interaction, crickets were collected for RT-qPCR expression analysis.

16. Data analysis

All statistical analyses were conducted using the GraphPad Prism version 6.0 for Mac (GraphPad Software, CA, USA; RRID: SCR_002798). Box plots indicate 25-75 percentile ranges and central

values, and '+' indicates mean. Error bars indicate 5-95 percentile ranges. Asterisks donate statistical significance (****, $p<0.0001$; ***, $p<0.001$; **, $p<0.01$; *, $p<0.05$). Details on statistical analyses are described in Table 3.

RESULTS

Gryllus egr homolog is expressed as a neuronal IEG in the cricket brain

Although the activity-regulated genes or neuronal IEGs have not been identified in any hemimetabolous insects including the cricket, recent genome-wide analyses of activity-regulated genes in vertebrates (Nedivi et al., 1993; Matsuo et al., 2000; Wada et al., 2006; Spiegel et al., 2014) and several insect species (Guan et al. 2005; Chen et al., 2016; Ugajin et al., *in press*) highlighted evolutionarily conserved transcription factor IEGs. On the basis of previous reports, we selected four transcription factor genes as our candidates for activity-regulated genes in the cricket as follows: *fos-related antigen* (*fra*; *c-fos* homolog), *jun-related antigen* (*jra*; *c-jun* homolog), *early growth response* (*egr*; *egr-1/zif268/NGFI-A* homolog), and *hormonal receptor 38* (*hr38*; *NGFI-B/nur77/NR4A1* homolog).

First, we cloned cDNAs encoding four candidate activity-regulated genes from the cricket brain cDNA library. We isolated cDNAs encoding two transcript variants of *Gryllus fra* (*fra-A* and *-B* isoforms; LC215243 and LC215244, respectively), *Gryllus jra* (LC215245), *Gryllus egr* (LC215246) and *Gryllus hr38* (LC341255) (see Figures 1-2, 1-3, 1-4, and 1-5). In *Drosophila*, the *stripe* gene (*Drosophila egr* homolog) has two alternative promoters to express alternative transcripts (*stripe-A* and *-B;* Frommer et al., 1996). The use of alternative promoters was also known in the honeybee. In the honeybee brain, variant I and II of *AmEgr* (equivalent to *egr*/*stripe*-A of other insects) are transiently expressed during metamorphosis, whereas the other isoform, variant III (equivalent to *egr*/*stripe*-B), is induced in an activity-dependent manner (Ugajin et al., 2016). Although we did not obtain cDNA clones of *egr-A* in the cricket, we concluded that the *egr* gene expressed in the cricket brain is a homolog of *egr-B* found in other insects based on the similar structural (Figure 1-4C) and expression characteristics.

Next, we examined whether the expressions of the candidate activity-regulated genes are rapidly induced in the cricket brain in an activity-dependent manner. In this study, the γ-aminobutyric acid A ($GABA_A$) receptor blocker picrotoxin (PTX) was used to cause neuronal hyperexcitability (Saffen et al., 1988; Ugajin et al., 2013). In the cricket, injection of 3 μl of 5 mM PTX caused immediate locomotor hyperactivity and, in turn, seizure-like behavior (e.g., abdominal contraction and shaking their appendages). The effect of PTX persisted at least 30 min, and most crickets recovered from the seizure within 90 min after injection. In addition, to clarify whether the candidate genes are induced as neuronal IEGs, 20 mM cycloheximide was injected 1 hr prior to PTX injection to block *de novo* protein synthesis (Matsumoto et al., 2003). Quantitative RT-PCR analysis revealed that, 30 min after PTX injection, the expression level of *Gryllus fra-B* and *egr-B* were significantly elevated in the brains of PTX-injected crickets (~1.5-fold up-regulation) than those of saline-injected crickets (Figure 1). These data demonstrated that, in the cricket brain, *Gryllus fra-B* and *egr-B* were rapidly induced as neuronal IEGs in response to neuronal hyperexcitability.

Next, we examined expression time-course of activity-regulated genes in the cricket brain after PTX injection, especially focused on *Gryllus egr-B* (Figure 2). In the brains of cycloheximide pre-treated crickets, the expression of *Gryllus egr-B* reached a maximum 60 min after PTX injection (~4.5-fold up-regulation relative to both the naïve animals and to the control (0 min after PTX injection)) and decreased to near baseline level by 120 min after PTX injection (Figure 2A). To ensure that the activity-dependent transcriptional activation of *Gryllus egr-B* is composed only of an immediate-early component, we compared the expression kinetics of *Gryllus egr-B* after PTX injection with or without the administration of cycloheximide (Figures 2A and B). The expression kinetics of *Gryllus egr-B* was not affected by blockade of *de novo* protein synthesis (two-way ANOVA, effect of pre-treatment: $F(1,98)=0.9604$, $p=0.3295$; effect of time: $F(6,98)=9.034$, $p<0.0001$; interaction: $F(6,98)=0.3634$, $p=0.9004$), indicating *Gryllus egr-B* was induced as an IEG in the cricket brain after PTX injection.

Next, we examined when PTX-induced *Gryllus egr-B* is actively transcribed by examining the expression time-course of *Gryllus egr-B* premature mRNA (pre-mRNA). In the brains of cycloheximide pre-treated crickets, the expression of *Gryllus egr-B* pre-mRNA significantly increased within 15 min after PTX injection (~7-fold up-regulation relative to both the naïve animals and to the control (0 min after PTX injection)), was sustained at a high level for >1 hr, and then decreased to near baseline level by 120 min after injection (Figure 2C). These data indicate that, after PTX injection, transcription of *Gryllus egr-B* is rapidly induced and sustained at a high level while its corresponding mRNA is elevated in the brain. Finally, we tested behaviorally-evoked expression of *Gryllus egr-B*. The expression of *Gryllus egr-B* in the brain was significantly increased 1 hr after feeding sucrose solution and agonistic interaction (Figures 2D and E). These data indicated that *Gryllus egr-B* is induced as a neuronal IEG in response to strong, sustained, widespread neuronal hyperexcitability caused by PTX, as well as behaviorally-evoked neuronal activation in the cricket brain.

Although, we could not detect significant increase of *Gryllus fra-A*, *jra*, and *hr38* 30 min after PTX injection in the initial expression analysis, these genes are slowly up-regulated in the brain under the influence of cycloheximide after PTX injection (see Figures 1-5 and 2-1). The expression of *Gryllus fra-A*, *fra-B*, and *hr38* were elevated 1 hr after feeding, whereas those of *Gryllus fra-B* and *hr38* were elevated 1hr after agonistic interaction. These data indicate that all candidate activity-regulated genes were expressed as neuronal IEGs in the brain with gene-specific expression kinetics and regulation mechanisms.

Cis-regulatory elements for stimulus-induced gene expression are enriched in the upstream genomic regions of insect/crustacean *egr-B* homologs

In order to construct an IEG promoter-driven reporter system, we determined the nucleotide sequences of the gene regulatory region of *Gryllus egr-B*. By using RNA ligase-mediated 5' rapid amplification of

534 cDNA ends (RACE) and inverse PCR, we obtained the transcription start site (TSS) and a ~1.6-kbp
535 upstream genomic fragment of *Gryllus egr-B*. Sequence comparison of the upstream regions of *egr-B*
536 homologs revealed that the core promoter region (~200 bases flanking the TSS) of *Gryllus egr-B*
537 showed a high level of sequence similarity with the corresponding part of most of *egr-B* homologs of
538 basal insects (hemimetabolous insects + Coleoptera [*Tribolium castaneum*]) and the crustacean
539 *Daphnia pulex* (Figure 3A; See Table 2). The core promoter region of basal insect/crustacean *egr-B*
540 contained *cis*-regulatory elements regulating stimulus-dependent gene expression (two CREs and
541 SRE), as well as sequence elements for core promoter function such as a CCAAT-box, initiator
542 element (Inr; consensus: YCATTC), and a downstream promoter element (DPE; consensus: AGTYY)
543 (see Figures 3A and B). Additionally, we found that two more sequence motifs, a GAGA motif and a
544 GC-rich motif, which are structurally related to the promoter elements associated with RNA
545 polymerase II (Pol II) stalling in *Drosophila* (Hendrix et al., 2008), are conserved in the upstream
546 sequences of *egr-B* homologs of most hemimetabolous insects (Figures 3A and B). The GAGA motif
547 (consensus: GRGAGGGRVGGAGAGS) is conserved in the upstream sequences of polyneopteran *egr-*
548 *B* homologs positioned at ~80 bp upstream to the TSS of *Gryllus egr-B*, but lacked in those of *egr-B*
549 homolog of the other taxa (i.e., Paraneoptera, Coleoptera, and crustacea). The GC-rich motif
550 (consensus: GCGCSSSGGCGCGC) is conserved in the upstream sequences of *egr-B* homolog of basal
551 insects positioned at ~30 bp downstream to the TSS of *Gryllus egr-B*.

552 Next, we expanded our analysis to the distal promoter region by predicting potential *cis*-
553 regulatory motifs (Figure 3C, see also Figure 3-1). Generally, the upstream regions of insect/crustacean
554 *egr-B* homologs were highly enriched with potential binding sites for transcription factors activated by
555 various intracellular signaling pathways (e.g., AP-1, C/EBP, CREB, SRF, and NF-AT), which are
556 known to regulate activity-dependent gene expression in vertebrates (Sealy et al., 1997; Hogan et al.,
557 2003; Benito et al., 2011; Nonaka et al., 2014). In addition, phylogenetic footprinting analysis of the

upstream region of *egr-B* homologs of polyneopteran insects revealed two conserved *cis*-regulatory modules (CRMs; CRM_{-800} and CRM_{-400} in Figure 4A), which consist of proximal, conserved potential binding sites for stimulus-regulated transcription factors (Figure 4B). These data indicate that the ~1.6 kbp upstream genomic fragment of *Gryllus egr-B* contains *cis*-regulatory elements for stimulus-regulated transcription factors, some of which constitute the core promoter and CRMs conserved across insect/arthropod species.

Construction of the IEG promoter-driven transgenic reporter system

An IEG promoter-driven reporter system was constructed using the upstream genomic fragment of *Gryllus egr-B* (Figure 5A and B). The IEG reporter cassette consisted of a nuclear-targeted destabilized EYFP (*EYFPnls:PEST*; Li et al, 1998) driven by a ~2.2 kb genomic DNA fragment containing the ~1.6 kbp upstream sequence and 5'UTR of *Gryllus egr-B*. Two repeats of a synthetic AU-rich element were inserted between the coding sequence of *EYFPnls:PEST* and the SV40 3' UTR to shorten mRNA half-life (Pham et al., 2008). The IEG reporter cassette along with a *3xP3-mCherry* expression cassette was integrated into the cricket genome using the *piggyBac* transposon. Successful transgenic crickets were selected according to mCherry fluorescence in the compound eyes, and reporter expression in the brain was checked by EYFP immunoreactivity after PTX treatment. In the *Drosophila* brain, Masuyama et al. (2012) reported that the inducible expression of EGFP protein under the control of activity-dependent synthetic transcription factor, CaLexA, became detectable ~4 hr after stimulation. In the present study, we collected brains of the IEG reporter line 6 hr after PTX injection to detect activity-dependent expression of reporter protein. Three out of 37 mCherry-expressing transgenic lines showed EYFP immunoreactivity in the brain after PTX injection. One transgenic line with a low background EYFP expression (line #19; Figure 5C) was selected, and a homozygous transgenic line was

established as an 'IEG reporter line' for subsequent analyses. The IEG reporter line develops normally, and we did not observe any behavioral abnormality.

We tested whether the expression of IEG reporter mimics innate IEG expression in the brains of the IEG reporter line. First, we compared baseline expression levels of the reporter gene (*EYFPnls:PEST*) and *Gryllus egr-B*. Quantitative RT-PCR expression analysis revealed that, in the brains of naïve animals, *EYFPnls:PEST* showed a higher baseline expression level than *Gryllus egr-B* (Figure 5D). Time-course analysis of PTX-induced expression of *EYFPnls:PEST* and *Gryllus egr-B* revealed significant differences in the expression kinetics of the genes after PTX injection. The expression of *EYFPnls:PEST* reached a maximum 60-90 min after PTX injection (~1.8-fold up-regulation relative to the naïve animals), and slowly decreased to near baseline by 120 min after injection (Figure 5E). On the other hand, the expression of *Gryllus egr-B* reached a maximum 90 min after PTX injection (~6.5-fold up-regulation relative to the naïve animals) and decreased to the baseline level by 120 min after PTX injection (Figure 5E). In addition, correlation analysis revealed that the expression level of the *EYFPnls:PEST* was strongly correlated with that of *Gryllus egr-B* in the brain of the IEG reporter line (Pearson's $r = 0.8269$, $p<0.0001$; Figure 5F). With these data, we concluded that the neuronal activity-driven expression of the reporter gene mimics that of *Gryllus egr-B* in the brain of the IEG reporter line.

