Identification of a reproductive-specific, putative lipid transport protein gene in a queenless ponerine ant *Diacamma* sp.

Yasukazu Okada[1,2], Satoshi Miyazaki[1,3], Shigeyuki Koshikawa[1,4], Richard Cornette[1,5], Kiyoto Maekawa[6], Kazuki Tsuji[7], Toru Miura[1,*]

[1]Laboratory of Ecological Genetics, Graduate School of Environmental Science, Hokkaido University, Sapporo 060-0810, Japan

[2]Laboratory of Evolutionary Ecology, Graduate School of Environmental Science, Okayama University, Okayama 700-8530, Japan

[3]Laboratory of Parasitology, Graduate School of Veterinary Medicine, Hokkaido University, Sapporo 060-0818, Japan

[4]Laboratory of Molecular Biology, University of Wisconsin-Madison, 1525 Linden Drive, Madison, Wisconsin 53706, USA

[5]Anhydrobiosis Research Unit, National Institute of Agrobiological Sciences, 1-2 Owashi, Tsukuba, Ibaraki 305-8634, Japan

[6]Graduate School of Science and Engineering, University of Toyama, Toyama 930-8555, Japan

[7]Faculty of Agriculture, University of Ryukyus, Naha 903-0213, Japan

[*]Correspondence: Toru Miura, Laboratory of Ecological Genetics, Graduate School of Environmental Science, Hokkaido University, Sapporo 060-0810, Japan. Tel & Fax: +81-11-706-4524, E-mail address: miu@ees.hokudai.ac.jp

Abstract

Of the various characteristics of social insects, communication for reproductive differentiation is one of the most important and basic social interactions among colony members. To elucidate the molecular basis underlying this process, genes responsible for reproductive differentiation in *Diacamma* were screened using fluorescent differential display. Differential display, together with real-time quantitative RT-PCR, revealed that a gene belonging to the family of cellular retinaldehyde-binding proteins was specifically expressed in the epidermis of the head, legs, and thorax in reproductives. The deduced protein sequence in the coding region, obtained by RACE PCR, was found to include CRAL-TRIO domain, suggesting that *DiaCRALDCP* functions in transportation of lipids, such as cuticular hydrocarbons. *DiaCRALDCP* transcript levels immediately decreased 1 day after the gemma mutilation, suggesting that *DiaCRALDCP* is involved in the physiological changes provoked by the behavioral regulation. Considering these results, the social functions of *DiaCRALDCP* in *Diacamma* are discussed.

Key words: reproductive differentiation, social communication, queenless ponerine ant, cellular retinaldehyde-binding protein (CRALBP)

Introduction

Colony organization of social insects is accomplished by behavioral regulation through various communications (Wilson 1971). Of the numerous social interactions among colony members, the caste differentiation between reproductives (queens) and non-reproductives (workers) is a topic of extensive research in the field of sociobiology. The recent advances in molecular biology and genomic analyses of various organisms, including the eusocial honeybee *Apis mellifera*, have enabled the clarification of the molecular bases of social behaviors and communications (Robinson et al. 2005). Therefore, the molecular basis of reproductive differentiation is now becoming an important target of evolutionary studies focusing on the molecular and physiological bases of social interactions in eusocial species.

The most primitive types of reproductive division of labor are observed in species that lack a morphological queen (i.e., primitively eusocial wasps and bees, queenless ponerine ants). In such species, social systems are mostly organized by dominance interactions that trigger physiological differentiation among adult individuals without any morphological specializations (West-Eberhard 1979; Turillazzi and West-Eberhard 1996; Gadagkar 1990). Such a system of behavioral differentiation represents the ancestral or basal organization of highly eusocial systems. Despite the evolutionary significance, however, the molecular basis of behavioral differentiation in such species has yet to be elucidated (Sumner et al. 2006; Toth et al. 2007; Pereboom et al. 2005).

In the queenless ponerine ant of the genus *Diacamma*, reproductive differentiation occurs via a highly specialized dominance interaction. All eclosed females (callow females) of *Diacamma* have the potential to become reproductives, but only 1 female possessing "gemmae," i.e., the vestigial wings specialized for the determination of reproductives (Fukumoto et al. 1989; Peeters and Higashi 1989; Gotoh et. al. 2005), becomes a dominant egg-layer called a "gamergate" (functional queen in queenless ants; Peeters and Crewe 1985; Peeters 1991), whereas the rest of the females, whose gemmae are mutilated by the gamergate, become

subordinate sterile workers (**Fig. 1**; Fukumoto et al. 1989; Peeters and Higashi 1989). We previously showed that the ovaries of subordinate sterile workers start to differentiate immediately after gemma mutilation (Okada et al. 2010). Furthermore, the cuticular hydrocarbon (CHC) profile of female adults also is known to play an important role in the social organization of *Diacamma*. Gamergates, workers, and eclosed callows possess specific CHC profiles that signal their social status (Cuvillier-Hot et al. 2001; 2002). In general, in most species of primitively-eusocial wasps and queenless ponerine ants, it is difficult to focus the process of reproductive differentiation, because the dominance interaction among females and the subsequent replacement of reproductives occur unpredictably.

The unique system of reproductive differentiation and social communication in *Diacamma* allows us to investigate the process of reproductive differentiation. In this study, a gene-screening method "differential display," was used to identify the specific gene expressions at the time of reproductive differentiation. For the analysis, we focused on the head part of the ant, because it contains the brain and several exocrine glands (e.g., postpharyngeal gland, propharyngeal gland, mandibular gland) (Billen 1998; Hölldobler and Wilson 1990), both of which were thought to be important for social communications.

Materials and Methods

Animal collection

Complete colonies of *Diacamma* sp. (the only species of *Diacamma* in Japan) were collected in Nakijin, Onna, and Naha, located on Okinawa Island, Japan. Gamergate-right colonies containing 50–200 workers were reared in plastic artificial nests (7.7 x 10.8 x 3.2 cm) filled with moistened plaster. The artificial nests were placed in plastic arenas. To obtain newly differentiated gamergates, orphan colonies without gamergates were induced by isolating 12–15 mutilated workers in a small artificial nest (4.8 x 6.2 x 2.5 cm). The nests were maintained at 23°C under a 12L/12D photoperiod and fed with chopped mealworms (*Tenebrio molitor* larvae) 3 times a week. Honey water was supplied *ad libitum*.

Induction of reproductives and sterile workers

A newly eclosed female before interactions with nestmates, defined as an "intact callow (IC)," can differentiate into either an egg-layer or a sterile worker depending on colony status. ICs were obtained by isolating female pupae. In an orphan colony, a newly eclosed female retains gemmae and becomes an unmated prospective egg-layer, which is defined as a future gamergate (FG) (**Fig. 1**; Fukumoto et al. 1989; Peeters and Higashi 1989). The induced FGs start to lay eggs within 14 days after eclosion (Okada et al. 2010). Mutilated females, which became sterile workers, of the following age classes were collected from gamergate-right colonies: day-1, day-7, day-30, and day-60. 'Day-60' workers included workers aged 60 days or older. Future gamergates of 3 age-categories were induced: day-1, day-7, and day-30. Gamergates aged 60 days or older were collected from stock colonies.

In orphan colonies, only one highly ranked mutilated female (dominant worker) can lay male-destined unfertilized haploid eggs, whereas the rest of the females are kept sterile (subordinate workers; Peeters and Tsuji 1993). We obtained dominant and subordinate workers by producing

orphan colonies consisting of 12–15 mutilated workers. These orphan colonies were kept for 2 weeks or more, and then all the workers were dissected to examine their ovarian development to classify them into dominant egg-layers or subordinate workers (Peeters and Tsuji 1993).

