Evolution, 59(4), 2005, pp. 858–870

EVOLUTION OF INSECT METAMORPHOSIS: A MICROARRAY-BASED STUDY OF LARVAL AND ADULT GENE EXPRESSION IN THE ANT *CAMPONOTUS FESTINATUS*

MICHAEL A. D. GOODISMAN,[1,2,3,4] JUN ISOE,[1,2] DIANA E. WHEELER,[1,5] AND MICHAEL A. WELLS[1,2]
[1]*Center for Insect Science, University of Arizona, Tucson, Arizona, 85721*
[2]*Department of Biochemistry and Molecular Biophysics, University of Arizona, Tucson, Arizona, 85721*
[3]*School of Biology, Georgia Institute of Technology, Atlanta, Georgia, 30332*
[4]*E-mail: michael.goodisman@biology.gatech.edu*
[5]*Department of Entomology, University of Arizona, Tucson, Arizona, 85721*

Abstract.—Holometabolous insects inhabit almost every terrestrial ecosystem. The evolutionary success of holometabolous insects stems partly from their developmental program, which includes discrete larval and adult stages. To gain an understanding of how development differs among holometabolous insect taxa, we used cDNA microarray technology to examine differences in gene expression between larval and adult *Camponotus festinatus* ants. We then compared expression patterns obtained from our study to those observed in the fruitfly *Drosophila melanogaster*. We found that many genes showed distinct patterns of expression between the larval and adult ant life stages, a result that was confirmed through quantitative reverse-transcriptase polymerase chain reaction. Genes involved in protein metabolism and possessing structural activity tended to be more highly expressed in larval than adult ants. In contrast, genes relatively upregulated in adults possessed a greater diversity of functions and activities. We also discovered that patterns of expression observed for homologous genes in *D. melanogaster* differed substantially from those observed in *C. festinatus*. Our results suggest that the specific molecular mechanisms involved in metamorphosis will differ substantially between insect taxa. Systematic investigation of gene expression during development of other taxa will provide additional information on how developmental pathways evolve.

Key words.—*Camponotus festinatus*, cDNA microarray, developmental regulation of expression, *Drosophila melanogaster*, eusocial, Formicidae.

Received August 19, 2004. Accepted January 4, 2005.

The holometabolous insects are among the most diverse and evolutionarily successful clades of animals (Whiting 2002). By definition, holometabolous insects, such as Diptera (flies), Hymenoptera (ants, bees, and wasps), Coleoptera (beetles), and Lepidoptera (moths and butterflies), pass through larval, pupal, and adult developmental stages (Kristensen 1991; Sehnal et al. 1996; Heming 2003). The larval stage typically represents a period of growth (Heming 2003). Holometabolous insect larvae often inhabit distinct niches from adults and do not possess wings, true eyes, or functional reproductive systems (Truman and Riddiford 1999). The larval phenotype likely evolved under conditions that selected against wings, because larvae may have safely developed within protected substrates or food sources, where their wormlike body form was practical (Sehnal et al. 1996; Truman and Riddiford 1999). In contrast, the adult holometabolous insects often disperse and reproduce. Consequently, features not necessary for larvae, such as functional wings and reproductive systems, are of paramount importance for adults.

The pupal stage, which occurs between the larval and adult stages, marks a period of remarkable transformation, whereby the morphologically undifferentiated larva metamorphoses into a highly specialized adult (Gilbert et al. 1996b; Truman and Riddiford 1999, 2002; Heming 2003). During metamorphosis, larval tissues are destroyed and adult tissues are generated through cell apoptosis and differentiation. Metamorphosis is governed by the titer of important insect hormones, such as ecdysone and juvenile hormone (Gilbert et al. 1996a). These hormones lead to changes in the genes expressed within cells, which ultimately lead to the formation of the adult phenotype.

Although the general process of metamorphosis occurs in all holometabolous insects, the specific genes underlying changes in phenotype during metamorphosis in distinct taxa likely differ. Not surprisingly, the patterns of gene expression linked to metamorphosis are most thoroughly studied in the fruitfly *Drosophila melanogaster* (White et al. 1999; Arbeitman et al. 2002). However, a more complete understanding of metamorphosis requires analysis of other, nonmodel organisms. In particular, comparative studies between *D. melanogaster* and other insect taxa could help reveal key genes fundamentally involved in metamorphosis.

An increased understanding of the molecular basis of metamorphosis in hymenopteran social insects would provide a strong contrast to that in dipteran insects. Hymenopteran social insects (social bees and wasps, as well as all ants) hatch from eggs as larvae and pass through several larval instars (Hölldobler and Wilson 1990; Ross and Matthews 1991; Seeley 1995). Subsequent to larval growth, they enter the quiescent pupal stage. The new adult ecloses following metamorphosis and takes on its role in colony maintenance and homeostasis, in the case of workers, or reproduction, in the case of queens and males.

As with all holometabolous insects, larval and adult hymenopteran workers represent striking contrasts in form and function. Larvae are often completely dependent on adults, as they are usually incapable of moving or feeding themselves. Female larvae are even incapable of determining if they will develop into fertile queens or sterile workers (Hartfelder and Engels 1998; Evans and Wheeler 2000). In addition, worker larvae typically provide little or no useful work for the colony. In contrast, adult workers are highly specialized. They frequently possess morphological and physi-

ological adaptations used in colony defense and food acquisition. In addition, adult workers possess a wealth of glands used for communication inside and outside of the colony (Hölldobler and Wilson 1990).

The striking differences in physiology, morphology, and behavior between larval and adult workers strongly suggest that these life stages differ on a molecular level. Indeed, larvae and adults show variation in lipid, carbohydrate, and protein profiles (Hartfelder and Engels 1998). Patterns of gene expression presumably differ between the life stages as well (Evans and Wheeler 1999). The purpose of this study was to use cDNA microarray-based technologies to further our understanding of the evolution of insect metamorphosis. In this study we focused on identifying genes that differed in expression between larva1 and adult ant workers and comparing patterns of expression observed in ants to those observed in *D. melanogaster*. Our goal was to gain a better understanding of the level of conservation of the molecular mechanisms underlying metamorphosis in divergent insect taxa.