IEG promoter-driven reporter system can be used for whole-brain activity mapping

First, we asked whether our reporter system can be used for whole-brain activity mapping. The brains of the IEG reporter line were collected 6 hr after PTX or vehicle injection, and the distribution of the reporter protein (EYFPnls:PEST) was examined by whole-mount fluorescent immunohistochemistry (Figure 6). In the brain of vehicle-injected animals (n=7), intense EYFP immunoreactivity was only detected in cell clusters located in the lateral parts of the supraesophageal

605 ganglion (white arrowheads in Figures 6A and C). The distribution of the EYFPnls:PEST protein in the
606 brain of naïve animals was similar to that in the vehicle-injected animals (data not shown). Six hours
607 after PTX injection, EYFP immunoreactivity was detected throughout the ganglia (Figures 6B and D;
608 n=6) and was restricted to the nucleus. In addition, we did not observe EYFP immunoreactivity in the
609 marginal glia (the cells located at the margins of the neuropil) or the epithelial glia after PTX injection,
610 suggesting the reporter protein is only induced in the neurons (see Movies 1 and 2). To examine the
611 expression time-course of the reporter protein, we conducted Western blot analysis using several anti-
612 GFP antibodies. However, we could not detect bands corresponding to EYFPnls:PEST before and after
613 PTX injection. We speculated that restricted localization of the reporter protein in the nuclei might
614 facilitate detection of the reporter protein in immunohistochemistry.

615 Next, we tested whether our reporting system can label behaviorally-evoked neural activation
616 under the physiological conditions. As our RT-qPCR analysis revealed that feeding of sucrose solution
617 induced expression of neuronal IEGs in the cricket brain, we examined the expression of the reporter
618 protein induced by feeding of sucrose solution in the brain of the IEG reporter line. In the insect brain,
619 the octopaminergic/tyraminergic neurons are activated by feeding of sucrose solution or presentation of
620 "reward" in the context of learning and memory studies. In the honeybee, an identified octopaminergic
621 neuron, VUMmx1, in the suboesophageal ganglion responds to sucrose solution applied to the antenna
622 and proboscis (Hammer, 1993). In *Drosophila*, the octopaminergic/tyraminergic neurons in the
623 suboesophageal ganglion control larval appetite (Zhang et al., 2013) and mediate reward signaling in
624 both larvae and adult flies (Burke et al., 2012; Selcho et al., 2014). Furthermore, in the cricket, the
625 pharmacological blockade of α-adrenergic-like octopamine receptor impaired appetitive learning,
626 indicating that the octopaminergic system is involved in reward signaling (Unoki et al., 2005; Unoki et
627 al., 2006; Matsumoto et al., 2015). These allow us to hypothesize that the feeding of sucrose solution
628 would activate the octopaminergic/tyraminergic system in the suboesophageal ganglion of the cricket.

629 According to this idea, we asked whether the feeding of sucrose solution induces the reporter protein
630 expression in the octopaminergic/tyraminergic neurons in the suboesophageal ganglion, especially
631 focused on the dorsal unpaired median (DUM) neurons (Pflüger and Stevenson, 2005).

632 To label the octopaminergic/tyraminergic neurons in the cricket brain, we first generated anti-
633 *Gryllus* tyrosine decarboxylase 2 (Tdc2) antibody. In *Drosophila*, Tdc2 is responsible for the
634 production of tyramine (a precursor of octopamine) in the nervous tissues (Cole et al., 2005; see Figure
635 7-1), and the *tdc2-Gal4* driver is used for targeted gene expression in the octopaminergic/tyraminergic
636 neurons in *Drosophila*. Our anti-*Gryllus* Tdc2 antibody successfully immunolabeled the
637 octopaminergic/tyraminergic DUM neurons in the suboesophageal ganglion of the Hokudai wild-type
638 strain (Figures 7A and B; see also Figure 7-2), which is confirmed by double immunostaining with the
639 anti-octopamine antibody (Figure 7C). Then, we examined whether reporter protein expression
640 becomes detectable in the nuclei of the DUM neurons in response to the feeding of sucrose solution. In
641 the suboesophageal ganglia of the naïve IEG reporter line, the DUM neurons did not show nuclear-
642 localized staining of EYFP immunoreactivity (n=4; Figure 7D upper columns). As we expected, 6 hr
643 after feeding of sucrose solution, the nuclear-localized reporter protein expression was detected in
644 several neurons in the DUM1 and DUM2 clusters (n=4; Figure 7D lower columns). These data
645 demonstrate that our IEG promoter-driven reporter system can visualize neuronal activation caused by
646 PTX-induced neuronal hyperexcitability, as well as behaviorally-evoked neuronal activation.

DISCUSSION

In the present study, we established an IEG promoter-driven transgenic reporter system for whole-brain activity mapping in the cricket. In the cricket brain, the neuronal IEGs, *Gryllus fra-B*, *egr-B* and *hr38*, were up-regulated within 1 hr after PTX-induced neuronal hyperexcitability, and after feeding of sucrose solution and agonistic interaction. We found that the region ~1.6 kbp upstream of *Gryllus egr-B* is sufficient to drive reporter gene expression in response to neuronal hyperexcitability. Inducible reporter protein expression was broadly detected in the brain of the IEG reporter line. Finally, we detected feeding-evoked neuronal activations in the cricket brain with our reporter system. Our study also revealed that the core promoter region of *egr-B* homologs show a high level of structural conservation in basal insects and crustaceans.

The *egr-1/zif268/NGFI-A* gene was first identified as an IEG induced by a variety of extracellular stimuli (e.g. serum, mitogens, and growth factors) in various cell types (Sukhatme et al., 1987; Milbrandt, 1987), and was later shown to be induced after depolarization in neurons (Sukhatme et al., 1988). Then, it became widely used as a marker for neuronal activation in various vertebrate species (Brennan et al., 1992; Mello and Clayton, 1994; Brennan et al., 1999; Hoke et al., 2004; Burmeister and Fernald, 2005). In the cricket brain, *Gryllus egr-B* was rapidly induced as an IEG in response to neuronal activation. Neuronal activity-regulated expression of *egr* homologs was also reported in the brain of *Drosophila* (Guan et al. 2005; Chen et al., 2016) and the honeybee (Ugajin et al., 2013). These data support the idea that the *egr-1/zif268/NGFI-A* homologs are evolutionarily-conserved, neuronal IEGs between vertebrates and insects.

In the present study, using the promoter region of *Gryllus egr-B* and the nuclear-targeted destabilized EYFP as a reporter, we established a retroactive, whole-mount, single-cell resolution activity mapping system to highlight behaviorally-relevant neuronal circuits in the cricket brain. Transgenic reporter systems for neuronal activity mapping have been constructed in the mouse and

671 *Drosophila* by using innate activity-dependent promoters (Smeyne et al., 1992; Wang et al., 2006;
672 Reijmers et al., 2007; Barth, 2007; Kim et al., 2015), as well as synthetic promoters (Kawashima et al.,
673 2013; Fujita et al., 2013; Sørensen et al., 2016) or a synthetic activity-regulated transcription factor
674 (Masuyama et al., 2012). In rodents, the activity-dependent promoters were used to express the
675 channelrhodopsin gene in the behaviorally-relevant neurons to investigate the formation of memory
676 engrams (Garner et al., 2012; Liu et al., 2012; Ramirez et al., 2013). In our future research, we will
677 combine our IEG promoter-based gene expression system with ontogenetic/chemogenetic tools or
678 genetically encoded calcium indicators to uncover the key neural circuits involving cricket social
679 behaviors. Moreover, recent advances in genetic engineering techniques (transposon-based germline
680 transgenesis and CRISPR/Cas-based knock-in; O'Brochta and Atkinson, 1996; Atkinson et al., 2001;
681 Kimura, 2001; Handler, 2002; Sun et al., 2017) allows us to introduce the IEG promoter-based reporter
682 system into non-*Drosophila* insects classically used in insect neuroethology.

683 When neuronal IEG mRNAs are induced in the vertebrate nervous system, expression reaches
684 its peak around 30 min after stimulation, and then rapidly returns to baseline by 60 min (Zangenehpour
685 and Chaudhuri, 2002, Mello and Clayton, 1994, Burmeister et al., 2005). On the contrary, in the cricket
686 brain, the expression levels of neuronal IEGs peaked more slowly (60-90 min after PTX injection;
687 Figures 2, 1-5, and 2-1). This is consistent with previous studies in other insect species (Kiya et al.,
688 2007; Ugajin et al., 2012; Fujita et al., 2013; Ugajin et al., 2016; Ugajin et al., *in press*). These data
689 indicate that, in general, mRNAs of insect neuronal IEGs show slower expression kinetics than those in
690 the vertebrate systems. The expression dynamics of inducible gene mRNA is determined depending on
691 the rates of transcription and mRNA decay (Shalem et al., 2008). In the cricket brain, transcription of
692 *Gryllus egr-B* was rapidly initiated upon stimulation, and promoter activity was sustained
693 at a high level for over 1hr after stimulation. In the vertebrate system, neuronal IEGs are categorized
694 into two subgroups: rapid and delayed IEGs (Saha and Dudek, 2013). The expression of delayed IEGs

695 requires several time-consuming steps to recruit Pol II to the promoter for the initiation of transcription
696 (e.g., signal transduction, post-translational modifications of transcription factors, and chromatin
697 remodeling), whereas transcription of rapid IEGs is initiated by stalled Pol II. In insects, Pol II stalling-
698 dependent rapid transcription initiation was found in *Drosophila* heat shock response genes (Teves and
699 Henikoff, 2013), where inducible gene expression becomes detectable within 5 min after stimulation
700 (Vazquez et al., 1993). Rapid initiation of *Gryllus egr-B* transcription suggests that this gene is
701 expressed as a rapid IEG in the cricket brain. This is also supported by the existence of promoter
702 elements for Pol II stalling in the core promoter region of *Gryllus egr-B*, which is discussed in detail
703 in the following section. The level of *Gryllus egr-B* mRNA is decreased when its promoter activity is
704 down-regulated, indicating *Gryllus egr-B* mRNA undergoes rapid degradation. Our expression analysis
705 also revealed that another neuronal IEG, *Gryllus hr38*, has a slower induction kinetics with
706 exceltionally long mRNA half-life when compared to the case of *Gryllus egr-B* (see Figure 1-5),
707 suggesting that *Gryllus hr38* is expressed as a delayed IEG in the cricket brain.

708 As mentioned above, the core promoter region of *Gryllus egr-B* contained promoter elements
709 associated with Pol II stalling, which are highly conserved in the upstream region of basal
710 insects/crustacean *egr-B* homologs. Pol II stalling was first reported on *Drosophila* heat shock genes
711 (Gilmour and Lis, 1986). Now, Pol II stalling is considered as a widespread mechanism of transcription
712 regulation for rapid and precise control of gene expression in metazoans (Core and Lis, 2008; Nechaev
713 et al., 2010; Levine, 2011). Genome-wide survey in *Drosophila* (Hendrix et al., 2008) revealed that
714 ~20% of stalled promoters contain an arrangement of core promoter elements, namely, GAGA motif,
715 Inr, GC-rich motif, and/or DPE. Our promoter analysis revealed that all of the core promoter elements
716 found in *Drosophila* stalled promoters are contained in the core promoter region of *Gryllus egr-B* and
717 basal insect *egr-B* homologs. The GAGA motif is found in the *Drosophila hsp70* promoter (Shopland
718 et al., 1995; Wilkins and Lis, 1997) and other stalled promoters (Lee et al., 2008; Hendrix et al., 2008).

719 In *Drosophila*, a GAGA transcription factor, GAF, occupies the GAGA motif in the stalled promoter,
720 and recruits transcription co-regulators, such as the nucleosome remodeling factor and the negative
721 elongation factor, to the core promoter region (Fuda et al., 2015; Tsai et al., 2016). In *Drosophila*, the
722 Inr and DPE are enriched by about 2-fold in the stalled promoters relative to the core promoter region
723 of constitutively expressed genes, respectively (Hendrix et al., 2008). These elements are highly
724 conserved in the core promoter of basal insect *egr-B* homologs, as are some sequence homologies of
725 *Drosophila* stalled promoters (Hendrix et al., 2008). In the present study, we found a GC-rich motif
726 between the TSS and DPE of the core promoter region of basal insect *egr-B* homologs. In *Drosophila*
727 embryo, ~25% of stalled promoters contain GC-rich region between +1 and +50, which could promote
728 Pol II stalling by attenuating the movement of the Pol II complex (Hendrix et al., 2008). To our
729 surprise, the putative core promoter region of *Daphnia egr-B* homolog shares the similar arrangement
730 of *cis*-regulatory and core promoter elements found in the upstream sequence of basal insect *egr-B*
731 homologs (Figure 3A). This finding indicates that ancestral arthropods acquired a prototypic *egr-B* core
732 promoter before the diversification of insects from crustaceans (~500 million years ago; Misof et al.,
733 2014). To our knowledge, the core promoter region of insect/crustacean *egr-B* homologs is the most
734 ancestral structurally conserved activity-regulated promoter.