Fluorescent differential display

Differential display is a gene screening method that allows the detection of differentially or specifically expressed genes among multiple sample categories (Liang and Pardee 1992). In this study, fluorescent differential display (FDD; Bauer et al. 1993; Ito et al. 1994) was performed to isolate genes expressed in the head during reproductive differentiation. The heads of 6–10 ants from multiple colonies (more than five) were lumped for total RNA extraction. Samples were obtained from following 5 categories: IC, day-1 FG (1FG), day-1 worker (1W), day-7 FG (7FG), and day-7 worker (7W). RNA extraction was performed with an SV total RNA isolation system (Promega, Madison, WI, USA). Then, 800 ng of total RNA was reverse transcribed with Super Script III reverse transcriptase (Invitrogen), using 50 pmol of Bam-TG primer (5'-CCCGGATCCT$_{15}$G-3'; Linskens et al. 1995). The resultant cDNA was used for PCR amplifications with 20 combinations of arbitrary 10-mers with a *Hin*dIII site (Linskens et al. 1995) and FITC-labeled Bam-TG primer. PCR reaction mixtures (20 µL) contained 1 µL of the RT products, 2 µL of 10 x PCR buffer, 2 mmol/L dNTP, 0.1 µL AmpliTaq Gold DNA polymerase (5 units/µL; Applied Biosystems, Foster, CA, USA), and 8 pmol of each primer. PCR conditions were as follows: 1 cycle of 94°C for 10 min, 37°C for 5 min, and 72°C for 5 min; then 35 cycles of 94°C for 30 s, 55°C for 1 min, and 72°C for 1 min; and finally, extension at 72°C for 5 min. PCR products were separated on a denaturing 6% polyacrylamide gel and the resultant electrophoresis profile was scanned by FluorImager (GE healthcare, Tokyo, Japan). To confirm that the electrophoresis profile was reproducible, RT–PCR and electrophoresis were duplicated.

Subcloning and sequencing

The bands specific to FGs or workers were excised and boiled at 95°C in 100 µL distilled water to extract DNA from gels. The extracted DNA was re-amplified by PCR with the *Hin*dIII and *Bam*-TG primers under the following conditions: 94°C for 1 min; 94°C for 30 s, 55°C for 30 s, and 72°C for 3 min for 40 cycles; and 72°C for 5 min. The reaction mixtures (100 µL) contained 1 µL of template DNA, 10 µL of 10 x PCR buffer, 20 mmol/L dNTP, 1 µL cloned *Pfu* DNA polymerase (2.5 units/µL; Stratagene, La Jolla, CA, USA), and 50 pmol of each primer. PCR products were purified with a Wizard SV Gel and PCR Clean-up System (Promega), and ligated in pGEM 3Zf(+) Vectors (Promega) at the *Sma*I site by means of the TaKaRa DNA Ligation Kit version 2 (Takara Bio, Otsu, Shiga, Japan). Ligated plasmids were transfected into *Escherichia coli* JM 109 (Takara Bio) and cultured at 37°C for 16 h, then purified with a Sigma Gene Elute Plasmid Miniprep Kit (Sigma-Aldrich, St Louis, MO, USA). The nucleotide sequences of insert cDNA fragments were determined by the Big Dye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster city, CA, USA) and ABI Prism 3100 Genetic Analyzer (Applied Biosystems). The determined nucleotide sequences were subjected to database search for homologues using the National Center for Biotechnology Information's (NCBI) BLASTX.

Real-time quantitative RT-PCR

To confirm specific gene expressions, real-time quantitative RT-PCR was performed for 9 identified gene candidates. cDNA was synthesized from the head-derived RNA as described in the former section, using Super Script III transcriptase (Invitrogen) and oligo-dT primer. The genes that exhibited more than 2 fold changes in transcript level were judged as reproductive-role specific genes. For further investigation of tissue- and reproductive-role specificity, another series of RNA and cDNA samples were prepared using 6-10 individuals from multiple colonies (more than five) in following experiments. In order to examine gene expression in the brain, RNA was extracted from the brain (including the subesophageal ganglion) and the

remaining part of the head, i.e., the “head excluding brain” (HEB) of ICs, and subjected to real-time qRT-PCR. Since the gene *DiaCRALDCP* was detected exclusively in the HEB (see Results), the RNA from the HEB in various social statuses (ICs, workers, and FGs of day-1, day-7, day-30, and day-60, dominant egg-laying workers, subordinate workers, and gamergates) was subjected to qRT-PCR to confirm the reproductive-specificity. Furthermore, in order to identify the detailed localization of the transcript in the HEB, we performed the combined analysis of tissue dissection and real-time qRT-PCR. The HEB of ICs was divided into the propharyngeal gland, postpharyngeal gland, eye, pharynx, “cuticle and epidermis,” and “muscle and tracheae.” Fat bodies were excluded from the analysis, because they were generally scarce. The whole ant body was also dissected into head, thorax (including petiole), gaster, antennae, and legs for transcript quantifications. Reverse transcription was performed as described above, except that the micro scale RNA extraction kit RNAqueous micro (Ambion) was used for the small tissues (propharyngeal gland, postpharyngeal gland, eye, pharynx, “cuticle and epidermis,” and “muscle and tracheae”).

Relative quantification of cDNAs was performed using SYBR Green PCR Master Mix and Power SYBR Green PCR Master Mix (Applied Biosystems) with sequence detection system ABI PRISM 7000 (Applied Biosystems). For relative quantification of the targeted transcripts, cycle threshold (Ct) values were calculated by the relative standard curve method. For the reference for gene expression, actin gene (Accession No. AB510469) whose relative stability has been confirmed in various tissues and stages in honeybee (Lourenco et al. 2008, Scharlaken et al. 2008), was applied for normalization. Because of the large number of sample categories required in the series of experiments, only one replicate of qRT-PCR per category was performed. Reactions were technically triplicated to ensure the accuracy of amplification. Primers were designed using the Primer Express software (ver. 2.0.0) (Applied Biosystems). Primer sequences used in qRT-PCR were available in Supplementary Table 1.

Rapid amplification of cDNA ends (RACE)

In order to obtain the 5' and 3' ends of the caste-specific gene sequence obtained by FDD (see Results), 5' and 3' RACE PCR was performed using the SMART RACE amplification kit (Clontech Laboratories, Palo Alto, CA, USA). The 5'-RACE PCR reaction was performed with the respective internal 5'-specific primer (5'- CTT AAG CCG TAC GCC TCT AAA CAC AGG -3'). Second nested PCR was then performed with the second 5'-specific primer (5'- GAG GAC GGT TGG CGT ATA TTT TAC AAG ATG -3'). The 3' RACE PCR was also performed with the 3'-specific primer (5'- GGA AAC AAC GGG GAT GGT TGG AGG AAC AG -3'). Subcloning and sequencing of the RACE products was performed with the pGEM-T vector system, as described above.

Phylogenetic analysis and evolutionary history of *DiaCRALDCP*

Phylogenetic analyses were performed on the reproductive-specific gene obtained in FDD (*DiaCRALDCP*, see Result) under Bayesian Inference (BI), Maximum Likelihood (ML), and Neighbor Joining (NJ) criteria. Estimations of tree topology were obtained under BI using the program MrBayes 3.1 (Huelsenbeck and Ronquist 2001). The most appropriate model of sequence evolution was determined using Modeltest 3.06 (Posada and Crandall 1998), and parameters for the selected model of substitution were estimated from the data. A total of 10,000 trees were obtained (ngen = 1000000, samplefreq = 100), and the first 3,000 of these were considered as the 'burn in' and discarded. A 50% majority-rule consensus tree of the remaining trees was produced. Branch lengths for this tree were calculated in the program PAUP*4.0b10 (Swofford 2000) under ML criteria, estimating parameters from the data. Two independent runs under the same model of sequence evolution were performed. For ML and NJ, bootstrap probabilities were obtained using PAUP* [100 (ML) and 1000 (NJ) replicates, HKY + G model, parameters estimated from the data]. Synonymous (dS) and non-synonymous (dN) substitution rates in the lipid binding domain and other regions were calculated by codeml (Goldman and Yang 1994), using the

program PAL2NAL (Suyama et al. 2006).