Materials and Methods

Insect System

We examined changes in gene expression between larval and adult *Camponotus festinatus* worker ants. *Camponotus festinatus* is a relatively common species that lives in a variety of habitats throughout the southwestern United States (Hunt and Snelling 1975). *Camponotus festinatus* laboratory colonies were initiated by single queens collected after mating flights. Colonies were allowed to grow and develop for three to five years under constant temperature (28–30°C) and ambient humidity. Ants were fed one of two diets. The first diet consisted of frozen *Nauphoeta cinerea* and *Manduca sexta* larvae as a protein source, and a 20% honey solution with 50 mg each of Vanderzants vitamins and Wesson salt mixture per 100 ml. The second diet consisted of ground tuna fish packed in water as a protein source, and 20 g table sugar, 100 mg Vanderzants, and 100 mg Wesson salt per 100 ml of water.

cDNA Library Construction

We constructed two independent cDNA libraries from distinct tissues obtained from *C. festinatus* laboratory colonies. The first library was composed of cDNA obtained from adult (imago) workers (hereafter referred to as A) and the second was composed of cDNA obtained from last instar larvae and pupae (hereafter referred to as L/P). To construct these two libraries, total RNA was obtained from approximately 100 whole *C. festinatus* adult workers and 100 fourth instar *C. festinatus* larvae and pupae from four distinct colonies. Total RNA was purified using the RNaid kit developed by Qbiogene (Irvine, CA). A total of 488 μg and 1100 μg of total RNA was obtained from the A and L/P tissue samples. The total RNA was concentrated by ethanol precipitation and then treated with DNAse I (Ambion, Austin, TX) to remove any potentially contaminating DNA. After DNAse I treatment, the total mass of RNA remaining was 251 μg for the A samples and 608 μg for the L/P samples. We then used the MicroPoly(A)Pure mRNA purification kit (Ambion) and purified 2.73 μg and 4.37 μg of mRNA from the A and L/P samples, respectively.

Double-stranded cDNA was made from the mRNA using the Superscript Choice System for cDNA synthesis (Invitrogen, Carlsbad, CA) and *Eco*R I adapters were ligated to the ends of the cDNA. Manufacturer's directions were followed with the exception that double-stranded cDNAs were size selected to be more than 500 bp using agarose gel purification rather than affinity columns. Double-stranded cDNAs were ligated into the plasmid vector pBluescript II SK(+), which was predigested to provide compatible *Eco*R I ends (Stratagene). XL10-Gold Ultracompetent Cells (Stratagene, La Jolla, CA) were then transformed with the plasmid library. Cells containing inserts were selected on ampicillin plates. In total, 101,850 independent colonies were collected resulting in a nondirectional cDNA. A worker library and 75,770 independent colonies were collected resulting in a nondirectional cDNA L/P worker library.

To determine sequence redundancy within each library, we initially obtained the sequence of 63 clones from the A library and 77 clones from L/P library (all sequencing procedures were conducted at the University of Arizona Genomic Analysis and Technology Core facility). Clone sequences were compared against existing protein and DNA sequences in the Genbank Nucleotide Sequence Database (nr = GenBank + RefSeq Nucleotides + EMBL + DDBJ + PDB sequences) and against predicted sequences from the *Apis mellifera* genome obtained from the *Apis mellifera* Gnomon gene prediction program, via BLASTX or BLASTN searches (Altschul et al. 1990). Clones that displayed a similarity of $E < 1 \times 10^{-10}$ to known or predicted GenBank sequences were considered putative homologues of those genes.

The distributions of clones homologous to sequences from various organismal taxa in the L/P and A cDNA libraries were compared with a *G*-test of independence. In addition, we measured the gene diversity of our libraries using a modification of Nei's formula:

$$d = \frac{n\left(1 - \sum_i f_i^2\right)}{(n - 1)}, \quad (1)$$

where n is the number of sequences analyzed and f_i is the frequency of sequence i (Nei 1987). The gene diversity measures the probability of randomly choosing two different sequences from the library and ranges from zero, indicating that all sequences within a library are identical, to one, indicating that all sequences within a library are unique.

Probe Preparation and Microarray Construction

We constructed a cDNA microarray composed of sequenced and anonymous cDNA clones. A total of 192 clones, including those sequenced in the preliminary sequencing screen described above, from each of the two libraries were amplified using polymerase chain reaction (PCR) in a final volume of 100 μl containing a stab of bacterial cellular matter from a single colony (containing the plasmid DNA of interest), 3.5 U Taq DNA polymerase (New England Biolabs,

Beverly, MA), and a final concentration of 200 μM dNTPs, 0.5 μM of each of plasmid-specific T3 and T7 PCR primers, and 1× New England Biolabs PCR buffer. The PCR cycling profile began with an initial denaturation at 94°C for 2 min, and then proceeded with 40 cycles of 94°C for 30 sec, 50°C for 30 sec, and 72°C for 2 min, followed by a final extension of 72°C for 10 min. Aliquots of PCR products were checked via agarose gel electrophoresis to confirm size and quality. Glycerol stocks were made of all clones printed on the microarray for subsequent analysis. PCR products were purified (Qiagen, Valencia, CA) to remove unincorporated primers and dNTPs and then resuspended in 10 μl of 3× SSC.

PCR products were printed on GAPS II coated slides (Corning) coated with amino-silane using a VersArray ChipWriter Pro System at the University of Arizona Genomic Analysis and Technology Core facility. Microarray spots measured 140–150 μm in diameter and were separated by 230 μm center to center. Each spot was printed on each array in quadruplicate.

Target Preparation

Total RNA was extracted from adult and larvae *C. festinatus* ants as described above and treated with DNAse according to manufacturer's recommendations (note that pupae were not included in the target sample). Approximately 20 μg of total RNA was then converted to single-stranded cDNA using oligo d(T) as a primer. The reverse trascription reaction took place in a volume of 30 μl and contained 400 U of Superscript II (Invitrogen) reverse transcriptase and a final concentration of 1× Superscript II Buffer (Invitrogen), 1× Aminoallyl-dNTP (Sigma, St. Louis, MO), and 0.02 M DTT. Reactions were incubated at 42°C for 3 h and terminated by the addition of 10 μl of 1 M NaOH and 10 μl of 0.5 M EDTA followed by incubation at 65°C for 15 min.

The single-stranded cDNA was then cleaned using the Qiaquick (Qiagen) system and eluted in 60 μl of 19:1 mixture of 1 M K_2HPO_4 and 1 M KH_2PO_4. The cDNA was then dried down by speed-vac and resuspended in 4.5 μl of 0.1 M Na_2CO_3. To the cDNA, 4.5 μl of either Cy3 or Cy5 dyes (Amersham, Piscataway, NJ) suspended in DMSO were added and incubated in the dark for 1 h at room temperature. Following coupling, 35 μl of 100 mM NaOAc was added to the reaction, which was then cleaned using the Qiaquick system. The labeled cDNA was then eluted in 60 μl of elution buffer (Qiagen) and 10 μl of Cot 1 DNA (Invitrogen) and 2 μl of poly-A DNA (Invitrogen) were added. This DNA cocktail was then dried down by speed-vac and resuspended in 5 μl of 10 mM EDTA and heated for 10 min. at 96°C.