735 In addition to the structural analysis of the core promoter, in the present study we investigated
736 the spatial distribution of *cis*-regulatory elements in the upstream region of insect *egr-B* homologs. In
737 vertebrates, structural and functional analyses of activity-regulated promoters/enhancers revealed
738 transcription factor binding sites essential for inducible gene expression (West et al., 2002; Benito and
739 Barco, 2015). For example, Kawashima et al. (2009) identified a synaptic activity-responsive CRM at
740 ~7 kb upstream of the TSS of the mouse *Arc* gene, which consists of closely localized binding sites for
741 CREB, MEF2, and SRF. The promoter region of the *c-fos* gene contains binding sites for SRF and
742 CREB, which play crucial roles in activity-regulated gene expression (Ghosh et al., 2004). The

743 promoter region of the *down syndrome critical region isoform 4* gene contains several binding sites for
744 NF-AT, which are required for depolarization-dependent expression of the gene (Vihma et al., 2016).
745 Although biochemical properties of stimulus-regulated transcription factors have not been
746 characterized in insects, our exploratory *in silico* analysis showed structural conservation in the DNA-
747 binding domains of stimulus-regulated transcription factors among insects and vertebrates. This
748 suggests insect stimulus-regulated transcription factors share similar DNA-binding properties to their
749 vertebrate counterparts (Table 3-1 and Figure 3-1). Structural analyses of the promoter region of *egr-B*
750 homologs first revealed that the putative core promoter region of basal insect/crustacean *egr-B*
751 homologs contains two conserved CREs and one SRF. Also, potential binding sites for stimulus-
752 regulated transcription factors are enriched in the distal upstream regions of insect/crustacean *egr-B*
753 homologs (Figure 3B). Comparative structural analysis of the upstream sequences of polyneopteran
754 *egr-B* homologs revealed two conserved CRMs, which contain binding sites for stimulus-regulated
755 transcription factors, such as AP-1, CREB, SRF, and C/EBP. CRMs integrate the multiple transcription
756 factors and their associated co-factors at a specific timing to perform elaborate and accurate regulation
757 (Jeziorska et al., 2009). In fact, when we stimulated an individual intracellular signaling pathway, the
758 expression of *Gryllus egr-B* was not induced (see Figure 2-2). Although the reason is unclear, it is
759 possible that synergic activation of multiple signaling pathways and their downstream transcription
760 factors are necessarily to induce expression of *Gryllus egr-B* in the cricket brain. Further studies are
761 necessary to understand the upstream signal transduction pathways involved in transcriptional
762 activation of insect *egr-B* homologs.

763 In the vertebrate nervous system, *egr* family genes are involved in long-term plasticity and the
764 formation and consolidation of long-term memory (Poirier et al., 2008). Genome-wide analysis for the
765 direct target of Egr-1 protein in PC12 neurons revealed that a majority of Egr-1 target genes are down-
766 regulated when *egr-1* is overexpressed, suggesting that the Egr-1 protein principally functions as a

767 transcription suppressor (James et al., 2005). However, other studies revealed that, the activity-induced
768 Egr-1 and Egr-3 proteins up-regulate the expression of $GABA_A$ receptor subunits in the hippocampal
769 neurons to regulate homeostatic excitatory/inhibitory balance (Roberts et al., 2005; Mo et al., 2015).
770 These controversial findings suggest that the target gene and transcriptional regulation of the Egr
771 family proteins are determined according to cell-type and/or in gene-dependent manners. In the insect
772 nervous system, the function of the *egr* homolog, as well as its downstream target genes are poorly
773 understood. Further studies are needed to address the biological significance of the activity-regulated
774 expression of the *egr* homolog in the insect brain.

ACKNOWLEDGEMENTS

We thank Ryan B. Pruett for reviewing. This work was supported, in part, by the Grant-in-Aid for Japanese Society for the Promotion of Science (JSPS) Research Fellow (24•3065) to TW; the JSPS KAKENHI (16K20875) to TW; and Akiyama Life Science Foundation research grant 2014 to TW. The white-eye mutant of *G. bimaculatus* (*gwhite* mutant) was kindly gifted by Prof. Sumihare Noji in the University of Tokushima. The pXL-BacII, pBSII-IFP2-orf, and pBSII-ITR1.1k-EYFP plasmids were kindly gifted by Prof. Malcolm J. Fraser, Jr. in the University of Notre Dame.

AUTHOR CONTRIBUTIONS

T.W. designed the study and wrote the manuscript with the help of A.U.. T.W. performed all experiments and analyzed the data. All the authors reviewed the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare that they have no conflicts of interest.

789 **REFERENCES**

790 Adamo SA, Linn CE, Hoy RR (1995) The role of neurohormonal octopamine during 'fight or flight'
791 behaviour in the field cricket *Gryllus bimaculatus.* J Exp Biol 198:1691-700.

792 Alaux C, Robinson GE (2007) Alarm pheromone induces immediate-early gene expression and slow
793 behavioral response in honey bees. J Chem Ecol 33:1346-1350.

794 Andersen CL, Jensen JL, Ørntoft TF (2004) Normalization of real-time quantitative reverse
795 transcription-PCR data: a model-based variance estimation approach to identify genes suited for
796 normalization, applied to bladder and colon cancer data sets. Cancer Res 64:5245-5250.

797 Atkinson PW, Pinkerton AC, O'Brochta DA (2001) Genetic transformation systems in insects. Annu
798 Rev Entomol 46:317-346.

799 Barth AL (2007) Visualizing circuits and systems using transgenic reporters of neural activity. Curr
800 Opin Neurobiol 17:567-571.

801 Benito E, Barco A (2015) The neuronal activity-driven transcriptome. Mol Neurobiol 51:1071-1088.

802 Benito E, Valor LM, Jimenez-Minchan M, Huber W, Barco A (2011) cAMP response element-binding
803 protein is a primary hub of activity-driven neuronal gene expression. J Neurosci 31:18237-18250.

804 Brennan PA, Hancock D, Keverne EB (1992) The expression of the immediate-early genes *c-fos*, *egr-1*
805 and *c-jun* in the accessory olfactory bulb during the formation of an olfactory memory in mice.
806 Neuroscience 49:277-284.

807 Brennan PA, Schellinck HM, Keverne EB (1999) Patterns of expression of the immediate-early gene
808 *egr-1* in the accessory olfactory bulb of female mice exposed to pheromonal constituents of male urine.
809 Neuroscience 90, 1463-1470.

810 Burke CJ, Huetteroth W, Owald D, Perisse E, Krashes MJ, Das G, Gohl D, Silies M, Certel S, Waddell

811 S (2012) Layered reward signalling through octopamine and dopamine in *Drosophila*. Nature 492:433-
812 437.

813 Burmeister SS, Fernald RD (2005) Evolutionary conservation of the *egr-1* immediate-early gene
814 response in a teleost. J Comp Neurol 481:220-232.

815 Burmeister SS, Jarvis ED, Fernald RD (2005) Rapid behavioral and genomic responses to social
816 opportunity PLoS Biol 3, e363.

817 Ciapponi L, Jackson DB, Mlodzik M, Bohmann D (2001) *Drosophila* Fos mediates ERK and JNK
818 signals via distinct phosphorylation sites. Genes Dev 15:1540-1553.

819 Chen X, Rahman R, Guo F, Rosbash M (2016) Genome-wide identification of neuronal activity-
820 regulated genes in *Drosophila*. Elife 5, e19942.

821 Clayton DF (2000) The genomic action potential. Neurobiol Learn Mem 74:185-216.

822 Cole SH, Carney GE, McClung CA, Willard SS, Taylor BJ and Hirsh J (2005) Two functional but
823 noncomplementing *Drosophila* tyrosine decarboxylase genes: distinct roles for neural tyramine and
824 octopamine in female fertility. J Biol Chem 280: 14948-14955.

825 Core LJ, Lis JT (2008) Transcription regulation through promoter-proximal pausing of RNA
826 polymerase II. Science 319:1791-1792.

827 Cress WD, Triezenberg SJ (1991) Critical structural elements of the VP16 transcriptional activation
828 domain. Science 251:87-90.

829 Dolmetsch RE, Lewis RS, Goodnow CC, Healy JI (1997) Differential activation of transcription factors
830 induced by Ca^{2+} response amplitude and duration. Nature 386:855-858.

831 Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput.
832 Nucleic Acids Res 32:1792-1797.

833 Flavell SW, Greenberg ME (2008) Signaling mechanisms linking neuronal activity to gene expression
834 and plasticity of the nervous system. Annu Rev Neurosci 31:563-590.

835 Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I (2004) VISTA: computational tools for
836 comparative genomics. Nucleic Acids Res 32:273-279.

837 Frommer G, Vorbrüggen G, Pasca G, Jäckle H, Volk T (1996) Epidermal egr-like zinc finger protein of
838 *Drosophila* participates in myotube guidance. EMBO J 15:1642-1649.

839 Fuda NJ, Guertin MJ, Sharma S, Danko CG, Martins AL, Siepel A, Lis JT (2015) GAGA factor
840 maintains nucleosome-free regions and has a role in RNA polymerase II recruitment to promoters.
841 PLoS Genet 11, e1005108.

842 Fujita N, Nagata Y, Nishiuchi T, Sato M, Iwami M, Kiya T (2013) Visualization of neural activity in
843 insect brains using a conserved immediate-early gene *Hr38*. Curr Biol 23:2063-2070.

844 Garner AR, Rowland DC, Hwang SY, Baumgaertel K, Roth BL, Kentros C, Mayford M (2012)
845 Generation of a synthetic memory trace. Science 335:1513-1516.

846 Ghosh A, Ginty DD, Bading H, Greenberg ME (1994) Calcium regulation of gene expression in
847 neuronal cells. J Neurobiol 25:294-303.

848 Gilmour DS, Lis JT (1986) RNA polymerase II interacts with the promoter region of the noninduced
849 hsp70 gene in *Drosophila melanogaster* cells. Mol Cell Biol 6:3984-3989 .

850 Gonzalez-Bellido PT, Wardill TJ (2012) Labeling and confocal imaging of neurons in thick
851 invertebrate tissue samples. Cold Spring Harb Protoc 2012:969-983.

852 Griffin R, Binari R, Perrimon N (2014) Genetic odyssey to generate marked clones in *Drosophila*
853 mosaics. Proc Natl Acad Sci USA 111:4756-4763.

854 Guan Z, Saraswati S, Adolfsen B, Littleton JT (2005) Genome-wide transcriptional changes associated

855 with enhanced activity in the *Drosophila* nervous system. Neuron 48:91-107.

856 Guzowski JF, Timlin JA, Roysam B, McNaughton BL, Worley PF, Barnes CA (2005) Mapping
857 behaviorally relevant neural circuits with immediate-early gene expression. Curr Opin Neurobiol
858 15:599-606.

859 Hammer M (1993) An identified neuron mediates the unconditioned stimulus in associative olfactory
860 learning in honeybees. Nature 366:59-63.

861 Handler AM (2002) Use of the *piggyBac* transposon for germ-line transformation of insects. Insect
862 Biochem Mol Biol 32:1211-1220.

863 Hedwig B (2006) Pulses patterns and paths: neurobiology of acoustic behaviour in crickets. J Comp
864 Physiol A Neuroethol Sens Neural Behav Physiol 192:677-689.

865 Hendrix DA, Hong JW, Zeitlinger J, Rokhsar DS, Levine MS (2008) Promoter elements associated
866 with RNA Pol II stalling in the *Drosophila* embryo. Proc Natl Acad Sci USA 105:7762-7767.

867 Hogan PG, Chen L, Nardone J, Rao A (2003) Transcriptional regulation by calcium calcineurin and
868 NFAT. Genes Dev 17:2205-2232.

869 Hoke KL, Burmeister SS, Fernald RD, Rand AS, Ryan MJ, Wilczynski W (2004) Functional mapping
870 of the auditory midbrain during mate call reception. J Neurosci 24:11264-11272.

871 Horch HW, Mito T, Popadić A, Ohuchi H, Noji S, ed (2017) The cricket as a model organism:
872 Development regeneration and behavior. Tokyo: Springer Japan.

873 Hu J, Roy SK, Shapiro PS, Rodig SR, Reddy SP, Platanias LC, Schreiber RD, Kalvakolanu DV (2001)
874 ERK1 and ERK2 activate CCAAAT/enhancer-binding protein-beta-dependent gene transcription in
875 response to interferon-gamma. J Biol Chem 276:287-297.

876 Huber F, Moore TE, Loher W (1989) Cricket Behavior and Neurobiology. New York : Cornell

University Press.

Hulboy DL, Lozano G (1994) Structural and functional analysis of p53: the acidic activation domain has transforming capability. Cell Growth Differ 5:1023-1031.

James AB, Conway AM, Morris BJ (2005) Genomic profiling of the neuronal target genes of the plasticity-related transcription factor -- Zif268. J Neurochem 95:796-810.

Jenett A, Rubin GM, Ngo TT, Shepherd D, Murphy C, Dionne H, Pfeiffer BD, Cavallaro A, Hall D, Jeter J, Iyer N, Fetter D, Hausenfluck JH, Peng H, Trautman ET, Svirskas RR, Myers EW, Iwinski ZR, Aso Y, DePasquale GM, Enos A, Hulamm P, Lam SC, Li HH, Laverty TR, Long F, Qu L, Murphy SD, Rokicki K, Safford T, Shaw K, Simpson JH, Sowell A, Tae S, Yu Y, Zugates CT (2012) A GAL4-driver line resource for *Drosophila* neurobiology. Cell Rep 2:991-1001.