Results

Differential display, together with real-time qRT-PCR, identified 1 reproductive-specific gene fragment that was abundant in future gamergates (supplementary Fig. 1, H11-2). RACE-PCR for this fragment successfully amplified a gene sequence of 2441 bases that contained a complete open reading frame coding a protein (base 416–1324, 303 amino acid residues). By BLASTX search, the product was found to belong to the cellular retinaldehyde-binding protein family (CRAL/TRIO domain-containing protein) that comprises carriers of lipophilic hydrocarbons. Protein structure analysis by Phyre (http://www.sbg.bio.ic.ac.uk/~phyre/index.cgi) confirmed that this protein possesses the CRAL-TRIO lipid-binding domain that characterizes the carrier function of this protein. Thus, we named this gene *DiaCRALDCP* (Accession no. AB523879). The amino acid sequence of inferred protein was closest to the *Aedes* CRAL-TRIO domain-containing protein (XP_001651428.1, E-value: e-46, 35.3% amino acid identity [107/303]) and *Anopheles* AGAP002559 (XP_312380.2, E-value: 5e-47, 36.3% amino acid identity [110/303]) (**Fig. 2;** Panagabko et al. 2003). In addition, *DiaCRALDCP* showed similarity to *Locusta* tfp-1, whose putative ligand are reported to be juvenile hormone (JH) (Zhou et al. 2006a; ABH 11711, E-value: 3e-29, 28.4% amino acid identity [86/303]). Amino acid sequence of CRALDCP was aligned with representative insect CRAL-TRIO domain-containing proteins (**Fig. 2**).

Molecular phylogeny based on DNA sequence was constructed using closely related genes selected from various insect lineages. *DiaCRALDCP* was most closely related to the *Apis melifera* ortholog, which is similar to CG2663 PB isoform B in *Drosophila melanogaster* (Fig. 3). Figure 3 shows the 50% majority-rule consensus tree obtained from Bayesian analysis, along with posterior probabilities (PP) from BI and bootstrap values (BV) from ML and NJ analyses respectively. The most appropriate model of nucleotide substitution for likelihood analyses selected by Modeltest 3.06 was HKY + G model. Synonymous (dS) and non-synonymous (dN) substitution rates between putative orthologs of *Diacamma* and *Apis* were

0.1175 (within the domain) and 0.0149 (other regions).

Among 8 other candidates obtained by differential display, one sequence named *DiaH9-9* exhibited worker-specificity at day-7 (supplementary Fig. 1). However, no coding region was detected in RACE-PCR amplicon (1494 bases, Accession No. AB523880), suggesting that *DiaH9-9* is a 3' UTR (untranslated region) or non-coding RNA. Since its gene function is uncertain, no further analysis was carried out for *DiaH9-9*. For others, since expression differences were less than two-fold (supplementary Fig. 1), they were judged as "false positives" that often occur in differential display (Zhang et al. 1998), and excluded from the further analyses. Unfortunately, we could not design appropriate primers for real-time qRT-PCR due to the small fragment sizes of two candidates, so that they were also omitted from the analysis.

Further, gene expression analysis by qRT-PCR revealed that *DiaCRALDCP* was specifically localized in the HEB (head excluding brain), suggesting that this gene functions in some head tissues other than brain (**Fig. 4a**). Greater levels of *DiaCRALDCP* transcripts were detected in the HEB of ICs and reproductives (i.e., FG, gamergate, and dominant worker) in comparison with workers at the same age (**Fig. 4b**). In reproductives, transcript levels decreased in an age-dependent manner, although day-7 FG had lower transcript level than that of day-30 FG. Furthermore, although their exact ages were unknown, dominant workers (egg-layer) had greater transcript levels than sterile subordinates, supporting the reproductive-specificity of this gene. The combined analysis of tissue dissection and qRT-PCR showed that *DiaCRALDCP* was expressed in the head epidermis (**Fig. 5a**). In addition to the HEB, *DiaCRALDCP* was strongly expressed in the thorax and legs in reproductives (**Fig. 5b**). Interestingly, transcript levels were immediately decreased 1 day after the mutilation in these body parts (**Fig. 5b arrowheads**). Because reference gene showed some variations between tissues and/or castes, analyses were repeated on unnormalized data, but the reproductive and tissue specificity was highly consistent with normalized data (Supplementary Figs. 2, 3).

Discussion

By differential display analysis, the gene *DiaCRALDCP* was identified and found to be strongly expressed in epidermal tissues of head, thorax, and legs of reproductives in *Diacamma* sp. Interestingly, the transcript levels of this gene rapidly decreased after gemma mutilation (**Fig. 5b**), suggesting that *DiaCRALDCP* is involved in the physiological changes of reproductive differentiation.

The protein structure of DiaCRALDCP indicates that it belongs to the cellular retinaldehyde-binding protein (CRALBP) family that comprises transporters of small lipophilic molecules, such as retinaldehyde, alpha-tocopherol (vitamin E) and juvenile hormone (JH) (**Fig. 3**; Panagabko et al. 2003, Zhou et al. 2006a). Molecular phylogeny based on DNA sequence indicated that *DiaCRALDCP* was closest to *Apis melifera* gene similar to CG2663. In addition, *DiaCRALDCP* was included in the clade containing *D. melanogaster* gene CG2663 and other insect orthologs, suggesting that this gene was orthologous to CRALBP family gene CG2663 (**Fig. 3**). Although dN/dS rations between DiaCRALCVP and Apis CG2663 were less than 1 implying that no positive selection was occurred in the lipid-biding domain, *DiaCRALDCP* had low sequence similarity to *Apis* and other insect orthologs (e.g. Amino acid sequence similarity between *Diacamma* and *Apis* was 34.2 % in domain, and 35.5 % in other region). DiaCRALDCP may have unique evolutionary history that may lead to the distinctive ligand affinity of this gene, although this requires further analysis. One of the well-known functions of CRALDCP is the transduction of visual substances. In *Drosophila*, a CRALBP family protein called PINTA (**Fig. 3**, blastx e-value: 1e-14) is expressed in retinal pigment cells and transfers retinol from extra-retinal neural cells to retinal pigment cells of the eye (Wang and Montell 2007). However, in our study, *DiaCRALDCP* transcript was scarce in eye tissue (**Fig. 5a**), suggesting that this gene has other functions.

In spite of the fact that a large number of genes encoding CRALDCP are found in the genomes of various insect species, the ligands for these CRALDCPs are largely unknown. Here we discuss 2 possible ligand

candidates for DiaCRALDCP. The localization of the gene to the cuticular epidermis and its immediate decline after gemma mutilation suggests that a candidate ligand of DiaCRALDCP is cuticular hydrocarbons (CHCs). The cuticular epidermis of insects contains a large number of oenocytes that synthesize CHCs (Diehl 1975; Ferveur et al. 1997; Noirot and Quennedey 1991; Fan et al. 2003), which are transferred by lipophilic carrier proteins, such as lipophorin (Lucas et al. 2004). CHC profiles are known to function as pheromones that signal individual social status, reproductive status in particular, in many ant species (Liebig et al. 2000; Cuvillier-Hot et al. 2001; Dietemann et al. 2003; Peeters and Liebig 2009). In *Diacamma*, the changes in CHC profile occur some time after gemma mutilation and therefore CHC profiles differ between egg-layers and sterile workers (Cuvillier-Hot et al. 2001). One hypothesis is that DiaCRALDCP may be involved in the alteration of CHC profiles via the transportation of CHC(s) in a caste-specific manner. The strong localization of the *DiaCRALDCP* transcript in the thorax and legs strengthens this hypothesis. Soon after callow eclosion, nestmates excitedly bite and antennate her cuticle, especially on legs and thorax, so that the nestmates surround and immobilize the callow (Fukumoto et al. 1989; Peeters and Higashi 1989; Baratte et al. 2006; Okada personal observation). This behavioral sequence finally results in the mutilation of the thoracic appendage, i.e. gemmae. Interestingly, however, nestmates lose their interest in the callow cuticle within a day after the mutilation.