Hybridization

Microarray slides were held over a beaker of heated water for 5 sec to rehydrate spots and then snap dried by placing the slide on a heat block for 5 sec. The DNA was then crosslinked to the slide using a Stratalinker (Stratagene). The slide was then incubated in an orbital shaker in an SDS bath for 2 min and then rinsed 10–15 times in MilliQ water (Billerica, MA). Hybridizations were conducted in a GeneTAC hyb station (Genomic Solutions, Ann Arbor, MI). A total of 140 μl of SlideHyb Glass Array Hybridization Buffers (Ambion) was heated to 68°C and added to the hybridization chamber. The target cDNA was then injected into the hybridization chamber and hybridized at 60°C for 16 h. Following hybridization, slides were washed with GeneTAC (Genomic Solutions) medium wash buffer (42°C for 40 sec), high stringency wash buffer (42°C for 40 sec), and twice with postwash buffer (42°C for 40 sec). Slides were then dunked in 70% ethanol five times and spun dry. Slides were scanned with an arrayWoRx Biochip Reader. Exposure time for the Cy3 channel was 0.8 sec and exposure time for the Cy5 channel was 0.2 sec. Spot features (i.e., mean, median, standard deviation, and background intensities in the two channels) were extracted from TIFF image using DigitalGENOME software (MolecularWare, Irvine, CA).

Experimental Design, Array Normalization, and Statistical Analyses

To obtain accurate readings of gene expression, we conducted five microarray hybridization assays, a number expected to yield sufficient statistical power to distinguish genes that differ by as little as 1.5-fold in expression (Gibson 2002). The five hybridizations represented independent RNA extractions and flip protocols. All raw microarray data are available from the authors.

After scanning of slides, mean intensities in the two channels were $\log_2$ transformed, yielding values defined as $\log_2(\text{Intensity Dye1}) = \text{LD1}$ and $\log_2(\text{Intensity Dye2}) = \text{LD2}$. The microarray data were then normalized to remove the nonlinear effects of the dyes on the ratio of expression. The normalization process involved comparing the difference in log-intensities, $R = \text{LD1} - \text{LD2}$, to the mean log-intensity, $I = 0.5(\text{LD1} + \text{LD2})$, for each hybridization independently. We assumed that the average gene would show no difference in expression, a result that should lead to no association between R and I. Deviations from this predicted association, which presumably arose from dye artifacts, were removed using nonlinear regression techniques. Our method of normalization shares strong similarities to standard loess adjustment. However, nonlinear regression confers the same correction function across all datapoints, resulting in a more consistent adjustment.

We fit a series of hierarchical models to the intensity data to normalize the microarray data. The simplest models included only a linear term, but higher order terms (quadratic, cubic, etc.) were allowed if they significantly improved fit. The residuals from the model, R', were then used in conjunction with the given values of I to calculate new adjusted estimates of the intensities defined as $\text{LD1}' = (2I + R')/2$ and $\text{LD2}' = (2I - R')/2$. To adjust variances in dye intensities across microarrays, we employed the quantile method of Bolstad et al. (2003). Their algorithm involves independently ranking the intensities for each hybridization assay, and then replacing the value of the highest rank in hybridization assays 2–5 with the highest value in hybridization assay 1, and then replacing the value the second most highly ranked value in hybridization assays 2–5 with the second most highly ranked value in assay 1, et cetera. This method resulted in new adjusted intensity values LD1* and LD2* for hybridization assays 2–5.

We used the analysis of variance (ANOVA) to detect dif-

TABLE 1. Polymerase chain reaction (PCR) primers used in quantitative reverse transcriptase–PCR (QRT-PCR) analyses of gene expression. Clone identifications beginning with L and A were obtained from the larval/pupal (L/P) and adult (A) *Camponotus festinatus* libraries, respectively. The primer sequences (5′–3′), annealing temperatures (T_a), and PCR product sizes obtained from cDNA template used in all QRT-PCR assays are provided.

Clone	T_a (°C)	Size (bp)	Forward primer	Reverse primer
L-A-B10	45	168	CTTCATCATGCCATTTTATCGT	TATTTAAGTGAATCAACACGAG
L-A-D6	50	107[1]	ATCTCCGCGAGTCCTACAAA	GGTGAAGACCGAGTCTGAG
L-B-B10	50	114	CGTTCGCTTTAAACGTGGTT	CCGTATCCACCACCGTATC
L-B-D3	45	104	CGTTGGGATACCCACTTGA	CAATTCCTCCACCTGCTTGT
L-B-E10	50	124	AATTGCCGGATTTGTCACTC	AGCAGAGACCTCGGGAACA
L-B-E7	50	118	GGGAATGACCTTCTTTGCAC	TGGCGATAGAGGTCGATTG
L-B-E8	50	118	GCCGTTGCGTGAGTTATCAT	AATGATCGAGGCGAAAGTCC
L-B-F5	50	112	GATGGCTTTGTGCAATTCCT	TGATCTGCTACGTGCTACC
L-B-G4	45	123	TTGGTAGTGAACGTCGAGCA	CAGATGAGCGAGGCCAAGA
A-C-A3	50	113	CACAACGACGCATACAATGA	TACCGCAGCTCATGACAAAG
A-C-C4	50	119	GTCCCAAATGTGAAATCATGT	CGCCAAATAAGTGCTGAATG
A-C-E3	50	119[1]	CTATTGGAAGTGGCACTGG	TAACTTCCTTCCGCCAATGA
A-C-F1	50	120	TAGTTATGCGCGTTTCCACA	AGCGACGACTTCTTCCTTC
A-C-F8	55	121	CGAAAATAAACCGCGAAAAA	TGACGACGAAAACGTGCTA
A-D-C6	45	103	ATTGGAGCTGACGAGAATGG	GTCAGGGGTAAGATCGTGT
A-D-D10	50	115	GCTCCTTGATGGGAATGAAA	TTACCGTCGGTCGTTGAAG
A-D-D7	50	129[1]	GGGTAGCATTGTTTGGAGGA	AAATATCGAGAAGCGGACCA
A-D-D9	50	110	AGATTTACGCACGGATGCAC	TCGCCGGTCTACCTAATGA
A-D-E11	50	111[1]	CGAGAAGACCGTCAAGAAG	GCGAATGTCGAGTCCATTTC
A-D-E6	45	136	CCGAGAGCATCAAGTCTTC	AATTGTGGTTTCGACGATTC
A-D-E8	45	120[1]	GAGTGAAGCGATAAAAGCTG	GATCTTTGCATCCTGGTTTG
A-D-F12	45	115[1]	ATCCCAGAATCGTGAATGGA	TTTTCCGTCAGGATTGAACC

[1] PCR product sizes obtained from genomic DNA (as opposed to cDNA) template were larger, indicating the presence of an intron.

ferences in expression for individual genes (e.g., Wolfinger et al. 2001; Kerr et al. 2002). Specifically, we fit the model $y_{adsr} = \mu + A_a + D_d + AD_{ad} + S_s + R_{r(a)} + \varepsilon_{adsr}$, to the data on a gene-by-gene basis. In this model, y_{adsr} is the quantile-adjusted fluorescent intensity for a given gene, which reflects the overall mean intensity μ, and the effects of array *a*, dye *d*, array by dye interaction *ad*, stage (larva or adult) *s*, replicate spots within array *r*(*a*), and error ε_{adsr}. The model was fit using the MIXED procedure of SAS (SAS Institute, Inc., Cary, NC) with array, array by dye, and spot within array treated as random effects, and stage and dye treated as fixed effects. The output of the procedure resulted in least-square mean estimates for the intensity of a gene in larvae and adults, the ratio of which was taken as an estimate of the ratio of expression of that gene in the two life stages. In addition, we obtained *P*-values that reflected the significance of the difference in intensities between larval and adult stages. To correct for the fact that analyses were performed on 384 genes, we used the Bonferroni method (Sokal and Rohlf 1995); consequently, genes were considered to show significant differences in expression between larvae and adults if $P < 0.05/384 = 0.00013$.