Jeziorska DM, Jordan KW, Vance KW (2009) A systems biology approach to understanding *cis*-regulatory module function. Semin Cell Dev Biol 20:856-862.

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30:772-780.

Kawashima T, Kitamura K, Suzuki K, Nonaka M, Kamijo S, Takemoto-Kimura S, Kano M, Okuno H, Ohki K, Bito H (2013) Functional labeling of neurons and their projections using the synthetic activity-dependent promoter E-SARE. Nat Methods 10:889-895.

Kawashima T, Okuno H, Bito H (2014) A new era for functional labeling of neurons: activity-dependent promoters have come of age. Front Neural Circuits 8:37.

Kawashima T, Okuno H, Nonaka M, Adachi-Morishima A, Kyo N, Okamura M, Takemoto-Kimura S, Worley PF, Bito H (2009) Synaptic activity-responsive element in the Arc/Arg31 promoter essential for synapse-to-nucleus signaling in activated neurons. Proc Natl Acad Sci USA 106:316-321.

Kim Y, Venkataraju KU, Pradhan K, Mende C, Taranda J, Turaga SC, Arganda-Carreras I, Ng L,

900 Hawrylycz MJ, Rockland KS, Seung HS, Osten P (2015) Mapping social behavior-induced brain
901 activation at cellular resolution in the mouse. Cell Rep 10:292-305.

902 Kimura K (2001) Transposable element-mediated transgenesis in insects beyond *Drosophila*. J
903 Neurogenet 15:179-192.

904 Kiya T, Kubo T (2011) Dance type and flight parameters are associated with different mushroom body
905 neural activities in worker honeybee brains. PLoS One 6:e19301.

906 Kiya T, Kunieda T, Kubo T (2007) Increased neural activity of a mushroom body neuron subtype in
907 the brains of forager honeybees. PLoS One 2, e371.

908 Lee C, Huang CH (2014) LASAGNA-Search 2.0: integrated transcription factor binding site search
909 and visualization in a browser Bioinformatics 30, 1923-1925.

910 Lee C, Li X, Hechmer A, Eisen M, Biggin MD, Venters BJ, Jiang C, Li J, Pugh BF, Gilmour DS
911 (2008) NELF and GAGA factor are linked to promoter-proximal pausing at many genes in *Drosophila*.
912 Mol Cell Biol 28:3290-3300.

913 Lee T (2014) Generating mosaics for lineage analysis in flies. Wiley Interdiscip Rev Dev Biol 3:69-81.

914 Lee T, Luo L (2001) Mosaic analysis with a repressible cell marker (MARCM) for *Drosophila* neural
915 development. Trends Neurosci 24:251-254.

916 Letunic I, Doerks T, Bork P (2012) SMART 7: recent updates to the protein domain annotation
917 resource. Nucleic Acids Res 40:D302-305.

918 Levine M (2011) Paused RNA polymerase II as a developmental checkpoint. Cell 145:502-511.

919 Li X, Zhao X, Fang Y, Jiang X, Duong T, Fan C, Huang CC, Kain SR (1998) Generation of
920 destabilized green fluorescent protein as a transcription reporter. J Biol Chem 273:34970-34975.

921 Liu X, Ramirez S, Pang PT, Puryear CB, Govindarajan A, Deisseroth K, Tonegawa S (2012)

922 Optogenetic stimulation of a hippocampal engram activates fear memory recall. Nature 484:381-385.

923 Loebrich S, Nedivi E (2009) The function of activity-regulated genes in the nervous system. Physiol
924 Rev 89:1079-1103.

925 Manning L, Heckscher ES, Purice MD, Roberts J, Bennett AL, Kroll JR, Pollard JL, Strader ME,
926 Lupton JR, Dyukareva AV, Doan PN, Bauer DM, Wilbur AN, Tanner S, Kelly JJ, Lai SL, Tran KD,
927 Kohwi M, Laverty TR, Pearson JC, Crews ST, Rubin GM, Doe CQ (2012) A resource for manipulating
928 gene expression and analyzing *cis*-regulatory modules in the *Drosophila* CNS. Cell Rep 2:1002-1013.

929 Maruska KP, Becker L, Neboori A, Fernald RD (2013) Social descent with territory loss causes rapid
930 behavioral endocrine and transcriptional changes in the brain. J Exp Biol 216:3656-3666.

931 Masuyama K, Zhang Y, Rao Y, Wang JW (2012) Mapping neural circuits with activity-dependent
932 nuclear import of a transcription factor. J Neurogenet 26:89-102 .

933 Matsumoto Y, Noji S, Mizunami M (2003) Time course of protein synthesis-dependent phase of
934 olfactory memory in the cricket *Gryllus bimaculatus*. Zoolog Sci 20:409-416.

935 Matsumoto Y, Matsumoto CS, Wakuda R, Ichihara S, Mizunami M (2015) Roles of octopamine and
936 dopamine in appetitive and aversive memory acquisition studied in olfactory conditioning of maxillary
937 palpi extension response in crickets. Front Behav Neurosci 9:230.

938 Matsuo R, Murayama A, Saitoh Y, Sakaki Y, Inokuchi K (2000) Identification and cataloging of genes
939 induced by long-lasting long-term potentiation in awake rats. J Neurochem 74:2239-2249.

940 Meeker ND, Hutchinson SA, Ho L, Trede NS (2007) Method for isolation of PCR-ready genomic
941 DNA from zebrafish tissues. Biotechniques 43:610, 612, 614.

942 Mello CV, Clayton DF (1994) Song-induced ZENK gene expression in auditory pathways of songbird
943 brain and its relation to the song control system. J Neurosci 14:6652-6666.

944 Milbrandt J (1987) A nerve growth factor-induced gene encodes a possible transcriptional regulatory
945 factor. Science 238:797-799.

946 Minatohara K, Akiyoshi M, Okuno H (2016) Role of immediate-early genes in synaptic plasticity and
947 neuronal ensembles underlying the memory trace. Front Mol Neurosci 8:78.

948 Misof B, Liu S, Meusemann K, Peters RS, Donath A, Mayer C, Frandsen PB, Ware J, Flouri T, Beutel
949 RG, Niehuis O, Petersen M, Izquierdo-Carrasco F, Wappler T, Rust J, Aberer AJ, Aspöck U, Aspöck
950 H, Bartel D, Blanke A, Berger S, Böhm A, Buckley TR, Calcott B, Chen J, Friedrich F, Fukui M,
951 Fujita M, Greve C, Grobe P, Gu S, Huang Y, Jermiin LS, Kawahara AY, Krogmann L, Kubiak M,
952 Lanfear R, Letsch H, Li Y, Li Z, Li J, Lu H, Machida R, Mashimo Y, Kapli P, McKenna DD, Meng G,
953 Nakagaki Y, Navarrete-Heredia JL, Ott M, Ou Y, Pass G, Podsiadlowski L, Pohl H, von Reumont BM,
954 Schütte K, Sekiya K, Shimizu S, Slipinski A, Stamatakis A, Song W, Su X, Szucsich NU, Tan M, Tan
955 X, Tang M, Tang J, Timelthaler G, Tomizuka S, Trautwein M, Tong X, Uchifune T, Walzl MG,
956 Wiegmann BM, Wilbrandt J, Wipfler B, Wong TK, Wu Q, Wu G, Xie Y, Yang S, Yang Q, Yeates DK,
957 Yoshizawa K, Zhang Q, Zhang R, Zhang W, Zhang Y, Zhao J, Zhou C, Zhou L, Ziesmann T, Zou S, Li
958 Y, Xu X, Zhang Y, Yang H, Wang J, Wang J, Kjer KM, Zhou X (2014) Phylogenomics resolves the
959 timing and pattern of insect evolution. Science 346:763-767.

960 Mo J, Kim CH, Lee D, Sun W, Lee HW, Kim H (2015) Early growth response 1 (Egr-1) directly
961 regulates $GABA_A$ receptor α2, α4, and θ subunits in the hippocampus. J Neurochem 133:489-500.

962 Nakamura T, Yoshizaki M, Ogawa S, Okamoto H, Shinmyo Y, Bando T, Ohuchi H, Noji S, Mito T
963 (2010) Imaging of transgenic cricket embryos reveals cell movements consistent with a syncytial
964 patterning mechanism. Curr Biol 20, 1641-1647.

965 Nechaev S, Fargo DC, dos Santos G, Liu L, Gao Y, Adelman K (2010) Global analysis of short RNAs
966 reveals widespread promoter-proximal stalling and arrest of Pol II in *Drosophila*. Science 327:335-338.

967 Nedivi E, Hevroni D, Naot D, Israeli D, Citri Y (1993) Numerous candidate plasticity-related genes
968 revealed by differential cDNA cloning. Nature 363:718-722.

969 Nonaka M, Kim R, Sharry S, Matsushima A, Takemoto-Kimura S, Bito H (2014) Towards a better
970 understanding of cognitive behaviors regulated by gene expression downstream of activity-dependent
971 transcription factors. Neurobiol Learn Mem 115:21-29.

972 O'Brochta DA, Atkinson PW (1996) Transposable elements and gene transformation in non-
973 drosophilid insects. Insect Biochem Mol Biol 26:739-753.

974 Okuno H (2011) Regulation and function of immediate-early genes in the brain: beyond neuronal
975 activity markers. Neurosci Res 69:175-186.

976 Owald D, Lin S, Waddell S (2015) Light heat action: neural control of fruit fly behavior. Philos Trans
977 R Soc Lond B Biol Sci 370, 20140211.

978 Pflüger HJ, Stevenson PA (2005) Evolutionary aspects of octopaminergic systems with emphasis on
979 arthropods. Arthropod Struct Dev 34:379-396.

980 Pham DH, Moretti PA, Goodall GJ, Pitson SM (2008) Attenuation of leakiness in doxycycline-
981 inducible expression via incorporation of 3' AU-rich mRNA destabilizing elements. Biotechniques
982 45:155-156, 158, 160 passim.

983 Poirier R, Cheval H, Mailhes C, Garel S, Charnay P, Davis S, Laroche S (2008) Distinct functions of
984 *egr* gene family members in cognitive processes. Front Neurosci 2:47-55.

985 Pollack G (2000) Who what where? Recognition and localization of acoustic signals by insects. Curr
986 Opin Neurobiol 10:763-767.

987 Ramirez S, Liu X, Lin PA, Suh J, Pignatelli M, Redondo RL, Ryan TJ, Tonegawa S (2013) Creating a
988 false memory in the hippocampus. Science 341:387-391.

989 Ramirez S, Tonegawa S, Liu X (2014) Identification and optogenetic manipulation of memory engrams
990 in the hippocampus. Front Behav Neurosci 7:226.

991 Regier JL, Shen F, Triezenberg SJ (1993) Pattern of aromatic and hydrophobic amino acids critical for
992 one of two subdomains of the VP16 transcriptional activator. Proc Natl Acad Sci USA 90:883-887.

993 Reijmers LG, Perkins BL, Matsuo N, Mayford M (2007) Localization of a stable neural correlate of
994 associative memory. Science 317:1230-1233.

995 Riemensperger T, Kittel RJ, Fiala A (2016) Optogenetics in *Drosophila*. Neuroscience Methods Mol
996 Biol 1408:167-175.

997 Roberts DS, Raol YH, Bandyopadhyay S, Lund IV, Budreck EC, Passini MA, Wolfe JH, Brooks-Kayal
998 AR, Russek SJ (2005) Egr3 stimulation of GABRA4 promoter activity as a mechanism for seizure-
999 induced up-regulation of $GABA_A$ receptor α4 subunit expression. Proc Natl Acad Sci USA 102:11894-
1000 11899.

1001 Roy SK, Hu J, Meng Q, Xia Y, Shapiro PS, Reddy SP, Platanias LC, Lindner DJ, Johnson PF,
1002 Pritchard C, Pagés G, Pouyssegur J, Kalvakolanu DV (2002) MEKK1 plays a critical role in activating
1003 the transcription factor C/EBP-beta-dependent gene expression in response to IFN-gamma. Proc Natl
1004 Acad Sci USA 99:7945-7950.

1005 Ruden DM (1992) Activating regions of yeast transcription factors must have both acidic and
1006 hydrophobic amino acids. Chromosoma 101:342-348.

1007 Saffen DW, Cole AJ, Worley PF, Christy BA, Ryder K, Baraban JM (1988) Convulsant-induced
1008 increase in transcription factor messenger RNAs in rat brain. Proc Natl Acad Sci USA 85:7795-7799.

1009 Saha RN, Dudek SM (2013) Splitting hares and tortoises: a classification of neuronal immediate early
1010 gene transcription based on poised RNA polymerase II. Neuroscience 247:175-181.

1011 Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C,

1012 Saalfeld S, Schmid B, Tinevez JY, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A
1013 (2012) Fiji: an open-source platform for biological-image analysis. Nat Methods 9:676-682.