Another hypothesis is that DiaCRALDCP has a binding function to JH, because JH is a sesquiterpenoid, which is similar to retinaldehyde and alpha-tocopherol (vitamin E) (Zhou et al. 2006a). In queenless ponerine ants, JH suppresses ovarian activity (Sommer et al. 1993; Cuvillier-Hot 2004). Recent studies have indicated that the sequestration of JH by JH-binding molecules, such as hexamerin, lowers the hemolymphatic JH titer and blocks JH signaling (Zhou et al. 2006b). It is possible that DiaCRALDCP binds to JH to suppress the JH signaling in reproductives, while the low expression of *DiaCRALDCP* in workers raises JH titer in workers.

DiaCRALDCP exhibited an age-dependent decrease, except that

day-7 FG had low transcript level. It is speculated that 1FG may have produced sufficient amount of protein for acute physiological change, so that 7FG may not have to produce additional amount of this protein. In *Diacamma*, gamergates and dominant workers have high JH level (Sommer et al. 1993). Further, ICs, FGs, gamergates and dominant workers have similar, but somewhat different CHC profiles (Cuvillier-Hot et al. 2001; 2002). The complex patterns of *DiaCRALDCP* may explain these physiological dynamics, but it needs further investigations.

The scarceness of isolated genes in this study may partially be due to the weakness of differential display method. Differential display has insensitivity in detecting rare cDNAs due to the competitive nature of PCR amplification (Bertioli et al. 1995; Zhang et al. 1998). Some changes in the rare cDNA amounts during reproductive differentiation might not be detected in our method. However, since reproductive differentiation in queenless ants only involves physiological changes without morphogenesis, a small but essential change in transcription may contribute to effective physiological alterations. In future studies, functional analyses of such candidate genes, including *DiaCRALDCP*, will provide us with additional knowledge of the molecular bases of social evolution.

Acknowledgements

We are grateful to T. Kikuchi, M. Suwabe, H. Ohnishi, and H. Shimoji for their assistance in our field sampling. We also thank K. Yamazaki and T. Tojo for their kind instruction on FluorImager. This work was supported by Grants-in-Aid for Scientific Research (Nos. 20033002, 20370030, and 21677001) from the Ministry of Education, Culture, Sports, Science and Technology of Japan. YO, SM, and SK were supported by the Japan Society for the Promotion of Science (JSPS) research fellowship for young scientists.

References

Baratte S, Cobb M, Peeters C (2006) Reproductive conflicts and mutilation in queenless *Diacamma* ants. Anim Behav 72: 305-311

Bauer D, Miiller H, Reich J, Riedel H, Ahrenkiel V, Warthoe P, Strauss M (1993) Identification of differentially expressed mRNA species by an improved disply technique (DDRT-PCR). Nucl Acids Res 21: 4272-4280.

Bertioli DJ, Schlichter UHA, Adams MJ, Burrows PR, Steinbiss HH, Antoniw JF (1995) An analysis of differential display shows a strong bias towards high copy number mRNAs. Nucl Acids Res. 23: 4520-4523.

Billen J (1998) The social insect as a glandular factory. Ins Soc Life 2: 9-14

Cuvillier-Hot V, Cobb M, Malosse C, Peeters, C (2001) Sex, age and ovarian activity affect cuticular hydrocarbons in *Diacamma ceylonense*, a queenless ant. J Insect Physiol 47: 485-493

Cuvillier-Hot V, Gadagkar R, Peeters C, Cobb M (2002) Regulation of reproduction in a queenless ant: aggression, pheromones and reduction in conflict. Proc R Soc Lond B 269: 1295-1300

Cuvillier-Hot V, Lenoir A, Peeters C (2004) Reproductive monopoly enforced by sterile police workers in a queenless ant. Behav Ecol 15: 970–975

Diehl PA (1975) Synthesis and release of hydrocarbons by the oenocytes of the desert locust, Schistocerca gregaria. J Insect Physiol 21: 1237–1246

Dietemann V, Peeters C, Liebig J, Thivet V., Hölldobler B. (2003) Cuticular hydrocarbons mediate recognitions of queens and reproductive workers in the ant *Myrmecia gulosa*. Proc Natl Acad Sci USA 100: 10341-10346

Fan YL, Zurek L, Dykstra MJ, Schal C (2003) Hydrocarbon synthesis by enzymatically dissociated oenocytes of the abdominal integument of the German cockroach, *Blatella germanica*. Naturwissenschaften 90: 121-126

Ferveur JF, Savarit F, O'Kane CJ, Sureau G, Greenspan RJ, Jallon JM (1997) Genetic feminization of pheromones and its behavioral consequences in Drosophila males. Science 276: 1555–1558

Fukumoto Y, Abe T, Taki A (1989) A novel form of colony organization in the 'queenless' ant *Diacamma rugosum*. Physiol Ecol Jpn 26: 55 - 61

Gadagkar R (1990) Origin and evolution of eusociality: A perspective from studying primitively eusocial wasps. J Genet 70: 1-31

Gotoh A, Sameshima S, Tsuji K, Matsumoto T, Miura T (2005) Apoptotic wing degeneration and formation of an altruism-regulating glandular appendage (gemma) in the ponerine ant *Diacamma* sp. from Japan (Hymenoptera, Formicidae, Ponerinae). Dev Genes Evol 215: 69 - 77

Goldman N, Yang Z (1994) A codon-based model of nucleotide substitution for protein-coding DNA sequences. Mol Biol Evo 11: 725-736.

Hölldobler B and Wilson EO (1990) The Ants. Harvard University Press. Cambridge, MA.

Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754-5.

Ito T, Kito K, Adati N, Mitsui Y, Hagiwara H, Sakaki Y (1994) Fluorescent differential display: arbitrarily primed RT-PCR fingerprinting on an automated DNA sequencer. FEBS Lett 351: 231–236

Liang P, Pardee AB (1992) Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science 257: 967–971

Liebig J, Peeters C, Oldham NJ, Markstadter C, Hölldobler B (2000) Are variations in cuticular hydrocarbons of queens and workers a reliable signal of fertility in the ant *Harpegnathos saltator*? Proc Natl Acad Sci USA 97: 4124-4131

Linskens MHK, Feng J, Andrews WH (1995) Cataloging altered gene expression in young and senescent cells using enhanced differential display. Nucl Acids Res 23: 3244-3251

Lourenço AP, Mackert A, dos Santos Christino A, Simões ZLP (2008) Validation of reference genes for gene expression studies in the honey bee, *Apis mellifera*, by quantitative real-time RT-PCR. Apidologie 39: 372–385.