Confirmation with Quantitative Reverse Transcriptase Polymerase Chain Reaction

To confirm the expression results obtained from our microarray experiments, we selected several genes for further expression analysis using quantitative reverse transcriptase PCR (QRT-PCR). Most clones were chosen because they showed strongly divergent patterns in expression between larval and adult ants. A total of 23 clones from L/P library and 18 clones from the A library were obtained from frozen stocks and sequenced. PCR primers were designed for 13 clones from the L/P library and 16 clones from the A library with the aid of the program Primer3 (available via http://frodo.wi.mit.edu.cgi-bin/primer3/primer3_www.cgi). The nine primer pairs from the L/P library and 13 primer pairs from the A library that gave consistent amplification products from genomic DNA and plasmid DNA libraries were used for further study in QRT-PCR assays (Table 1; hereafter clone designations beginning with ''L'' or ''A'' were obtained from the larval/pupal or adult cDNA libraries, respectively).

For QRT-PCR, total RNA was extracted from larval and adult *C. festinatus* ants as described above on two separate occasions from distinct sets of colonies. This yielded four RNA samples, larval-1 (L1), adult-1 (A1), larval-2 (L2), and adult-2 (A2). All samples were treated with DNAse I so that no PCR product could be obtained from any of the samples, indicating an absence of significant genomic DNA contamination. A total of 1.81 μg (L1), 1.28 μg (A1), 1.40 μg (L2), and 1.18 μg (A2) of total RNA were used to make single-stranded cDNA in a 30-μl reaction using a poly-T oligonucleotide to prime the reaction (Applied Biosystems, Foster City, CA) according to the manufacturer's directions. To control for differences in cDNA synthesis efficiency or errors in RNA quantification (Bustin 2002; Klein 2002), dilutions of the single-stranded cDNA from larvae and adults were visualized on an agarose gel and quantitated using digitized imaging (Kodak, Rochester, NY) to obtain accurate readings of the amounts of cDNA used as templates in the PCR.

To quantitate differences in expression levels for distinct genes, we conducted PCRs on an ABI PRISM 7700 (Applied Biosystems) thermocycler using Cybr Green to detect double-stranded PCR product during amplification. PCRs were conducted in a final volume of 10 μl containing 1 μl of 10-fold or 100-fold diluted cDNA, 0.5 μM of each of clone-specific

forward and reverse PCR primers (Table 1), and 5 μl of either Applied Biosystems or Invitrogen 2× Cybr Green mix (containing Taq DNA Polymerase, dNTPs, Passive Reference, and buffer components). The PCR cycling profile began with an initial denaturation at 95°C for 5 min, and then proceeded with 40 cycles of 94°C for 30 sec, clone-specific annealing temperature (Table 1) for 30 sec, and 60°C extension for 1 min, in which time amplicon concentrations were determined by amount of Cybr Green intercalation. All reactions were carried out in triplicate.

To accurately account for differences in amplification efficiency among clones, standard amplification curves were generated for each clone using 10-fold serial dilutions of clone-specific plasmid DNA ranging from approximately 10^{-1} to 10^{-8} ng/μl of plasmid. In several cases, very high or very low plasmid copy numbers resulted in aberrant amplification and were not included in standard curve analyses. Standard curves relating the threshold cycle (C_t) to the initial number of molecules present were constructed from the remaining datapoints. The standard curves were then used to distinguish relative clone initial copy number in experimental cDNAs by comparing amount of product generated by cDNA templates during exponential amplification.

Ratios of expression (larval/adult) of genes were obtained by interpolating values obtained from cDNAs L1, A1, L2, and A2 on the gene-specific standard curves during exponential PCR amplification. The ratios of cDNA for the investigated genes were normalized (L1/A1 and L2/A2) using the relative amounts of cDNA starting material within a given set of RNA extractions (RNA extractions 1 vs. 2). The significance of differences in expression between larvae and adults was assessed using two-factor ANOVA under the model $y_{rs} = \mu + R_r + S_s + RS_{rs} + \varepsilon_{rs}$, where the expected number of copies of particular gene y_{rs} reflects the overall mean intensity μ, and variation due to RNA extraction *r*, stage (larva or adult) *s*, RNA extraction by stage interaction *rs*, and error ε_{rs} (in a few cases other factors were included in the model to account for variation in cDNA dilution). Statistical calculations were performed using the GLM procedure of SAS (SAS Institute, Inc.) treating both RNA extraction and stage as fixed effects. The ratios of the least square means for the stages (larval/adult) were taken as the best estimate of ratio in gene expression obtained and probability that the number of starting copies of cDNA were the same was given by the *P*-value. Bonferroni corrections were implemented to adjust *P*-values for the multiple tests performed. Consequently, genes were considered to show significant differences in expression between larvae and adults if $P < 0.05/22 = 0.0023$.

Results

cDNA Library Sequence Analysis

We constructed two *C. festinatus* ant cDNA libraries, one from late instar worker larvae and pupae (L/P) and the other from adult workers (A), to analyze changes in gene expression that occurred during ant development. Initially, 62 sequences were obtained from the A library. Twenty-eight of these sequences showed significant homology to protein or nucleotide sequences in GenBank as measured by BLASTX or BLASTN similarity of $E < 1 \times 10^{-10}$. In addition, no two of the 62 sequences were identical giving an effective library diversity of $d = 1.0$. We obtained sequences from 77 random clones from the L/P library, 39 of which matched known sequences in GenBank. In contrast to the A library, the L/P library contained multiple copies of a few sequences including some ribosomal RNA. Nevertheless, a majority of the sequences in the library were unique, and library diversity was still very high ($d = 0.986$).

Subsequent to the initial cDNA library screen and analysis of gene expression (see below), we ultimately obtained the sequences of 91 clones from the L/P library and 76 clones from the A library (Fig. 1; GenBank accession numbers CK656423–CK656590; see Appendix available online only at http://dx.doi.org/10.1554/04-514.1.s1). Approximately 40% of the clones in both libraries failed to display homology to any known sequences. Another large proportion of clones showed strongest similarity to sequences in the *A. mellifera* genome. Most other clones exhibited similarity to sequences from other insects. In addition, a small percentage of sequenced clones from the libraries showed strongest similarities to more distantly related taxa, such as vertebrates, bacteria, plants, or fungi. Overall, the distribution of clones differed significantly between the L/P and A libraries ($G_7 = 19.97$, $P < 0.01$; Fig. 1).