1014 Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis.
1015 Nat Methods 9:671-675.

1016 Schultz J, Milpetz F, Bork P, Ponting CP (1998) SMART, a simple modular architecture research tool:
1017 identification of signaling domains. Proc Natl Acad Sci USA 95:5857-5864.

1018 Sealy L, Malone D, Pawlak M (1997) Regulation of the cfos serum response element by C/EBPβ. Mol
1019 Cell Biol 17:1744-1755.

1020 Selcho M, Pauls D, Huser A, Stocker RF, Thum AS (2014) Characterization of the octopaminergic and
1021 tyraminergic neurons in the central brain of *Drosophila* larvae. J Comp Neurol 522:3485-3500.

1022 Shalem O, Dahan O, Levo M, Martinez MR, Furman I, Segal E, Pilpel Y (2008) Transient
1023 transcriptional responses to stress are generated by opposing effects of mRNA production and
1024 degradation. Mol Syst Biol 4:223.

1025 Shopland LS, Hirayoshi K, Fernandes M, Lis JT (1995) HSF access to heat shock elements in vivo
1026 depends critically on promoter architecture defined by GAGA factor TFIID, and RNA polymerase II
1027 binding sites. Genes Dev 9:2756-2769 .

1028 Smeyne RJ, Schilling K, Robertson L, Luk D, Oberdick J, Curran T, Morgan JI (1992) Fos-lacZ
1029 transgenic mice: mapping sites of gene induction in the central nervous system. Neuron 8:13-23.

1030 Sørensen AT, Cooper YA, Baratta MV, Weng FJ, Zhang Y, Ramamoorthi K, Fropf R, LaVerriere E,
1031 Xue J, Young A, Schneider C, Gøtzsche CR, Hemberg M, Yin JC, Maier SF, Lin Y (2016) A robust
1032 activity marking system for exploring active neuronal ensembles. Elife 5:e13918.

1033 Spiegel I, Mardinly AR, Gabel HW, Bazinet JE, Couch CH, Tzeng CP, Harmin DA, Greenberg ME
1034 (2014) Npas4 regulates excitatory-inhibitory balance within neural circuits through cell-type-specific

1035 gene programs. Cell 157:1216-1229.

1036 Stevenson PA, Schildberger K (2013) Mechanisms of experience dependent control of aggression in
1037 crickets. Curr Opin Neurobiol 23:318-323.

1038 Sukhatme VP, Cao XM, Chang LC, Tsai-Morris CH, Stamenkovich D, Ferreira PC, Cohen DR,
1039 Edwards SA, Shows TB, Curran T, Le Beau MM, Adnmson ED (1988) A zinc finger-encoding gene
1040 coregulated with *c-fos* during growth and differentiation and after cellular depolarization. Cell 53:37-
1041 43.

1042 Sukhatme VP, Kartha S, Toback FG, Taub R, Hoover RG, Tsai-Morris CH (1987) A novel early
1043 growth response gene rapidly induced by fibroblast epithelial cell and lymphocyte mitogens. Oncogene
1044 Res 1:343-355.

1045 Sun D, Guo Z, Liu Y, Zhang Y (2017) Progress and prospects of CRISPR/Cas systems in insects and
1046 other arthropods. Front Physiol 8:608.

1047 Teves SS, Henikoff S (2013) The heat shock response: A case study of chromatin dynamics in gene
1048 regulation. Biochem Cell Biol 91:42-48.

1049 Tsai SY, Chang YL, Swamy KB, Chiang RL, Huang DH (2016) GAGA factor a positive regulator of
1050 global gene expression modulates transcriptional pausing and organization of upstream nucleosomes.
1051 Epigenetics Chromatin 9:32.

1052 Tullai JW, Schaffer ME, Mullenbrock S, Sholder G, Kasif S, Cooper GM (2007) Immediate-early and
1053 delayed primary response genes are distinct in function and genomic architecture. J Biol Chem
1054 282:23981-23995.

1055 Ugajin A, Kiya T, Kunieda T, Ono M, Yoshida T, Kubo T (2012) Detection of neural activity in the
1056 brains of Japanese honeybee workers during the formation of a "hot defensive bee ball". PLoS One
1057 7:e32902.

1058 Ugajin A, Kunieda T, Kubo T (2013) Identification and characterization of an Egr ortholog as a neural
1059 immediate-early gene in the European honeybee (*Apis mellifera* L). FEBS Lett 587:3224-3230.

1060 Ugajin A, Uchiyama H, Miyata T, Sasaki T, Yajima S, Ono M (in press) Identification and initial
1061 characterization of novel neural immediate early genes possibly differentially contributing to foraging-
1062 related learning and memory processes in the honeybee. Insect Mol Biol doi: 101111/imb12355.

1063 Ugajin A, Watanabe T, Uchiyama H, Sasaki T, Yajima S, Ono M (2016) Expression analysis of Egr-1
1064 ortholog in metamorphic brain of honeybee (*Apis mellifera* L): Possible evolutionary conservation of
1065 roles of Egr in eye development in vertebrates and insects. Biochem Biophys Res Commun 478:1014-
1066 1019.

1067 Unoki S, Matsumoto Y, Mizunami M (2005) Participation of octopaminergic reward system and
1068 dopaminergic punishment system in insect olfactory learning revealed by pharmacological study. Eur J
1069 Neurosci 22:1409-1416.

1070 Unoki S, Matsumoto Y, Mizunami M (2006) Roles of octopaminergic and dopaminergic neurons in
1071 mediating reward and punishment signals in insect visual learning. Eur J Neurosci 24:2031-2038.

1072
1073 Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F (2002)
1074 Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple
1075 internal control genes. Genome Biol 3:RESEARCH0034.

1076 Vazquez J, Pauli D, Tissières A (1993) Transcriptional regulation in *Drosophila* during heat shock: a
1077 nuclear run-on analysis. Chromosoma 102:233-248.

1078 Vihma H, Luhakooder M, Pruunsild P, Timmusk T (2016) Regulation of different human NFAT
1079 isoforms by neuronal activity. J Neurochem 137:394-408.

1080 Wada K, Howard JT, McConnell P, Whitney O, Lints T, Rivas MV, Horita H, Patterson MA, White

1081 SA, Scharff C, Haesler S, Zhao S, Sakaguchi H, Hagiwara M, Shiraki T, Hirozane-Kishikawa T, Skene
1082 P, Hayashizaki Y, Carninci P, Jarvis ED (2006) A molecular neuroethological approach for identifying
1083 and characterizing a cascade of behaviorally regulated genes. Proc Natl Acad Sci USA 103:15212-
1084 15217.

1085 Wang KH, Majewska A, Schummers J, Farley B, Hu C, Sur M, Tonegawa S (2006) In vivo two-photon
1086 imaging reveals a role of arc in enhancing orientation specificity in visual cortex. Cell 126:389-402.

1087 Watanabe T, Aonuma H (2014) Tissue-specific promoter usage and diverse splicing variants of found
1088 in neurons an ancestral Hu/ELAV-like RNA-binding protein gene of insects in the direct-developing
1089 insect *Gryllus bimaculatus*. Insect Mol Biol 23:26-41.

1090 Watanabe T, Ochiai H, Sakuma T, Horch HW, Hamaguchi N, Nakamura T, Bando T, Ohuchi H,
1091 Yamamoto T, Noji S, Mito T (2012) Non-transgenic genome modifications in a hemimetabolous insect
1092 using zinc-finger and TAL effector nucleases. Nat Commun 3:1017.

1093 Watanabe T, Sadamoto H, Aonuma H (2011) Identification and expression analysis of the genes
1094 involved in serotonin biosynthesis and transduction in the field cricket *Gryllus bimaculatus*. Insect Mol
1095 Biol 20:619-635.

1096 Watanabe T, Takeuchi H, Kubo T (2010) Structural diversity and evolution of the N-terminal isoform-
1097 specific region of ecdysone receptor-A and -B1 isoforms in insects. BMC Evol Biol 10:40.

1098 West AE, Griffith EC, Greenberg ME (2002) Regulation of transcription factors by neuronal activity.
1099 Nat Rev Neurosci 3:921-931.

1100 Wilkins RC, Lis JT (1997) Dynamics of potentiation and activation: GAGA factor and its role in heat
1101 shock gene regulation. Nucleic Acids Res 25:3963-3968.

1102 Zagurovskaya M, Shareef MM, Das A, Reeves A, Gupta S, Sudol M, Bedford MT, Prichard J,
1103 Mohiuddin M, Ahmed MM (2009) EGR-1 forms a complex with YAP-1 and upregulates Bax

1104 expression in irradiated prostate carcinoma cells. Oncogene 28:1121-1131.

1105 Zangenehpour S, Chaudhuri A (2002) Differential induction and decay curves of *c-fos* and *zif268*
1106 revealed through dual activity maps. Brain Res Mol Brain Res 109:221-225.

1107 Zhang T, Branch A, Shen P (2013) Octopamine-mediated circuit mechanism underlying controlled
1108 appetite for palatable food in *Drosophila*. Proc Natl Acad Sci USA 110:15431-15436.

FIGURE LEGENDS

Figure 1. Expression of candidate neuronal IEGs in the brain of cycloheximide pre-treated crickets 30 min after PTX injection.

(A) PTX-induced neuronal hyperexcitability in the cricket. Crickets show seizure-like behavior about 2 minutes after PTX injection. One hour before PTX/vehicle injection, 20 mM cycloheximide was injected to block *de novo* protein synthesis. (B-G) Expression of (B) *Gryllus fra* total transcript, (C) *fra-A* isoform, (D) *fra-B* isoform, (E) *jra*, (F) *egr-B*, and (G) *hr38* in the brains of cycloheximide pre-treated crickets 30 min after injection of vehicle (5% DMSO in saline) or PTX. Expression levels of each target gene was normalized with that of Gryllus *ef1α* gene (See Figure 1-1). RT-qPCR analyses were performed on eight biological replicates. Box plots indicate the 25-75 percentile ranges and central values. Error bars indicate the 5-95 percentile ranges. The "+" denotes the mean. Asterisks denote statistical significance (*, $p<0.05$). See Table 3 for the details of statistical analysis. See Figure 1-2, 1-3, 1-4, and 1-5 for the structures of the encoded proteins of candidate neuronal IEGs.

Figure 2. Expression characteristics of *Gryllus egr-B* in the cricket brain.

(A, B) Expression time-course of *Gryllus egr-B* after PTX injection in the brains of (A) cycloheximide and (B) saline pre-treated crickets. The expression levels were normalized to the mean of those of naïve animals (baseline expression level). Asterisks indicate statistical significance to control (0 min after PTX injection). (C) Expression time-course of *Gryllus egr-B* pre-mRNA in the brain of cycloheximide pre-treated crickets after PTX-injection. (D, E) Behaviorally evoked expression of *Gryllus egr-B* in the brain of crickets 1 hr after (D) sucrose feeding and (E) agonistic interaction. RT-qPCR analyses were performed on eight biological replicates. Box plots indicate the 25-75 percentile ranges and central

values. Error bars indicate the 5-95 percentile ranges. The "+" denotes the mean. Asterisks donate statistical significance to the control (0 min after PTX injection) (A-C) or to the naïve animals (D, E) (*, $p<0.05$; **, $p<0.01$; ***, $p<0.001$; ****, $p<0.0001$). See Table 3 for the details of statistical analysis. See Figure 2-1, 2-2, and 1-5 for the expression characteristics of other candidate neuronal IEGs.

Figure 3. Gene regulatory regions of the insect/crustacean *egr-B* homologs.

(A) Putative core promoter regions of basal insect and crustacean *egr-B* homologs share a high-level sequence similarity. The upstream sequences of insect/crustacean *egr-B* homologs are aligned with the core promoter region of *Gryllus egr-B*. The conserved bases are marked with asterisks under the alignment. *Cis*-regulatory elements and sequence motifs that are conserved are indicated above the alignment. CRE, cAMP-responsive element; SRE, serum response element; Inr, initiator element; DPE, downstream promoter element. (B) Sequence logo representation of the conserved motifs in the core promoter region of insect *egr-B* homologs. The sequence logo of the GAGA motif was generated by multiple alignment of the upstream sequences of polyneopteran *egr-B* homologs. The other sequence logos were generated by multiple alignment of the upstream sequences of insect *egr-B* homologs. The positions of conserved motifs are indicated by black bars under the logo. (C) Schematic representation of the gene regulatory regions of insect/crustacean *egr-B* homologs. The genomic regions were aligned to the position of the +1 site of *Gryllus egr-B* or the 5'-end of the putative core promoter region. The red bars indicate genomic regions aligned in Figure 3A. Positions of transcription factor binding sites predicted using the LASAGNA-Search 2.0 program (score above 8.0) are indicated by arrowheads. The phylogenetic relationship of insect/crustacean species is indicated as a phylogram tree. AP-1, activator protein 1; CREB, cAMP response element-binding protein; C/EBP, CCAAT-enhancer-binding protein; MEF2, myocyte enhancer factor 2; NF-AT, nuclear factor of activated T-cells; SRF, serum response

1156 factor. See Table 2 for the details of genomic sequences used for promoter analysis. See Figure 3-1 and
1157 Table 3-1 for the structural conservations of the transcription factors used for the binding site
1158 prediction.