Lucas C, Pho DB, Fresneau D, Jallon JM, (2004) Hydrocarbon circulation and colonial signature in *Pachycondyla villosa.* J Insect Physiol 50: 595-607

Noirot C, Quennedey A (1991) Glands, gland cells, glandular units: some comments on terminology and classification. Ann Soc Entomol Fr 27: 123–128

Okada Y, Miyazaki S, Miyakawa H, Ishikawa A, Tsuji K, Miura T (2010) Ovarian development and insulin-signaling pathways during reproductive differentiation in the queenless ponerine ant *Diacamma* sp. J Insect Physiol 56:288-295

Panagabko C, Morley S, Hernandez M, Cassolato P, Gordon H, Parsons R, Manor D, Atkinson J (2003) Ligand specificity in the CRAL-TRIO protein family. Biochemistry 42: 6467-6474

Peeters C (1991) The occurrence of sexual reproduction among ant workers.

Biol J Lin Soc 44: 141-152

Peeters C, Crewe R (1985) Worker reproduction in the ponerine ant *Ophthalmopone berthoudi*: an alternative form of eusocial organization. Behav Ecol Sociobiol 18: 29-37

Peeters C, Liebig J. (2009) Fertility signaling as a general mechanisms regulating reproductive division of labor. : Organization of insect societies: from genome to sociocomplexity (Gadau J. and Fewell J., Eds), Harvard University Press, Cambridge, MA. pp 212-220

Peeters C, Higashi S (1989) Reproductive dominance controlled by mutilation in the queenless ant *Diacamma australe*. Naturwissenschaften 76: 177-180

Peeters C, Tsuji K (1993) Reproductive conflict among ant workers in *Diacamma* sp. from Japan: dominance and oviposition in the absence of the gamergate. Insect Soc 40: 119-136

Pereboom JJM, Jordan WC, Sumner S, Hammond RL, Bourke AFG (2005) Differential gene expression in queen-worker caste determination in bumble-bees. Proc R Soc Lond B 272: 1145-1152

Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14:817-818.

Robinson GE, Grozinger CM, Whitefield CW (2005) Sociogenomics: social life in molecular terms. Nat Rev Genet 6: 257-270

Scharlaken B, de Graaf DC, Goossens K, Brunain M, Peelman LJ, Jacobs FJ (2008) Reference gene selection for insect expression studies using quantitative real-time PCR: The head of the honeybee, *Apis mellifera*, after a bacterial challenge. J Insect Sci 8: 10.

Sommer K, Hölldobler B, Rembold H (1993) Behavioral and physiological aspects of reproductive control in a *Diacamma* species from Malaysia (Formicidae, Ponerinae). Ethology 94: 162-170

Sumner S, Pereboom JJM, Jordan WC (2006) Differential gene expression and phenotypic plasticity in behavioural castes of the primitively eusocial wasp, *Polistes canadensis*. Proc R Soc Lond B 273: 19- 26

Suyama M, Torrents, D, Bork P (2006) PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucl Acids Res 34: W609-W612.

Swofford, DL (2000) PAUP*: Phylogenetic Analysis Using Parsimony (*and other methods). Sinauer Associates, Sunderland, Massachusetts.

Toth AL, Varala K, Newman TC, Miguez FE, Hutchison SK, Willoughby DA, Simons JF, Egholm M, Hunt JM, Hudson, ME, Robinson GE (2007) Wasp Gene Expression Supports an Evolutionary Link Between Maternal Behavior and Eusociality. Science 318: 441-444

Turillazzi S, West-Eberhard MJ (1996) Natural history and evolution of paper-wasps. Oxford University Press. New York, NY.

Wang T, Montell C (2007) Phototransduction and retinal degeneration in *Drosophila*. Eur J Physiol 454: 821-847

West-Eberhard MJ (1979) Sexual selection, social competition and evolution. Proc Am Phil Soc 123: 222-234

Wilson EO (1971) The Insect Societies. Harvard University Press. Cambridge, MA.

Zhang JS, Duncan EL, Chang ACM, Reddel RR (1998). Differential display of mRNA. Mol. Biotechnol. 10: 155–165.

Zhou S, Tejada M, Wyatt GR, Walker VK (2006a) DNA-binding protein, tfp1, involved in juvenile hormone-regulated gene expression in *Locusta migratoria*. Insect Biochem Mol Biol 36: 726-734

Zhou X, Oi FM, Scharf ME (2006b) Social exploitation of hexamerin: RNAi reveals a major caste-regulatory factor in termites. Proc Natl Acad Sci USA 103: 4499–4504

Figure legends

Fig. 1

Schematic diagram of reproductive differentiation in *Diacamma*. At eclosion, callow females of *Diacamma* sp. posses a pair of thoracic appendages called gemmae. Gemma-mutilated females differentiate into workers, whereas females that retain gemmae become future gamergates with developed ovaries.

Fig. 2

Amino acid sequence alignment of DiaCRALDCP and insect CRAL-TRIO family proteins. Identical residues to the DiaCRALDCP sequence are highlighted in black. The CRAL-TRIO lipid-binding domain is boxed. *Aedes* CRAL-TRIO domain containing protein, *Drosophila* CG2663 PB isoform B, *Locusta* tfp1, *Apis* similar to CG2663 PB isoform B were aligned to DiaCRALDCP.

Fig. 3

Phylogenetic relationships among CRAL-TRIO family proteins including *Diacamma* CRAL-TRIO domain containing protein (*DiaCRALDCP*), based on coding DNA sequences. The topology and branch lengths shown were obtained by Bayesian inference of phylogeny, with the HKY + G model of substitution (-InL = 10137.17742). Posterior probabilities (PP), expressed as percentages, are shown above branches to indicate the level of support for each node. Only one number was given if the PP value was identical at that node for the two independent analyses. Bootstrap values (expressed as percentages) from ML and NJ analyses are shown below nodes. An asterisk indicates node that was not supported in greater than 50% of ML and NJ bootstrap replicates.

Fig. 4

Expression levels of *DiaCRALDCP* in *Diacamma* depending on reproductive role. Relative transcript levels normalized based on *Actin* expression are

shown. **a** Comparison of *DiaCRALDCP* levels between the brain and HEB (head excluding brain) in intact callows (ICs). Relative expression levels (fold changes) to that in the brain are shown. **b** Comparison of *DiaCRALDCP* levels in the HEB of workers and reproductives. Relative expression levels (fold changes) to that in ICs are shown. An arrow indicates the timing of gemma mutilation. Grey: ICs, White: workers, Black: reproductives; W: worker, FG: future gamergate, G: gamergate, sub: subordinate worker, dom: dominant egg-laying worker. Error bars (SD) and statistical analyses are based on the experimental triplications. * $P < 0.05$, *** $P < 0.001$, Student's t-test.

Fig. 5

Tissue- and reproductive-role specific *DiaCRALDCP* expressions in *Diacamma*. Relative transcript levels normalized based on *Actin* expression are shown. **a** Tissue-specificity of *DiaCRALDCP* in intact callows (ICs). Relative expression levels (fold changes) to that in the eye are shown. Different letters indicate significant statistical differences in the transcript level ($P < 0.05$, ANOVA, Tukey's HSD test). **b** Tissue-specificity of *DiaCRALDCP* in individuals with different reproductive roles. Arrowheads indicate timing of gemma mutilation. Grey bar: intact callow (IC), White bar: day-1 worker (1W), Black bar: day-1 future gamergate (1FG). Relative expression levels (fold changes) to that in IC antennae are shown. Error bars (SD) and statistical analyses are based on the experimental triplications. ($P < 0.05$, ANOVA, Tukey's HSD test).

Supplementary Table 1

Primer sequences used in real-time qPCR.

Supplementary Fig. 1

Expression levels of candidate genes obtained by differential display. Relative transcript levels in heads were normalized based on *Actin* expression. Relative expression levels (fold changes) to that in ICs are shown. Grey: ICs, White: workers, Black: FGs; W: worker, FG: future

gamergate, Error bars (SD) and statistical analyses originate from experimental triplications. * P < 0.05, *** P < 0.001, Student's t-test.