Analysis of Gene Expression

To accurately measure differences in transcript abundance between larval and adult *C. festinatus* ants, we first needed to normalize the microarray data to remove potential sources of bias. All five microarray hybridizations required normalization (raw microarray data available from the authors). For hybridizations 1, 2, and 3, the relationship between the log-ratios of intensity, *R*, and the mean intensity, *I*, was linear, with $R = 1.846 - 0.178I$, $R = -0.843 + 0.150I$, and $R = -0.015 - 0.163I$, respectively. For hybridization 4, a more complex model was required and the relationship between *R* and *I* was given by $R = 4.764 - 0.839I + 0.029I^2$. Finally, hybridization 5 required a third-order term to yield relatively unbiased data, with the association between *R* and *I* given by $R = -3.833 - 2.510I + 0.434I^2 - 0.020I^3$. Although the transformation required differed substantially among hybridizations, none was particularly major in effect.

We next used standard statistical techniques to determine the proportion of genes that showed significant changes in expression over time. Of the 384 genes analyzed, 219 showed a significant difference between larvae and adults when the unadjusted cutoff level of $\alpha = 0.05$ was used (Fig. 2). This is in great excess to the number of tests expected to be significant by chance alone ($384 \times 0.05 = 19.2$), indicating that many genes were differentially expressed between the life stages. After correcting for the multiple tests performed using the conservative Bonferroni procedure, 91 genes displayed a $P < 0.05/384 = 0.00013$, and were deemed as showing significant differences in gene expression between the life stages. As expected, genes showing the highest statistical significance in expression (lowest *P*-values) also showed relatively large ratios of expression. However, several genes showed significant differences in expression with a less than

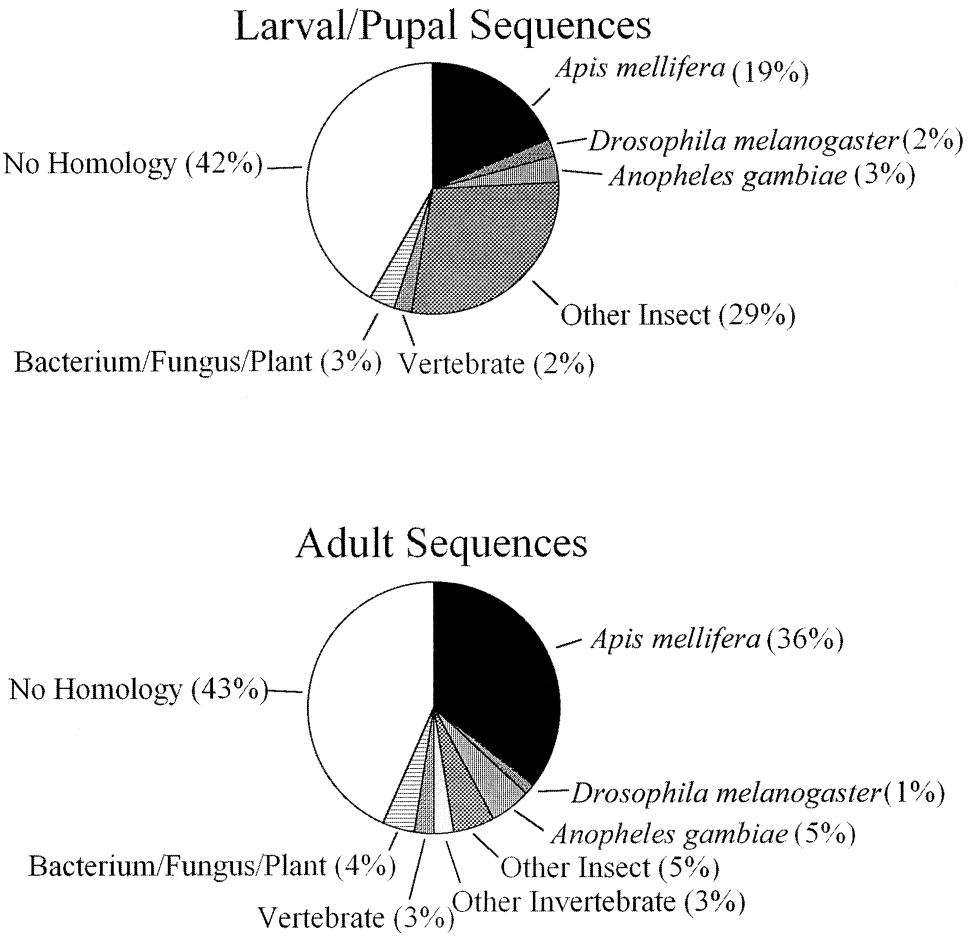

FIG. 1. Percentage of *Camponotus festinatus* sequences matching known gene sequences of different taxa.

two-fold difference in expression between larval and adult *C. festinatus* samples (Fig. 2).

Confirmation with Quantitative Reverse Transcriptase Polymerase Chain Reaction

We used quantitative RT-PCR to determine if the gene expression results obtained from the cDNA microarray were robust. We first designed PCR primers for 22 clones, most of which showed divergent patterns of gene expression between the larval and adult life stages (Table 1, Fig. 2). Analysis of real-time PCR amplification from known numbers of template molecules indicated that the amplification efficiencies (slope) and sensitivities (*y*-intercept) for each clone relating the threshold cycle (C_t) to the $\log_{10}$ (copy number) were fairly standard with the exception of clone A-D-D9, which showed unusually high PCR efficiency (Table 2; Bustin 2002).

To correctly assess the relative abundance of transcripts in larval and adult *C. festinatus* cDNA, it was first important to accurately estimate the starting mass of cDNA in each template. The alternative of using a housekeeping gene to normalize the results was not available to us because the existence of such genes has not been demonstrated in ants. Consequently, we used digital analysis of intensities to quantitate the initial mass of cDNA templates for QRT-PCR. By comparing the intensities of single-stranded cDNA (L1 vs. A1 and L2 vs. A2) on agarose gels, we determined that the mass ratios of larval to adult cDNA were 1.25 (L1/A1) and 1.71 (L2/A2). These ratios differed from those expected from the amount of total RNA used to synthesize the cDNA template, which equaled 1.41 and 1.18 for RNA extraction sets 1 and 2, respectively. This suggested that reverse trascription of RNA did not occur at equal efficiency for all reactions (Freeman et al. 1999).