1159

1160 **Figure 4. Phylogenetic footprinting revealed conserved *cis*-regulatory modules in the upstream**
1161 **regions of polyneopteran *egr-B* homologs.**

1162 (A) mVISTA plot of the upstream regions of polyneopteran *egr-B* homologs based on MLAGAN
1163 alignment using the upstream region of *Gryllus egr-B* as a reference sequence. Positions of potential
1164 transcription factor binding sites in the upstream region of *Gryllus egr-B* are indicated by arrowheads
1165 (see Figure 3B). The horizontal and vertical axes of the plot represent the position in the sequences and
1166 the percent identity, respectively. Two conserved *cis*-regulatory modules (CRMs; CRM_{-800} and CRM_{-400}) and the conserved core promoter region are shaded blue and red on the plot, respectively. (B)
1168 Nucleotide sequence alignments of two conserved CRMs (CRM_{-800} and CRM_{-400}) found in the upstream
1169 region of polyneopteran *egr-B* homologs. The conserved bases are marked with asterisks under the
1170 alignment. *Cis*-regulatory elements conserved among most of the sequences are indicated above the
1171 alignment. Black bars under the alignments indicate sequence motifs conserved across species where
1172 no transcription factor is assigned. AP-1, binding site for activator protein 1; AP-4, binding site for
1173 activating enhancer binding protein 4; ATF2, binding site for activating transcription factor 2;
1174 CDP/Cut, binding site for CCAAT-displacement protein/cut homeobox; C/EBP, binding site for
1175 C/EBP; CRE, cAMP-responsive element; SRE, serum response element. See Table 2 for the details of
1176 genomic sequences used for promoter analysis.

1177

1178 **Figure 5. IEG promoter-driven transgenic reporter system in the cricket brain.**

(A) A flow chart of the experimental procedures to establish the IEG reporter line. See the material and methods section for detail. (B) Schematic representation of the *piggyBac* transgenic vector for the IEG promoter-driven transgenic reporter system. The vector harbors the expression cassette of *EYFPnls:PEST* driven by the *Gryllus egr-B* promoter. *3xP3-mCherry* was used as a visible selection marker. A *gypsy* insulator sequence was inserted between two expression cassettes. ARE, AU-rich element; LTR, long terminal repeat. (Ci) Schematic representation of the *piggyBac* insertion in the IEG reporter line. A 5,629-bp insertion was inserted into the *piggyBac* donor TTAA site (highlighted in red). To conduct genotyping PCR, two primers, line19_fw and line19_rv, were designed at the 5' and 3' flanking region of the insertion sites, respectively. (Cii) The nucleotide sequence of the genomic region flanking the *piggyBac* insertion in the IEG reporter line. The *piggyBac* donor TTAA site is highlighted in red. The positions of the annealing site of primers for genotyping PCR are indicated by white arrows under the sequence. (D) Basal mRNA expressions of *EYFPnls:PEST* and *Gryllus egr-B* in the brain of naïve IEG reporter line. (E) Expression time-course of (Ei) *EYFPnls:PEST* and (Eii) *Gryllus egr-B* in the brain of the IEG reporter line after PTX injection. RT-qPCR analyses were performed on eight biological replicates. Box plots indicate the 25-75 percentile ranges and central values. Error bars indicate the 5-95 percentile ranges. The "+" denotes the mean. Asterisks donate statistical significance (0 min after PTX injection; *, $p<0.05$; **, $p<0.01$; ***, $p<0.001$; ****, $p<0.0001$). (F) Correlation plot between the expression levels of *EYFPnls:PEST* and *Gryllus egr-B* in the brains of the IEG reporter line. The data from PTX-injected crickets (n=75; black circles), saline pre-injected crickets (n=8; gray circles), and naïve crickets (n=8; white circles) were plotted. See Table 3 for the details of statistical analysis.

Figure 6. PTX-induced reporter protein expression in the brain of the IEG reporter line.

Distribution of the reporter protein (EYFPnls:PEST) in the brain of the IEG reporter line was examined

by whole-mount fluorescent immunohistochemistry. (A, B) Frontal views of the supraesophageal ganglion stained with anti-EYFP antibody. (A) EYFP immunoreactivity 6 hr after vehicle injection observed in cells is indicated by white arrowheads. (B) EYFP immunoreactivity 6 hr after PTX injection was observed throughout the ganglion. (C, D) Ventral views of the suboesophageal ganglion stained with anti-EYFP antibody. (C) EYFP immunoreactivity was not observed 6 hr after vehicle injection. (D) EYFP immunoreactivity was observed throughout the ganglion 6 hr after PTX injection. Dorsoventral (D-V) or rostrocaudal (R-C) axes were indicated. Scale bars represent 200 µm. See Movies 1 and 2 for the full stack of optical sections of the supraesophageal ganglia shown in (A) and (B).

Figure 7. Sucrose feeding-evoked reporter protein expression in the DUM neurons of the IEG reporter line.

(A) Dorsal view of the subesophageal ganglion of the wild-type cricket stained with anti-*Gryllus* antibody. The outline of the ganglion is surrounded by the white dotted line. The depth of the cells is color coded as indicated in the inset. Rostrocaudal (R-C) axes were indicated. Scale bar represents 200 µm. See Figure 7-1 for the octopamine biosynthesis pathway and the structures of the Tdc proteins in insects. See Figure 7-2 for the frontal view of the supraesophageal ganglion and the ventral view of the subesophageal ganglion stained with anti-*Gryllus* antibody. (B) Schematic drawing of the positions and numbers of the cell bodies of three DUM clusters (DUM1, DUM2, DUM3) on the dorsal side of the supraesophageal ganglion. (C) Double fluorescent immunostaining confirmed that the DUM neurons contain octopamine. The DUM neurons were stained with the anti-*Gryllus* Tdc2 antibody (green) and anti-octopamine antibody (magenta). Scale bar represents 50 µm. (D) Distribution of the reporter protein (EYFPnls:PEST) in the DUM neurons of the IEG reporter line before and 6 hr after feeding of sucrose solution (n=4 each). The DUM neurons were stained with the anti-*Gryllus* Tdc2 antibody

(green) and anti-GFP antibody (magenta). The cell bodies of *Gryllus* Tdc2 immunoreactive DUM neurons are surrounded by the white dotted line. The DUM neurons with nuclear EYFP immunoreactivity are indicated by white arrowheads. Scale bar represents 50 µm.

Figure 1-1. *Gryllus ef1α* is the most stable internal control for RT-qPCR analysis in the cricket brain.

Validation of housekeeping genes as reference genes for RT-qPCR expression analysis in the cricket brain. (A) Expression levels of three housekeeping genes (*Gryllus β-actin*, *ef1α*, and *rpl32*) in the brain of wild-type adult crickets used for initial expression analysis (n=16 including the PTX-injected and vehicle-injected crickets [n=8, respectively]; see Figure 1). The expression level of each housekeeping gene was normalized to the geometric mean of three housekeeping genes ($GM_{\beta\text{-}actin/ef1\alpha/rpl32}$; Vandesompele et al., 2002). Box plots indicate 25-75 percentile ranges and central values, and '+' indicates mean. Error bars indicate 5-95 percentile ranges. In the cricket brain, *Gryllus β-actin* and *ef1α* genes were expressed at the same levels, but *Gryllus rpl32* was expressed more weakly than other two housekeeping genes. Among three housekeeping genes, the coefficient of variation (CV) of *Gryllus ef1α* was the lowest (CV of *Gryllus β-actin*/$GM_{\beta\text{-}actin/ef1\alpha/rpl32}$, 20.99%; CV of *Gryllus ef1α*/$GM_{\beta\text{-}actin/ef1\alpha/rpl32}$, 8.85%; CV of *Gryllus rpl32*/$GM_{\beta\text{-}actin/ef1\alpha/rpl32}$, 22.5%), indicating *Gryllus ef1α* is the most stably expressed in the cricket brain. (B) Gene expression stability of three housekeeping genes calculated using the Normfinder algorithm (Andersen et al., 2004). NormFinder algorithm supported that *Gryllus ef1α* is the most stable housekeeping gene (stability value of 0.036). The stability value of *Gryllus ef1α* was lower than that of the best combination pair (*Gryllus β-actin* and *ef1α*; stability value of 0.058).

Figure 1-2. *Gryllus fra* gene encodes two protein isoforms closely related to insect and vertebrate

Fos/Fra homologs.

(A) Protein domain structures of *Gryllus* Fra proteins and Fos/Fra homologs of other species. Conserved domains and sequences important for transcriptional regulation are indicated by color boxes. DBD, DNA-binding domain; HOB, homology box; Leu zip, Leucine zipper domain; TBM, TATA-binding protein (TBP)-binding motif. *Gryllus fra-A* encodes a 382-amino-acid protein containing the DBD, Leu zip domains and the C-terminal regulatory domain. Vertebrate c-Fos proteins contain N-terminal and C-terminal transactivation domains and a repression domain at its C-terminus. The core motif of the N-terminal transactivation domain (HOB1-N) is well conserved among vertebrate Fra proteins, but not conserved in insect Fra proteins. Instead, insect Fra proteins contain sequence motifs which might play important roles in transactivation function: *Drosophila* Kayak-A isoform contains a δ-like motif (Ciapponi et al., 2001), and *Drosophila* Kayak-D/F isoforms contain glutamine-rich regions in the isoform-specific N-terminal region. *Gryllus* Fra-A isoform and its related insect Fra proteins contain an acidic patch structurally resembles the acidic activation domains of many eukaryotic transcriptional activators such as Gal4, VP16, p53 and EcR-B1 (Cress and Triezenberg, 1991; Ruden, 1992; Regier et al., 1993; Watanabe et al., 2010). In addition, the T/P-rich region is conserved in most insect Fra proteins. *Gryllus fra-B* encodes an N-terminal truncated 284-amino-acid protein. (B) Sequence alignment of the N-terminal region of *Gryllus* Fra-A and insect Fra/Kayak isoforms, and the C-terminal regulatory domain of Fos/Fra proteins. The conserved residues are marked with asterisks above the alignments. The amino acid residues are represented in the default color scheme of ClustalX. Positions of conserved domains/motifs were indicated by bars under the alignments. GeneBank IDs of proteins are following: *Apis mellifera* Kayak-X1, XP_006564216; *Bombyx mori* Kayak-X2, XP_004921825; *Drosophila melanogaster* Kayak-A, NP_001027579; *D. melanogaster* Kayak-B, NP_001027578; *D. melanogaster* Kayak-D, NP_001027580; *D. melanogaster* Kayak-F, NP_001027577; *Homo sapience* c-Fos, NP_005243; *H. sapience* Fra1, NP_005429: *H.*

sapience Fra2, NP_005244; *H. sapience* Fos-B, NP_006723; *Tribolium castaneum* Kayak-C, NP_001164294.

Figure 1-3. *Gryllus jra* gene encodes a protein closely related to insect and vertebrate Jun/Jra homologs.

(A) Protein domain structures of *Gryllus* Jra protein and Jun/Jra homologs of other species. Conserved domains and sequences important for transcriptional regulation are indicated by color boxes. DBD, DNA-binding domain; HOB; homology box; Leu zip, Leucine zipper domain. (B) Sequence alignment of the conserved domains for transcriptional regulation (δ domain and HOB motifs). The conserved residues are marked with asterisks above the alignments. The amino acid residues are represented in the default color scheme of ClustalX. Positions of conserved domains/motifs were indicated by bars under the alignment. GeneBank IDs of proteins are following: *D. melanogaster* Jra, NP_476586; *H. sapience* c-Jun, NP_002219; *H. sapience* Jun-B, NP_002220; *H. sapience* Jun-D, NP_005345.

Figure 1-4. *Gryllus egr* gene encodes a protein closely related to insect and vertebrate Egr-1 homologs.