Supplementary Fig. 2

Unnormalized expression levels of *DiaCRALDCP* in *Diacamma* depending on reproductive role. Transcript levels obtained from the same amount of RNA are shown. **a** Comparison of *DiaCRALDCP* levels between the brain and HEB (head excluding brain) in intact callows (ICs). Relative expression levels (fold changes) to that in the brain are shown. **b** Comparison of *DiaCRALDCP* levels in the HEB of workers and reproductives. Transcript levels obtained from same amount of RNA are shown. An arrow indicates timing of gemma mutilation. Grey: ICs, White: workers, Black: reproductives; W: worker, FG: future gamergate, G: gamergate, sub: subordinate worker, dom: dominant egg-laying worker. Error bars (SD) and statistical analyses originate from experimental triplications. * P < 0.05, *** P < 0.001, Student's t-test.

Supplementary Fig. 3

Unnormalized tissue- and reproductive-role specific *DiaCRALDCP* expression in *Diacamma*. Transcript levels obtained from the same amount of RNA are shown. **a** Tissue-specificity of *DiaCRALDCP* in intact callows (ICs). Relative expression levels (fold changes) to that in the eye are shown. Different letters indicate statistical differences in the transcript level (P < 0.05, ANOVA, Tukey's HSD test). **b** Tissue-specificity of *DiaCRALDCP* in individuals with different reproductive roles. Arrowheads indicate timing of gemma mutilation. Grey bar: intact callow (IC), White bar: day-1 worker (1W), Black bar: day-1 future gamergate (1FG). Relative expression levels (fold changes) to that in IC antennae are shown. Error bars (SD) and statistical analyses originate from experimental triplications. Different letters indicate statistical differences (P < 0.05, ANOVA, Tukey's HSD test).

Fig. 1.

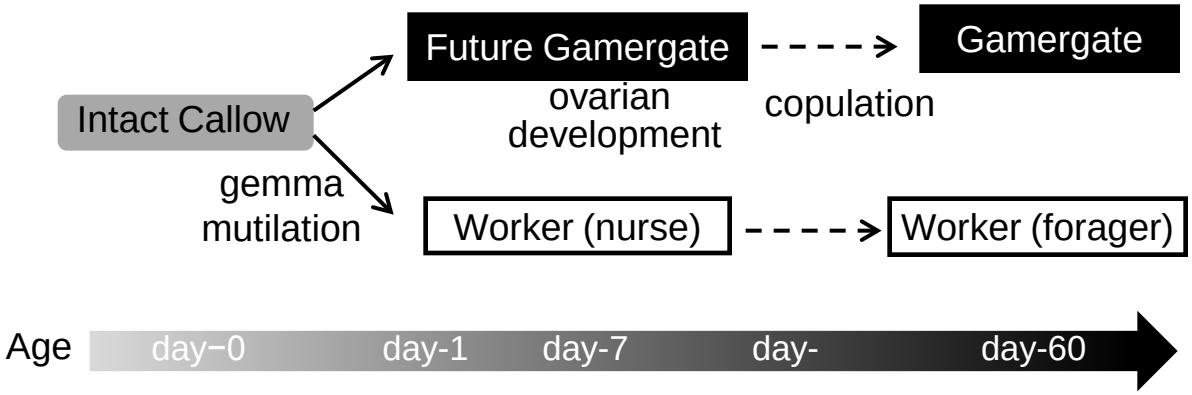

Future Gamergate
Gamergate
Intact Callow
ovarian
development
copulation
gemma
mutilation
Worker (nurse)
Worker (forager)
Age
day−0
day-1
day-7
day-
day-60

Fig. 2.