In total, 21 of the 22 clones analyzed using QRT-PCR showed differences in expression in the same direction as those observed from the microarray analyses (Fig. 3). That is, genes that were overexpressed in larvae or adults as judged from the analyses of the cDNA microarray data showed the same qualitative patterns when analyzed using QRT-PCR. The patterns of significance between the two techniques were also largely similar; most genes that showed significant differences in expression as determined by microarray analysis also showed significant difference as determined by QRT-PCR. The exceptions to this pattern were clones L-A-B10, A-C-F1, A-D-D9, A-D-E6, A-D-E8, and A-D-F12. However, it should be noted that three clones, L-A-D6, L-B-G4, and A-D-D10, which did not show significant differences as de-

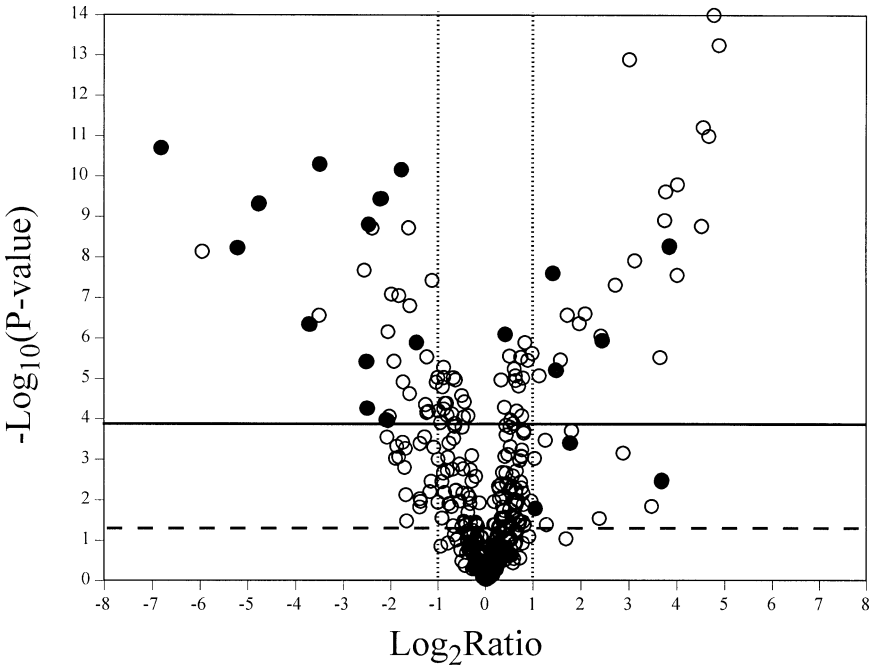

FIG. 2. Volcano plot of gene expression differences between larval and adult ants. The magnitude of the difference in gene expression between larval and adult *Camponotus festinatus* workers is given on the x-axis as the $\log_2$ ratio (larval/adult) of intensity, and the significance of the difference is given as the $-\log_{10}(P\text{-value})$ obtained from ANOVA. Points falling outside the dotted vertical lines represent genes showing two-fold differences in expression. Points falling above the dashed and solid horizontal lines represent genes showing significant differences in expression before ($P < 0.05$) and after Bonferroni corrections ($P < 0.00013$), respectively. Solid points symbolize genes used in reverse transcriptase–polymerase chain reaction analysis.

TABLE 2. Polymerase chain reaction (PCR) efficiencies and sensitivities for amplicons used to study gene expression by quantitative reverse transcriptase–PCR. Slopes (efficiencies) and *y*-intercepts (sensitivities) provide the linear relationships between threshold cycle (C_t) and the $\log_{10}$ of the number of initial starting template molecules.

Clone	Slope	*y*-intercept
L-A-B10	−3.44	36.19
L-A-D6	−2.96	40.23
L-B-B10	−2.93	38.79
L-B-D3	−3.46	41.45
L-B-E10	−3.31	40.72
L-B-E7	−3.00	41.29
L-B-E8	−3.28	41.91
L-B-F5	−3.13	40.56
L-B-G4	−3.56	46.37
A-C-A3	−3.22	39.70
A-C-C4	−3.58	43.91
A-C-E3	−3.07	37.67
A-C-F1	−3.52	42.24
A-C-F8	−3.11	37.64
A-D-C6	−3.22	37.69
A-D-D10	−3.38	38.92
A-D-D7	−3.63	37.84
A-D-D9	−2.19	35.27
A-D-E11	−3.53	27.38
A-D-E6	−3.21	41.25
A-D-E8	−3.43	39.59
A-D-F12	−3.73	34.86

termined by microarray analysis, did show significant differences when studied using QRT-PCR. Clone A-D-D9 was the major outlier from the comparisons, as it showed strongly divergent patterns of expression between the microarray and QRT-PCR results. We note, however, that this clone also displayed kinetics inconsistent with standard PCR (Table 2) and suggest that the results may be incorrect due to amplification problems. Consequently, the QRT-PCR results strongly support the microarray analyses overall and indicate that relative expression levels determined by the microarray are valid.

Genes Showing Differential Expression between Life Stages

We were interested in understanding the types of genes that showed variation in gene expression between larval and adult ants. Consequently, we obtained sequences for many of the clones showing significant differences in expression between these two *C. festinatus* stages. Many of these sequences displayed strong similarity to sequences in GenBank (Table 3). One pattern arising from examination of these sequences is that clones that were significantly overexpressed in larvae relative to adults tended to be derived from the L/P cDNA library (clone names beginning with L). Likewise, clones that were significantly overexpressed in adults were generally derived from the adult cDNA library (clone names beginning with A). This result was expected from the presumed tendency to clone genes from the L/P and A libraries that were relatively highly expressed in those stages. In addition, the lack of similarity between genes cloned from the two libraries supported the hypothesis that the expression

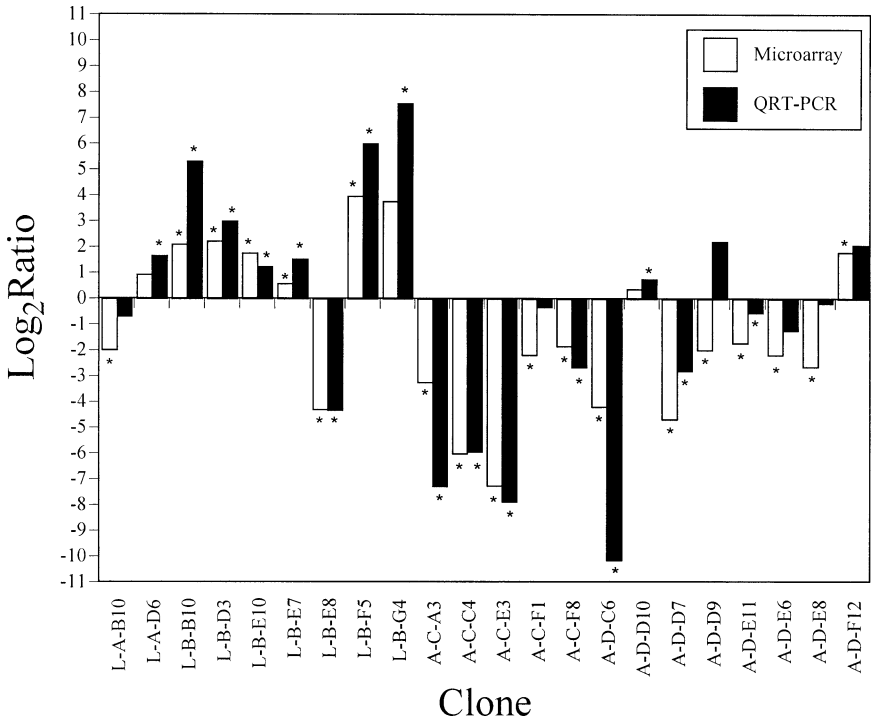

FIG. 3. Comparison of differences in gene expression as determined by cDNA microarray and quantitative reverse transcriptase–polymerase chain reaction analyses. The $\log_2$ ratio (larva/adult) of expression for 22 genes were obtained from least square means from ANOVA procedure. Bars with asterisk indicate significant difference in expression between larval and adult ants.