(A) Protein domain structures of *Gryllus* Egr-B protein and Egr homologs of other species. Conserved domains and sequences important for transcriptional regulation are indicated by color boxes. Three C_2H_2-type zinc finger domains, as well as a nuclear localization signal (NLS) and a potential acetylation site (Ac), were highly conserved across vertebrate and invertebrate Egr homologs. On the other hand, we found low sequence conservation in the repressor domain between Egr homologs of insect and other species. The WW binding motif (PPxY, where x = any amino acid), which involved in protein-protein interaction with the Yes kinase-associated protein 1 (Zagurovskaya et al., 2009), was

conserved across vertebrate and invertebrate Egr homologs. (B) Sequence alignment of the C-terminal region of Egr proteins. The conserved residues are marked with asterisks above the alignment. Positions of functional domains important for DNA-binding, protein localization, and transcriptional regulation are indicated by bars under the alignment. (C) Sequence alignment of the N-terminal region of insect/crustacean Egr-B proteins. The N-termini of insect/crustacean Egr-B proteins were highly conserved (conserved N-terminal motif 1; consensus sequence: MIM(D/E)FΨ(D/E)TL, where Ψ = bulky hydrophobic residues). Another conserved motif was found at residues from 105 to 140 of *Gryllus* Egr (conserved N-terminal motif 2). These two conserved motifs were only found in the N-terminal region of the insect/crustacean Egr-B proteins, but not in the vertebrate Egr homologs. Another Egr isoform (Egr-A or Stripe-A) found in several insect species (e.g. fruit flies and honeybees) contains an N-terminal extension with polyglutamine stretch (Ugajin et al., 2016). The conserved residues are marked with asterisks above the alignment. The amino acid residues are represented in the default color scheme of ClustalX. The positions of conserved motifs were indicated by bars under the alignment. GeneBank IDs of proteins are following: *Acyrthosiphon pisum* Egr, XP_001943786; *Anoplophora glabripennis* Egr-B, XP_018579268; *Apis mellifera* Egr-B (*Am*Egr variant III), ANS58852; *Aplysia californica* Egr-1-like, NP_001268725; *Calliphora vicina* Stripe-B, AAZ95459; *D. melanogaster* Stripe-B, NP_732289; *Tribolium castaneum* Egr, XP_0158379681; *Mus musculus* Egr-1, NP_031939. The N-terminal sequences of *Homarus americanus* Egr-B, *Periplaneta americana* Egr-B, and *Procambarus clarkii* Egr-B were deduced from following transcriptome shotgun assembly sequences: GEBG01017003.1, GEIF01013459.1, and GBEV01045599.1, respectively.

Figure 1-5. Molecular cloning and expression characteristics of *Gryllus hr38* in the cricket brain

(A) Comparison of the amino acid sequences of *Gryllus* hr38 deduced from its partial cDNA and its corresponding part of hr38 homologs in other insects. The conserved residues are marked with asterisks

above the alignments. The amino acid residues are represented in the default color scheme of ClustalX. The domain structure of *Drosophila* DHR38-D is represented above the alignment. GeneBank IDs of proteins are following: *D. melanogaster* DHR38 isoform D (DHR38-D), NP_001163024; *Bombyx mori* HR38, P49870.1; *Tribolium castaneum* HR38, XP_008194320.1; *Apis mellifera* HR38, XP_016773251.1; *Zootermopsis nevadensis* HR38, KDR09534.1. (B, C) Expression of *Gryllus hr38* after PTX injection in the brains of (B) cycloheximide and (C) saline pre-treated crickets. Asterisks indicate statistical significance to control (0 min after PTX injection). In both pre-treatment groups, *Gryllus hr38* reached a maximum 60-90 min after PTX injection (~60-fold and ~15-fold up-regulation relative to the naïve animals and to the control (0 min after PTX injection), respectively) and remained at a high level 120 min after PTX injection. The expression kinetics of *Gryllus hr38* was not affected by blockade of *de novo* protein synthesis (two-way ANOVA, Effect of pre-treatment: $F(1,98)=0.2142$, $p=0.6445$; Effect of time: $F(6,98)=14.45$, $p<0.0001$; interaction: $F(6,98)=0.5547$, $p=0.7652$). (D) Expression time-course of the pre-mRNA of *Gryllus hr38* in the brain of cycloheximide and saline pre-treated crickets after PTX-injection. Asterisks and daggers indicate statistical significance to control (0 min after PTX injection) within each pre-treatment group (***, $p<0.001$; ††††, $p<0.0001$). In both pre-treatment groups, the expression of *Gryllus hr38* pre-mRNA was significantly elevated 30 min after PTX injection (200~300-fold up-regulation relative to the naïve animals), and rapidly decreased to near baseline level by 60-90 min after injection. The expression kinetics of *Gryllus hr38* pre-mRNA was not affected by blockade of *de novo* protein synthesis (two-way ANOVA, Effect of pre-treatment: $F(1,98)=0.5631$, $p=0.4548$; Effect of time: $F(6,98)=16.11$, $p<0.0001$; interaction: $F(6,98)=1.733$, $p=0.1212$). (E) Correlation plot between the expression levels of *Gryllus hr38* and *Gryllus egr-B* in the brains of cycloheximide pre-treated crickets. The data from the cycloheximide and PTX-injected crickets (n=48; black circles), cycloheximide pre-treated crickets (n=8; gray circles), and naïve crickets (n=8; white circles) were plotted. (F, G) Behaviorally evoked expression of *Gryllus hr38* in the brain of crickets 1 hr after (F) sucrose feeding and (G) agonistic interaction. The expression levels were

normalized to the mean of those of the naïve animals (baseline expression level). An asterisk indicates statistical significance between the indicated groups. RT-qPCR analyses were performed on eight biological replicates. Box plots indicate the 25-75 percentile ranges and central values. Error bars indicate the 5-95 percentile ranges. The "+" denotes the mean. Asterisks donate statistical significance to the control (0 min after PTX injection) (B, C) or to the naïve animals (F, G) (*, $p<0.05$; **, $p<0.01$; ***, $p<0.001$; ****, $p<0.0001$). See Table 3 for the details of statistical analysis.

Figure 2-1. Expression characteristics of bZip transcription factor genes in the cricket brain.

Expression characteristics of (A-D) *Gryllus fra-A*, (E-H) *fra-B*, and (I-K) *jra*. (A, E, I) Expression time-course of *Gryllus fra-A*, *fra-B*, and *jra* after PTX-injection in the brains of cycloheximide pre-treated crickets. The expression levels were normalized to the mean of those of the naïve animals (baseline expression level). The expressions of *Gryllus fra-A* and *fra-B* reached a maximum 60 min after PTX injection (*fra-A*, ~3.5-fold and ~2.5-fold up-regulation relative to the naïve animals and to the control (0 min after PTX injection), respectively; *fra-B*, ~2-fold up-regulation relative to both the naïve animals and to the control (0 min after PTX injection)), whereas that of *Gryllus jra* reached a maximum 45 min after PTX injection (~6-fold and ~2-fold up-regulation relative to the naïve animals and to the control (0 min after PTX injection), respectively). (B, F, J) Correlation plot between the expression levels of *Gryllus fra-A*, *fra-B*, and *jra* with *Gryllus egr-B* in the brains of cycloheximide pre-treated crickets. The data from the cycloheximide and PTX-injected crickets (n=48; yellow circles) and naïve crickets (n=8; black circles) were plotted. Behaviorally evoked expression of *Gryllus fra-A*, *fra-B*, and *jra* in the brain of crickets 1 hr after (C, G, K) sucrose feeding and (D, H, L) agonistic interaction. The expression levels were normalized to the mean of those of the naïve animals (baseline expression level). RT-qPCR analyses were performed on eight biological replicates. Box plots indicate the 25-75 percentile ranges and central values. Error bars indicate the 5-95 percentile ranges. The "+" denotes the

mean. Asterisks donate statistical significance to the control (0 min after PTX injection) (A, E, I) or to the naïve animals (C, D, G, H, K, L) (*, $p<0.05$; **, $p<0.01$). See Table 3 for the details of statistical analysis.

Figure 2-2. Pharmacologically-induced expression of neuronal IEGs in the cricket brain.

Pharmacologically induced expression of (A) *Gryllus fra-A*, (B) *fra-B*, (C) *jra,* (D) *egr-B*, and (E) *hr38* in the brain of crickets. Following activators were used to stimulate intracellular signaling pathways: forskolin, an activator for adenylyl cyclases; 12-O-Tetradecanoylphorbol-13-acetate (TPA), an activator for protein kinase C; anisomycin, an activator for c-Jun N-terminal kinases; S-nitroso-N-acetylpenicillamine (SNAP), a nitric oxide donor; and a calcium-selective ionophore A23187. PTX was used as a positive control. The expression levels were normalized to the mean of those of the naïve animals (baseline expression level). Injection of 200 µM TPA resulted in significant increases in the expression levels of all neuronal IGEs (A-D). Injection of 200 µM forskolin and 200 µM A23187 resulted in significant increases in the expression levels of *fra-A* and *jra*, respectively. No obvious change was observed with any drug other than TPA in the expression levels of *Gryllus fra-B* and *egr-B* (B and D). The expression level of *Gryllus hr38* was drastically affected by stimulation of intracellular signaling pathways (E). RT-qPCR analyses were performed on eight or sixteen biological replicates. Box plots indicate the 25-75 percentile ranges and central values. Error bars indicate the 5-95 percentile ranges. The "+" denotes the mean. Letters above the plots (a and b) indicate statistical significance ($p<0.05$) to the naïve control and vehicle control, respectively. See Table 3 for the details of statistical analysis.

Figure 3-1. Sequence comparison of the DNA-binding domain of stimulus-regulated transcription

factors.

The amino acid sequence of the DNA-binding domain of stimulus-regulated transcription factors of mouse and several insect species were aligned. (A) Fos family, (B) Jun family, (C) ATF2, (D) ATF3, (E) ATF4/5, (F) ATF6, (G) large Maf family, (H) small Maf family, (I) CREB1 family, (J) insect CREB-B family, (K) C/EBPs except for C/EBP□ and C/EBPζ, (L) C/EBP□, (M) Egr family, (N) NFAT family, (O) MEF2 family, and (P) SRF. The residues for protein-DNA interaction were highly conserved in Jun family proteins, large/small Maf family proteins, CREB-like proteins, C/EBPs except for C/EBP□, MEF2 and SRF (see A, G-K, and O-P). One or few substations were detected in Fos family proteins, Egr and NFAT homologs (see A, M, and N). Extensive substitutions were found in the DBD of ATF4/5, ATF6 and C/EBP□ (see E and F). A lineage-specific occurrence of substitutions in the DBD of insect ATF2 homologs (see C). That is, extensive substitutions were detected in ATF2 homologs of *Drosophila*, *Apis* and *Tribolium* ('advanced' holometabolous insects), whereas the amino acid sequence of the ATF2 DBD is conserved between mouse and *Zootermopsis* (a 'basal' hemimetabolous insect). The conserved residues are marked with asterisks above the alignments. Residues important for nucleotide binding were indicated by black circles under the alignments. Red circles indicate the residues important for nucleotide binding where amino acid substitutions were found in most insect homologs. Residues important for protein-protein interaction (i.e. dimerization) were indicated by white circles under the alignments. The amino acid residues are represented in the default color scheme of ClustalX. GenBank IDs of proteins used for sequence comparison are listed in Table 3-1.

Table 3-1. List of stimulus-regulated transcription factors (TFs) in mammals and their homologs in insects.

Figure 7-1. Molecular cloning of *Gryllus Tdc* genes.

(A) Octopamine biosynthesis pathway in the insect brain. Tdc catalyzes the first step of octopamine biosynthesis by converting L-tyrosine into tyramine. (B and C) Insect Tdc proteins contain subtype-specific C-terminal extensions. Comparison of the amino acid sequences of the C-terminal extensions of (B) *Gryllus* Tdc1 and (C) Tdc2 with their corresponding parts of Tdc homologs in other insects. The conserved residues are marked with asterisks above the alignments. The amino acid residues are represented in the default color scheme of ClustalX. Positions of the C-terminal portion of the catalytic domains were indicated by bars under the alignment. GeneBank IDs of proteins are following: *D. melanogaster* Tdc1, NP_610226; *D. melanogaster* Tdc2, NP_724489; *A. aegypti* Tdc1, XP_001656851; *A. aegypti* Tdc2, XP_001656857; *T. castaneum* Tdc1, XP_972728; *T. castaneum* Tdc2, XP_972688.

Figure 7-2. Distribution of the Tdc2-expressing neurons in the cricket brain.

(A) Frontal view of the supraesophageal ganglion and (B) ventral view of the suboesophageal ganglion stained with anti-*Gryllus* Tdc2 antibody. The outlines of the ganglia are indicated by the white dotted lines. The depth of the cells is color coded as indicated in the inset. Dorsoventral (D-V) or rostrocaudal (R-C) axes were indicated. Scale bars represent 200 μm.

Movie 1. EYFP immunoreactivity in the optical sections of the brain of IEG reporter line 6 hr after vehicle injection.

Movie 2. EYFP immunoreactivity in the optical sections of the brain of IEG reporter line 6 hr after PTX injection stained with anti-GFP antibody.

Table 1. Nucleotide sequences of primers used in this study.