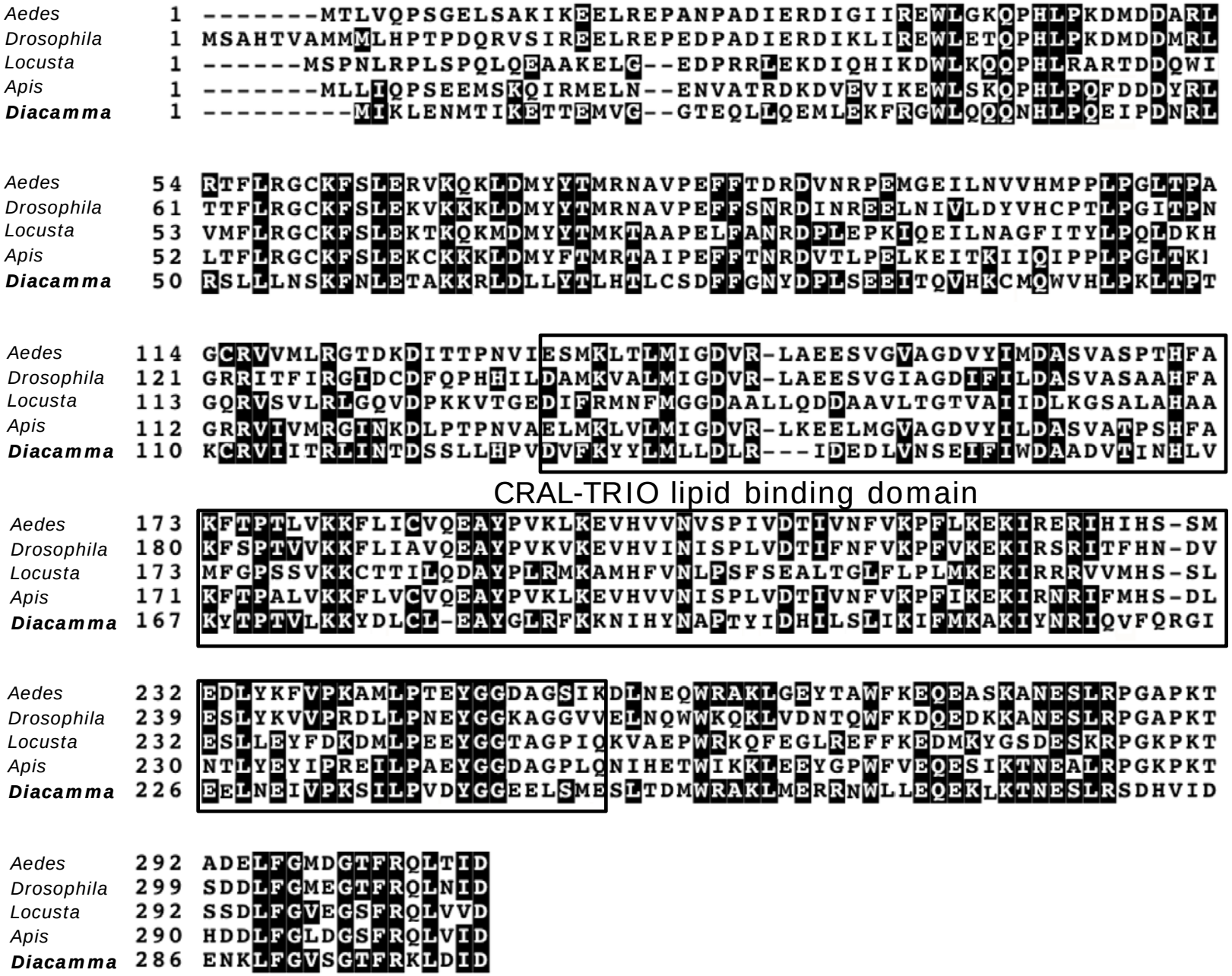

Aedes 1 -------MTLVQPSGELSAKIKEELREPANPADIERDIGIIREWLGKQPHLPKDMDDARL
Drosophila 1 MSAHTVAMMMLHPTPDQRVSIREELREPEDPADIERDIKLIREWLETQPHLPKDMDDMRL
Locusta 1 ------MSPNLRPLSPQLQEAAKELG--EDPRRLEKDIQHIKDWLKQQPHLRARTDDQWI
Apis 1 -------MLLIQPSEEMSKQIRMELN--ENVATRDKDVEVIKEWLSKQPHLPQFDDDYRL
Diacamma 1 ----------MIKLENMTIKETTEMVG--GTEQLLQEMLEKFRGWLQQQNHLPQEIPDNRL
Aedes 54 RTFLRGCKFSLERVKQKLDMYYTMRNAVPEFFTDRDVNRPEMGEILNVVHMPPLPGLTPA
Drosophila 61 TTFLRGCKFSLEKVKKKLDMYYTMRNAVPEFFSNRDINREELNIVLDYVHCPTLPGITPN
Locusta 53 VMFLRGCKFSLEKTKQKMDMYYTMKTAAPELFANRDPLEPKIQEILNAGFITYLPQLDKH
Apis 52 LTFLRGCKFSLEKCKKKLDMYFTMRTAIPEFFTNRDVTLPELKEITKIIQIPPLPGLTKI
Diacamma 50 RSLLLNSKFNLETAKKRLDLLYTLHTLCSDFFGNYDPLSEEITQVHKCMQWVHLPKLTPT
Aedes 114 GCRVVMLRGTDKDITTPNVIESMKLTLMIGDVR-LAEESVGVAGDVYIMDASVASPTHFA
Drosophila 121 GRRITFIRGIDCDFQPHHILDAMKVALMIGDVR-LAEESVGIAGDIFILDASVASAAHFA
Locusta 113 GQRVSVLRLGQVDPKKVTGEDIFRMNFMGGDAALLQDDAAVLTGTVAIIDLKGSALAHAA
Apis 112 GRRVIVMRGINKDLPTPNVAELMKLVLMIGDVR-LKEELMGVAGDVYILDASVATPSHFA
Diacamma 110 KCRVIITRLINTDSSLLHPVDVFKYYLMLLDLR---IDEDLVNSEIFIWDAADVTINHLV
CRAL-TRIO lipid binding domain
Aedes 173 KFTPTLVKKFLICVQEAYPVKLKEVHVVNVSPIVDTIVNFVKPFLKEKIRERIHIHS-SM
Drosophila 180 KFSPTVVKKFLIAVQEAYPVKVKEVHVINISPLVDTIFNFVKPFVKEKIRSRITFHN-DV
Locusta 173 MFGPSSVKKCTTILQDAYPLRMKAMHFVNLPSFSEALTGLFLPLMKEKIRRRVVMHS-SL
Apis 171 KFTPALVKKFLVCVQEAYPVKLKEVHVVNISPLVDTIVNFVKPFIKEKIRNRIFMHS-DL
Diacamma 167 KYTPTVLKKYDLCL-EAYGLRFKKNIHYNAPTYIDHILSLIKIFMKAKIYNRIQVFQRGI
Aedes 232 EDLYKFVPKAMLPTEYGGDAGSIKDLNEQWRAKLGEYTAWFKEQEASKANESLRPGAPKT
Drosophila 239 ESLYKVVPRDLLPNEYGGKAGGVVELNQWWKQKLVDNTQWFKDQEDKKANESLRPGAPKT
Locusta 232 ESLLEYFDKDMLPEEYGGTAGPIQKVAEPWRKQFEGLREFFKEDMKYGSDESKRPGKPKT
Apis 230 NTLYEYIPREILPAEYGGDAGPLQNIHETWIKKLEEYGPWFVEQESIKTNEALRPGKPKT
Diacamma 226 EELNEIVPKSILPVDYGGEELSMESLTDMWRAKLMERRNWLLEQEKLKTNESLRSDHVID
Aedes 292 ADELFGMDGTFRQLTID
Drosophila 299 SDDLFGMEGTFRQLNID
Locusta 292 SSDLFGVEGSFRQLVVD
Apis 290 HDDLFGLDGSFRQLVID
Diacamma 286 ENKLFGVSGTFRKLDID

Fig. 3.

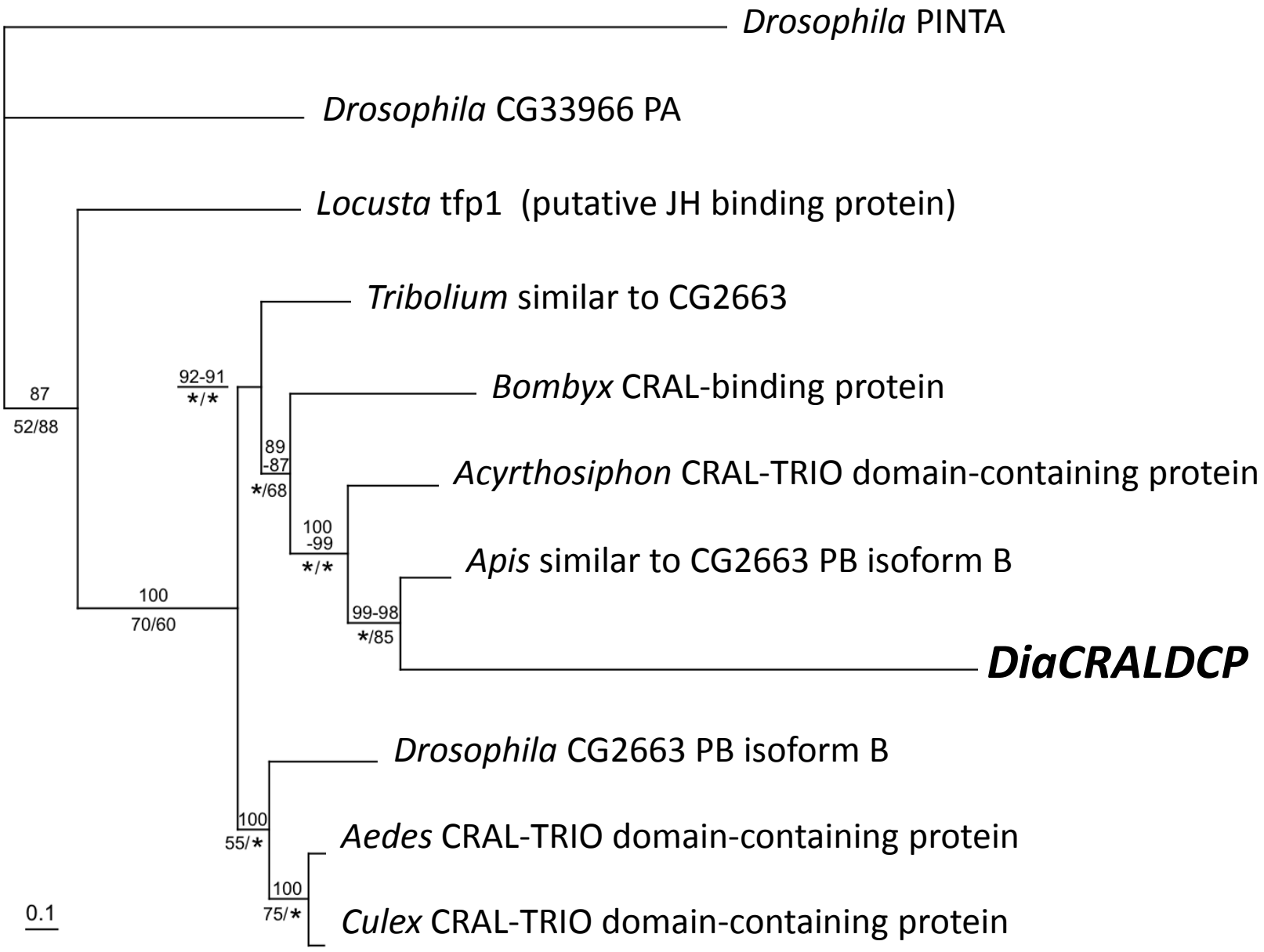

Drosophila PINTA
Drosophila CG33966 PA
Locusta tfp1 (putative JH binding protein)
Tribolium similar to CG2663
Bombyx CRAL-binding protein
Acyrthosiphon CRAL-TRIO domain-containing protein
Apis similar to CG2663 PB isoform B
DiaCRALDCP
Drosophila CG2663 PB isoform B
Aedes CRAL-TRIO domain-containing protein
Culex CRAL-TRIO domain-containing protein
87
52/88
92-91
/
89
-87
*/68
100
-99
/
100
70/60
99-98
*/85
100
55/*
100
75/*
0.1

Fig. 4.