TABLE 3. Sequenced clones displaying significant differences in expression between larva1 and adult *Camponotus festinatus* ants. Clones are divided into two groups based on whether they were more highly expressed in larvae (larval/adult ratio = L/A > 1) or adults (L/A < 1). The gene and associated E-value for the sequences showing the greatest similarity for each clone are also provided.

Clone	Ratio (L/A)	Best match	E-value
L-A-D10	24.99	*Blatella germanica* major allergen Bla g 1.02	4×10^{-30}
L-B-D5	13.23	*Bombyx mori* cuticle protein	2×10^{-27}
L-A-B6	12.37	*Periplaneta americana* Cr-PII allergen	5×10^{-15}
L-A-G1	8.58	*Anopheles gambiae* ENSANGP00000022136	7×10^{-27}
L-B-D1	6.49	*Schistocerca gregaria* endocuticular protein SgAbd-1	3×10^{-22}
L-B-D3	5.28	*Apis mellifera* hexamerin 70b	2×10^{-39}
L-B-B10	2.94	*Drosophila melanogaster* CG9269-PA	2×10^{-11}
L-A-C1	2.77	*A. gambiae* ENSANGP00000009917	2×10^{-47}
L-B-E10	2.63	*A. mellifera* predicted gene hmm1465	3×10^{-56}
A-D-G9	1.94	*Camponotus festinatus* hexamerin 2	1×10^{-113}
L-A-F6	1.83	*A. mellifera* predicted gene hmm3256	1×10^{-125}
A-C-A1	1.66	*A. mellifera* elongation factor-1 alpha F2	8×10^{-62}
L-A-H3	1.50	*Chironomus tentans* ribosomal YL10 protein homologue	1×10^{-102}
L-B-E7	1.31	*Spodoptera frugiperda* ribosomal protein L8	1×10^{-125}
4 clones	1.40–5.34	No significant homology	$>1 \times 10^{-10}$
A-C-E3	0.0088	*Mus musculus* cellular repressor of E1A-stimulated genes	1×10^{-35}
A-D-D7	0.036	*A. mellifera* predicted gene hmm7788	1×10^{-44}
A-D-C6	0.076	*Todarodes pacificus* cathepsin D	6×10^{-35}
A-D-E6	0.22	*A. mellifera* predicted gene hmm631	1×10^{-28}
A-D-F5	0.25	*A. mellifera* predicted gene hmm2950	8×10^{-52}
A-C-H4	0.26	*Crassostrea gigas* superoxide dismutase	2×10^{-39}
A-D-E11	0.29	*A. mellifera* predicted gene hmm155	1×10^{-101}
A-D-F8	0.32	*Arabidopsis thaliana* gene_id:F1D9.26~unknown protein	1×10^{-23}
A-C-A12	0.43	*D. melanogaster* CG12811-PA	3×10^{-17}
A-C-F12	0.49	*A. mellifera* predicted gene hmm5760	8×10^{-51}
A-C-G2	0.62	*A. mellifera* predicted gene hmm6424	7×10^{-58}
L-A-A3	0.72	*A. mellifera* predicted gene hmm957	2×10^{-45}
A-C-F4	0.77	*Danio rerio* DNA sequence from clone	3×10^{-20}
18 clones	0.027–0.72	No significant homology	$>1 \times 10^{-10}$

TABLE 4. Molecular functions and biological processes assigned to *Camponotus festinatus* cDNA clones under the gene ontology (GO) classification system. Results are based on putative homologies to sequences in *Drosophila melanogaster* (Fban ID and BLASTX match probability are provided). Ratios (larval/adult) of gene expression for which larval and adult values differed significantly after Bonferroni corrections are marked by an asterisk. Genes involved in unknown functions or processes are denoted by a dash.

Clone	*Dmel* ID	*Dmel* gene	E-value	L/A ratio	GO molecular function	GO biological process
L-B-F9	CG4409		1×10^{-10}	23.03*	—	—
L-A-G6	CG7658		3×10^{-14}	15.85*	structural constituent of larval cuticle	—
L-B-D5	CG8505		3×10^{-22}	13.23*	structural constituent of cuticle	—
L-A-F1	CG16885		1×10^{-19}	7.93*	—	—
L-B-D1	CG30045		2×10^{-18}	6.49*	—	—
L-B-D3	CG17285	*Fbp1*	5×10^{-11}	5.28*	protein transporter activity	storage protein uptake
A-D-F12	CG30028	γ*Try*	3×10^{-40}	3.4	—	—
L-B-B10	CG9269		2×10^{-11}	2.94*	—	—
L-A-C1	CG4326	*mRpS17*	2×10^{-47}	2.77*	structural constituent of ribosome	protein biosynthesis
L-B-E10	CG3922	*RpS17*	2×10^{-48}	2.63*	structural constituent of ribosome; nucleic acid binding	protein biosynthesis
L-A-F6	CG3401	β*Tub60D*	1×10^{-132}	1.83*	structural constituent of cytoskeleton; tubulin binding	chromosome segregation; protein transport; mitosis
L-A-D7	CG11808		8×10^{-13}	1.71	—	—
L-A-E10	CG2070		1×10^{-17}	1.68	oxidoreductase activity	—
L-A-D3	CG30392		1×10^{-37}	1.66	—	—
A-C-A1	CG1873	*Ef1α100E*	2×10^{-64}	1.66*	translation elongation factor activity; GTPase activity	regulation of translation; translational elongation
L-A-H6	CG8327		1×10^{-66}	1.54	spermidine synthase activity	—
L-A-H2	CG7939	*RpL32*	2×10^{-54}	1.53	structural constituent of ribosome; nucleic acid binding	protein biosynthesis
L-A-B4	CG9762	*l(3)neo18*	1×10^{-17}	1.38	NADH dehydrogenase activity	mitochondrial electron transport
A-C-A5	CG14314		8×10^{-17}	1.33	—	—
L-B-E7	CG1263	*RpL8*	1×10^{-128}	1.32*	structural constituent of ribosome; RNA binding	protein biosynthesis
A-C-D1	CG3314	*RpL7A*	1×10^{-63}	1.31	structural constituent of ribosome; nucleic acid binding	protein biosynthesis
L-A-A4	CG4111	*RpL35*	3×10^{-25}	1.24	structural constituent of ribosome; nucleic acid binding	protein biosynthesis
L-A-D1	CG4651	*RpL13*	3×10^{-70}	1.21	structural constituent of ribosome; nucleic acid binding	protein biosynthesis
L-A-F2	CG5920	*sop*	5×10^{-99}	1.21	structural constituent of ribosome; nucleic acid binding	protein biosynthesis
A-C-B7	CG14792	*sta*	2×10^{-72}	1.17	structural constituent of ribosome; nucleic acid binding	protein biosynthesis; protein metabolism
A-C-C2	CG7490	*RpLP0*	2×10^{-97}	1.17	structural constituent of ribosome; RNA binding	protein biosynthesis
L-A-A2	CG1646		1×10^{-19}	1.15	pre-mRNA splicing factor activity	nuclear mRNA splicing
A-C-F11	CG40049		1×10^{-40}	1.14	—	—
A-C-A10	CG31305		3×10^{-34}	1.11	citrate transporter activity	mitochondrial citrate transport
A-C-D11	CG12157	*Tom40*	3×10^{-61}	1.11	protein translocase activity; carrier activity	protein-mitochondrial targeting
A-C-H3	CG4464	*RpS19*	1×10^{-54}	1.05	structural constituent of ribosome; nucleic acid binding	protein biosynthesis
	CG7070			0.88	pyruvate kinase activity; carbohydrate kinase activity	glycolysis
L-A-C3		*PyK*	4×10^{-36}			
L-A-B11	CG32145	*ome*	5×10^{-70}	0.84	dipeptidyl-peptidase IV activity	signal transduction; proteolysis and peptidolysis
A-C-D8	CG6838		1×10^{-16}	0.81	ARF GTPase activator activity	intracellular protein transport; vesicle-mediated transport
A-C-G5	CG3246		2×10^{-42}	0.79	—	—
L-A-A3	CG11314		6×10^{-27}	0.72*	—	mesoderm development
A-C-G6	CG5939	*Prm*	4×10^{-85}	0.63	structural constituent of muscle; motor activity	mesoderm development
A-C-G2	CG5799	*dve*	7×10^{-33}	0.62	AT DNA binding; transcription factor activity	regulation of transcription; wing morphogenesis
A-C-G10	CG18143		4×10^{-56}	0.6	guanine deaminase activity	purine base metabolism
A-C-E11	CG11390	*PebIII*	1×10^{-38}	0.54	carrier activity	sensory perception
A-C-B4	CG18519		4×10^{-47}	0.53	oxidoreductase activity	defense response; nucleic acid metabolism