Degenerate primers		
	forward primer	reverse primer
egr	5'-GGA GTN CAR CTN GCH GAR TA-3'	5'-GAN CGC ATG CAD ATN CGR CAY TG-3'
	5'-ACS AGN AAR GGN CAY GAR AT-3'	5'-TTN AGR TGN ACY TTN GCR TG-3'
	5'-CCA RTG YCG NAT HTG CAT GCG-3'	
hr38	5'-AAC CGC TGC CAR TTY TGC-3'	5'-AGA AGC TCC TGR TCR TNG C-3'
tdc2	5'-TCG AGT ACG CSG AYT CKT TCA ACA C-3'	5'-GGA TCR CTS ACC ATN CGN ACG AAG AA-3'
Primers for full-length ORF amplification		
	forward primer	reverse primer
fra-A isoform	5'-CGC GGG AGT AAG GAC GTG-3'	5'-CCC CAT TGT CCA AAT CCT CC-3'
fra-B isoform	5'-GGC GGC TTG TGT GTT TGT G-3'	5'-CCC CAT TGT CCA AAT CCT CC-3'
jra	5'-GAC GGT CGC GGA GAG TC-3'	5'-GAT CTC ATA TGT ATA TGC ATG TGT TCA C-3'
egr-B	5'-TTC ATT CAT AAA AGT GTT GTA GAG CG-3'	5'-ATA TAT ACG AAT CGA GGA GAA CAC-3'

tdc1	5'-CAT CTG GCG TTC GCT C-3'	5'-CGC AGT CCC AGA AGA G-3'
tdc2	5'-CGA CGC CCG ACG ACA TTC G-3'	5'-CCG GCT CGT ATG TTG TGT GG-3'
Primers for RT-qPCR		
	forward primer	reverse primer
fra total transcripts	5'-GGA CGG CCT CAA TTC GGG-3'	5'-GGA TTC CAC CTC GCA CTG C-3'
fra-A	5'-CCT GCC TTC ATC TGC GTA CG-3'	5'-GTC TCA CTG GGC GAA ACG TG-3'
fra-B	5'-GGC GGC TTG TGT GTT TGT G-3'	5'-GGA TTC CAC CTC GCA CTG C-3'
jra	5'-GAG CGG ACG GTT GTG TTA GG-3'	5'-GCA GTT GCG TAC CAT CTA AAT CC-3'
egr-B (for initial expression analysis)	5'-GAC CTA GGC GTC GAA CCC-3'	5'-GTT CCA AGG ATC CTG TGA TGG G-3'
egr-B	5'-GTT TGG AAA CGC TGA GCC C-3'	5'-CCT GAC GCT GTA GAG GCA C-3'
egr pre-mRNA	5'-GTG ACA CAT GTA ATT GGC GTA AC-3'	5'-CAA TTC CTC GGG TTC CAA GG-3'
hr38	5'-CCA ACC TCG ACT ATT	5'-CCG GAA TCT TAT CAG CAA

	CAC AGT ATC-3’	ACG TG-3’
hr38 pre-mRNA	5’-GAA GCA TCT ACT CCA GTC TCA TAA TAG-3’	5’-GTA GGC TCA CGA TAC TGG AAA TG-3’
EYFPnls:PEST	5’-CGA GGA GCT GTT CAC CGG-3’	5’-GGT GCA GAT GAA CTT CAG GG-3’
β-actin	5’-CGT AAA CTC AAC TAC TAA CCA TGT GC-3’	5’-GCC CTG GGT GCA TCA TCG-3’
ef1α	5’-CGA CTC CGG TAA ATC TAC GAC C-3’	5’-CAC CCA GGC ATA CTT GAA AGA AC-3’
rpl32	5’-CGC CCA GTT TAT CGT CCA AC-3’	5’-GCC TGC GAA CTC TGT TGT C-3’
Primers used to amplify the core promoter regions of orthopteran *egr-B* homologs		
	forward primer	reverse primer
	5’-GTT ACG TCA TTT GAC GT CA-3’	5’-GTC CCA TAT TTG GAA GTC G-3’
	5’-GTT ACG TCA TTT TGA CGT CA-3’	5’-CAM CAS TTT TAT GAA TGA AG-3’
Primers used to amplify the genomic DNA fragment upstream to the coding sequence of the orthopteran *egr-B* homologs		
	forward primer	reverse primer
Gryllus bimaculatus	5’-CAG GGG TTG TTT ATT CGC CG-3’	5’-CTG TGA TGG GAG GCG GTT CAA C-3’

Acheta domesticus	5'-AAA TTC GAA AGC CTT GAC AGT GG-3'	5'-ACG ATG GAC GAG CGT CGT G-3'
Gampsocleis buergeri	5'-ATG TTC CCC CTC CAT GCC AG-3'	5'-ACA TGC TGA CGC GCA ACA C-3'
Locusta migratoria	5'-CAG TGT TGC CAG CCT CC-3'	5'-CCG ACG AGT ACA GGC AGT C-3'

Table 2. Genomic sequences used for promoter analysis.

Species	GenBank ID	Genomic region targeted for TFBS prediction	Position of the conserved core promoter	Position of the +1 site	Position of CRM_{-800}	Position of CRM_{-400}
Drosophila melanogaster	NT_033777.2	base 13945983 to 13948265	-	base 13947783	-	-
Apis mellifera	NC_007084.3	base 6690515 to 6692794	-	base 6691064**	-	-
Tribolium castaneum	NC_007417.3	base 2995987 to 2998141	base 2997452 to 2997620 (169 bp)	-	-	-
Acyrthosiphon pisum	ABLF02030506.1	base 25286 to 24071	base 25155 to 24964 (192 bp)	-	-	-
Pediculus humanus	NW_002987224.1	base 14669 to 12120	base 12809 to 12649 (161 bp)	-	-	-
Blattella germanica	JPZV01078734.1	base 44,303 to 41,904	base 42626 to 42446 (181 bp)	-	base 43369 to 43205 (165 bp)	base 42891 to 42809 (83 bp)
Zootermopsis nevadensis	AUST01012629.1	base 2539 to 150	base 954 to 753 (197 bp)	-	base 1667 to 1493 (83 bp)	base 1248 to 1170 (79 bp)
Gryllus bimaculatus	LC215247*	base 1 to 2574	base 1518 to 1711 (194 bp)	base 1668**	base 786 to 937 (152 bp)	base 1235 to 1316 (82 bp)
Acheta domesticus	LC215248*	base 1 to 2781	base 2113 to 2306 (194 bp)		base 1358 to 1509 (152 bp)	base 1830 to 1911 (82 bp)
Gampsocleis buergeri	LC215249*	base 1 to 2098	base 1223 to 1413 (191 bp)		base 498 to 652 (155 bp)	base 945 to 1024 (80 bp)
Locusta migratoria	LC215250*	base 1 to 1909	base 1166 to 1404 (239 bp)		-	-
Daphnia pulex	ACJG01000376.1	base 1585904 to 1583495	base 1584357 to 1584191 (167 bp)		-	-

* Genomic sequences were determined in this study.

** Positions of the TSS(s) were determined in this study.

Table 3. Experimental conditions and statistics.

Figure	Experimental conditions	Statistical test
1A-F	Adult male crickets one week after the imaginal molt were isolated for 3 days. Crickets were injected with 20 mM cycloheximide. One hour later, 5 mM PTX or saline were injected to the crickets. Thirty minuets later, brains were dissected for RNA extraction (n=8 in each group).	*Gryllus fra* total saline vs *Gryllus fra* total cycloheximide, U=23; *Gryllus fra*-A saline vs *Gryllus fra*-A cycloheximide, U=24, *Gryllus fra*-B saline vs *Gryllus fra*-B cycloheximide, U=15; *Gryllus jra* saline vs *Gryllus jra* cycloheximide, U=32; *Gryllus egr-B* saline vs *Gryllus egr-B* cycloheximide, U=11; *Gryllus hr38* saline vs *Gryllus hr38* cycloheximide, U=29, Mann-Whitney U test.
2A, B	Adult male crickets one week after the imaginal molt were isolated for 3 days. Crickets were injected with (A) 20 mM cycloheximide or (B) saline. One hour later, 5 mM PTX was injected to the crickets. Brains were dissected for RNA extraction before PTX injection (0 min), or 15, 30, 45, 60, 90, or 120 min after PTX injection (n=8 in each group).	Effect of pre-treatment: $F(1,98)=0.9604$, $p=0.9004$; Effect of time: $F(6,98)=9.034$, $p<0.0001$; interaction: $F(6,98)=0.3634$, $p=0.9004$, Two-way ANOVA, Dunnett's post-hoc test.
2C	Same as in Figure 2A.	$H=29.21$, $p<0.0001$, Kruskal-Wallis test, Dunn's post-hoc test.
2D	Adult male crickets one week after the imaginal molt were isolated for 3 days. Cricket were fed with 20% sucrose solution 3 times with 5 min intervals. Brains were dissected for RNA extraction before feeding (naïve), or 1 hr after feeding (n=8 in each group).	$t=3.051$, $df=8.491$, Welch's unpaired t-test.
2E	Adult male crickets one week after the imaginal molt were isolated for 3 days. Weight-matched crickets were introduced into an arena to interact for 5 min, then re-isolated. Brains were dissected for RNA extraction before interaction (naïve), or 1 hr after start of interaction (n=8 in each group).	$H=12.62$, $p<0.01$, Kruskal-Wallis test, Dunn's post-hoc test.
5C	Adult male crickets one week after the imaginal molt were isolated for 3 days. Then, brains were dissected for RNA extraction (n=8 in each group).	$t=15.95$ $df=7.094$, Welch's unpaired t-test.

5D	Adult male crickets one week after the imaginal molt were isolated for 3 days. 5 mM PTX was injected to the crickets. Brains were dissected for RNA extraction before PTX injection (0 min), or 15, 30, 45, 60, 90, or 120 min after PTX injection (n=8 in each group).	H=25.87, p<0.001, Kruskal-Wallis test, Dunn's post-hoc test.
5E	Same as in Figure 5C.	H=39.42, p<0.0001, Kruskal-Wallis test, Dunn's post-hoc test.
5F	Pooled expression data of *EYFPnls:PEST* and *Gryllus egr-B* in the brain within 120 min after PTX injection (including data presented in Figures 5C, D), 60 min after saline injection, and naïve crickets were analyzed (n=72, 8, and 8, respectively).	r=0.8269, p<0.0001, Pearson's correlation analysis.
2-1A, E, I	Same as in Figure 2A.	*Gryllus fra-A*, F(6,49)=3.529, p<0.001; *Gryllus fra-B,* F(6,49)=3.327, p<0.01; *Gryllus jra*, F(6,49)=5.305, p<0.001, One-way ANOVA, Dunnett's post-hoc test.
2-1B, F, J	Pooled expression data of *Gryllus fra-A, fra-B, jra* and *egr-B* in the brain of cycloheximide pre-treated crickets within 120 min after PTX injection (including data presented in Figures 2A, D and G) and naïve crickets were analyzed (n=48 and 8, respectively).	*Gryllus fra-A* vs *egr-B,* r=0.7856, p<0.0001; *Gryllus fra-B* vs *egr-B,* r=0.5250, p<0.0001*; Gryllus jra* vs *egr-B*, r=0.6222, p<0.0001, Pearson's correlation analysis.
2-1C, G, K	Same as in Figure 2D.	*Gryllus fra-A*, U=13; *Gryllus fra-B, U*=7; *Gryllus jra*, U=22, Mann-Whitney U test.
2-1D, H, L	Same as in Figure 2E.	*Gryllus fra-A*, H=10.24, p<0.05; *Gryllus fra-B, H*=13.09, p<0.01; *Gryllus jar*, H=2.284, p=0.5156, Kruskal-Wallis test, Dunn's post-hoc test.
1-5B, C	Adult male crickets one week after the imaginal molt were isolated for 3 days. Crickets were injected with (B) 20 mM cycloheximide or (C) saline. One hour later, 5 mM PTX was injected to the crickets. Brains were dissected for RNA extraction before PTX injection (0 min), or 15, 30, 45, 60, 90, or 120 min after PTX injection (n=8 in each group).	Effect of pre-treatment: F(1,98)=2.060, p=0.1544; Effect of time: F(6,98)=14.57, p<0.0001; interaction: F(6,98)=0.7991, p=0.5729, Two-way ANOVA, Šidák's post-hoc test.

1-5D	Same as in Figures 2A, B.	Effect of pre-treatment: $F(1,98)=0.5631$, $p=0.4548$; Effect of time: $F(6,98)=16.11$, $p<0.0001$; interaction: $F(6,98)=1.733$, $p=0.1212$, Two-way ANOVA, Šidák's post-hoc test.
1-5E	Pooled expression data of *Gryllus hr38* and *Gryllus egr-B* in the brain of saline and cycloheximide pre-treated crickets within 120 min after PTX injection (including data presented in Figures 2A, B and S6B, C), and naïve crickets were analyzed (n=48, 48, and 8, respectively).	Saline pre-treated crickets: $r=0.8498$, $p<0.0001$; cycloheximide pre-treated crickets: $r=0.8446$; $p<0.0001$; all data included: $r=0.8498$, $p<0.0001$, Pearson's correlation analysis
1-5F	Same as in Figure 2D.	$t=2.501$, $df=13.18$, Welch's unpaired t-test.
1-5G	Same as in Figure 2E.	$H=17.61$, $p<0.001$, Kruskal-Wallis test, Dunn's post-hoc test.
2-2	Adult male crickets one week after the imaginal molt were isolated for 3 days. Crickets were received injection of PTX, forskolin, TPA, SNAP, aniosmycin, A23187, or saline. Brains were dissected for RNA extraction 60 min after injection (PTX, n=8; the other treatments, n=16).	*Gryllus fra-A*, $H=62.94$, $p<0.0001$; *Gryllus fra-B,* $H=40.22$, $p<0.0001$; *Gryllus jar*, $H=56.30$, $p<0.0001$; *Gryllus egr-B,* $H=39.41$, $p<0.0001$; *Gryllus hr38,* $H=67.07$, $p<0.0001$, Kruskal-Wallis test, Dunn's post-hoc test.