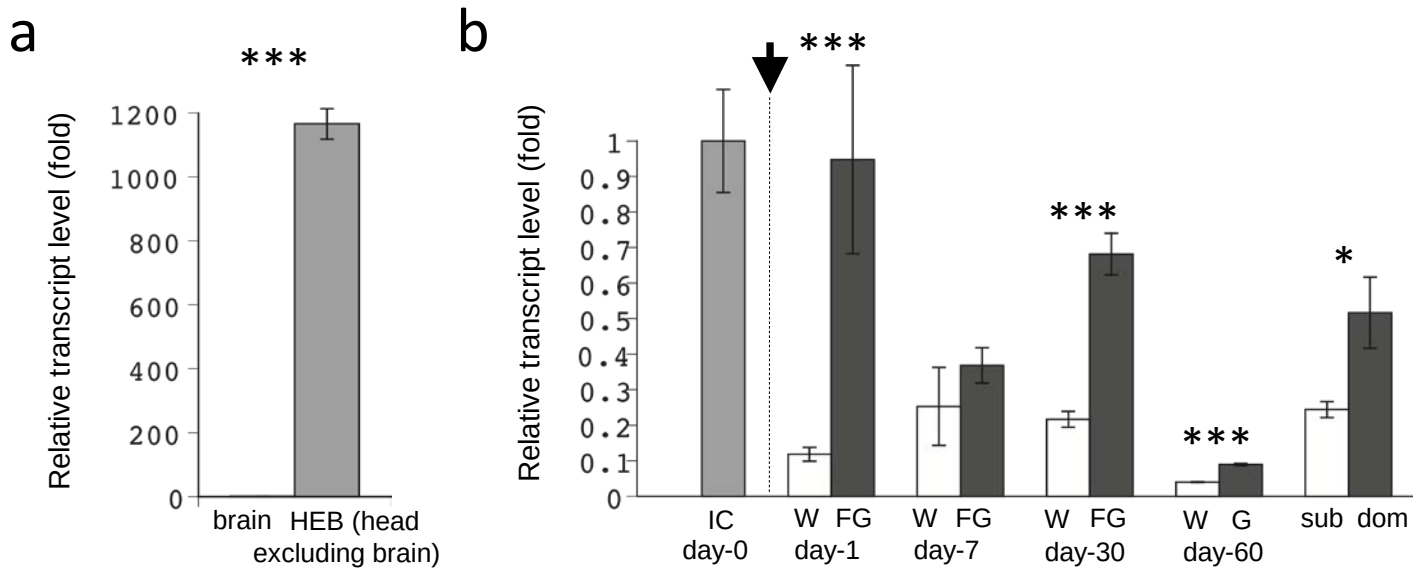
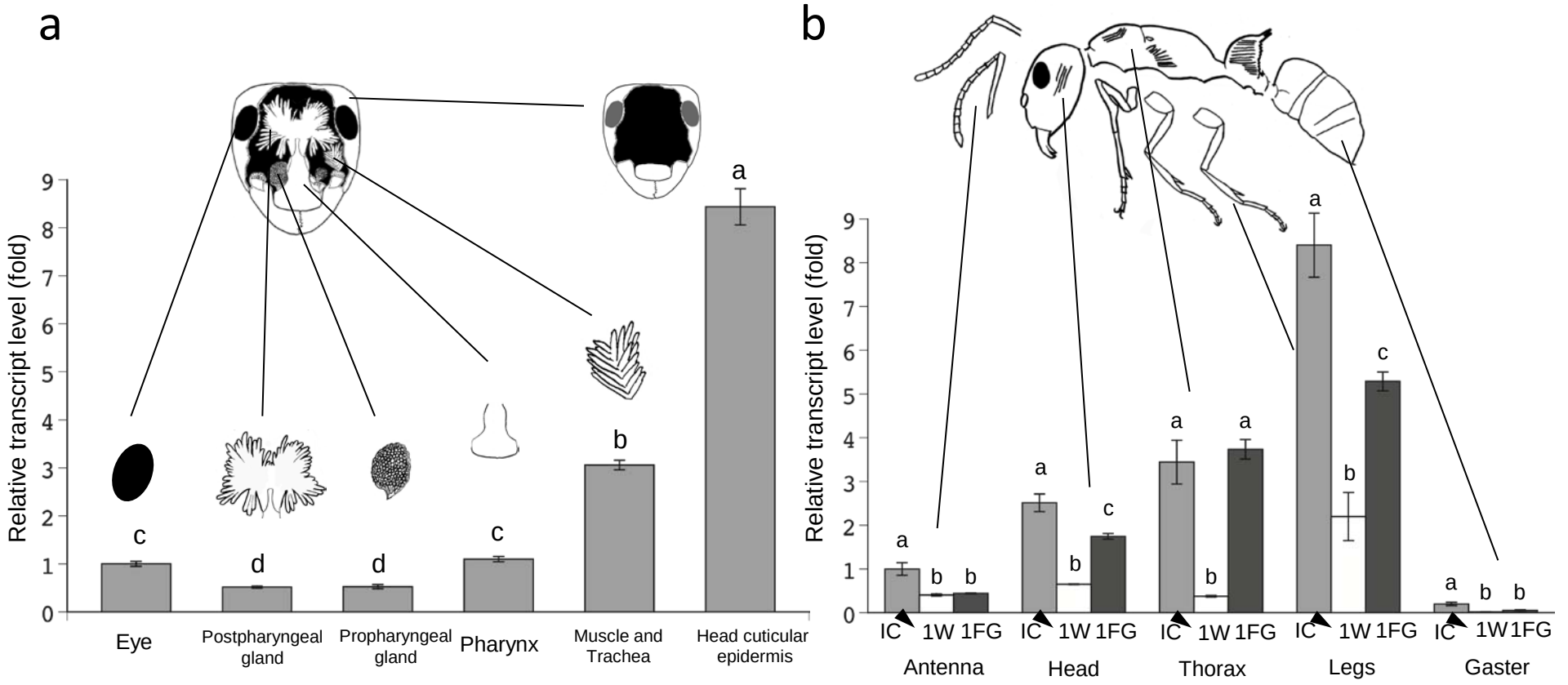

a
***
Relative transcript level (fold)
1200
1000
800
600
400
200
0
brain
HEB (head excluding brain)
b
***
***
*
***
Relative transcript level (fold)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
IC day-0
W FG day-1
W FG day-7
W FG day-30
W G day-60
sub dom

Fig. 5.


a
Relative transcript level (fold)
0
1
2
3
4
5
6
7
8
9
c
d
d
c
b
a
Eye
Postpharyngeal gland
Propharyngeal gland
Pharynx
Muscle and Trachea
Head cuticular epidermis
b
Relative transcript level (fold)
IC
1W
1FG
Antenna
Head
Thorax
Legs
Gaster