TABLE 4. Continued.

Clone	*Dmel* ID	*Dmel* gene	E-value	L/A ratio	GO molecular function	GO biological process
A-C-F5	CG1915	*sls*	2×10^{-12}	0.52	structural constituent of cytoskeleton; kinase activity	cytoskeleton organization; muscle contraction
A-C-F12	CG2145		2×10^{-29}	0.49*	serine-type peptidase activity	—
A-C-F7	CG4696	*Mp20*	8×10^{-79}	0.46	structural constituent of cytoskeleton; actin binding	muscle contraction
A-C-A12	CG12811		2×10^{-17}	0.43*	—	—
A-D-F8	CG2839		2×10^{-12}	0.32*	—	development; gametogenesis
A-C-B8	CG11715	*Cyp4g15*	8×10^{-94}	0.3	electron transporter activity; oxidoreductase activity	steroid biosynthesis
A-D-E11	CG4843	*Tm2*	2×10^{-92}	0.29*	actin binding; structural constituent of cytoskeleton	muscle contraction; pole plasm assembly
A-C-H4	CG9027		7×10^{-36}	0.26*	superoxide dismutase act	defense response
A-D-F5	CG7178	*wupA*	2×10^{-58}	0.25*	structural constituent of cytoskeleton; actin binding	muscle contraction
A-D-E6	CG14207		3×10^{-22}	0.22*	—	—
A-D-C6	CG1548	*cathD*	2×10^{-37}	0.08*	cathepsin D activity	proteolysis and peptidolysis
A-D-D7	CG7390	*smp-30*	1×10^{-31}	0.04*	calcium ion binding	anterior/posterior axis specification; signal transduction
A-C-E3	CG5413	*CREG*	5×10^{-17}	0.01*	—	—

profiles from the two development stages differed substantially.

We used information from the gene ontology (GO) classifications, which provide a vocabulary for classifying genes based on their molecular functions, cellular components, and biological processes (Gene Ontology Consortium 2000), to further understand the differences in expression patterns between larval and adult ants. We considered patterns of expression only for genes that showed significant similarity to genes in *D. melanogaster*, the insect taxon in which gene function is best understood. GO classifications were determined by their suggested function in *D. melanogaster* as described through the FlyBase Genome Annotation Database (Flybase Consortium 2003).

Many *C. festinatus* genes showed similarity to *D. melanogaster* genes and were involved in a particular biological process or possessed a specific molecular function under the GO classification system (Table 4). Genes that were upregulated in larvae were often involved in biological processes associated with protein metabolism. Those that were upregulated in adults were more varied in their activities and functions.

Similarity to Drosophila melanogaster Expression Patterns

We compared the patterns of gene expression observed in the ant *C. festinatus* to those found in a previous study of development in *D. melanogaster* to determine if homologous genes showed similar patterns of expression in the two insect taxa. In the *Drosophila* study, variation in gene expression was monitored using cDNA microarray technology and assayed from egg through late adulthood (Arbeitman et al. 2002). In total, 27 of the genes printed on the array in this study showed strong ($E < 1 \times 10^{-10}$) similarity to genes from *D. melanogaster* that were included in the study of Arbeitman et al. (2002).

Surprisingly, relatively little relationship existed between the observed ratios of expression in *C. festinatus* and *D. melanogaster*, regardless of whether the time-point adjusted ratios (late larval stage before prepupation [L96h] and a late female adult stage [Af25d], which were expected to best approximate the life stages examined in this study) or the overall mean-stage-specific ratios were used (Table 5). Indeed, Pearson's correlation between the ratio of expression in *C. festinatus* and the mean ratio and time-point adjusted ratio in *D. melanogaster* was -0.17 ($P = 0.40$) and 0.10 ($P = 0.61$), respectively. That is, genes that were more highly expressed in *C. festinatus* larvae than adults were not necessarily more highly expressed in *D. melanogaster* larvae than adults. This lack of relationship was even observed among genes showing significant variation in expression between the larval and adult *C. festinatus* stages.

To further understand differences between larval and adult holometabolous insects, we examined the functions of *D. melanogaster* genes that differed in expression between late larval (L96h) and late female adult (Af25d) stages (Arbeitman et al. 2002). Our intention was to compare gross-level function of genes showing expression differences between life stages in *C. festinatus* and *D. melanogaster*. We examined genes that showed strong (at least eight-fold) difference in expression between larvae and adult *D. melanogaster* samples. We then compared the distribution of GO molecular functions and biological processes for genes relatively upregulated in the two stages.

In total, 67 genes were at least eight-fold more highly expressed in larval than adult *D. melanogaster*, whereas 27 genes were at least eight-fold more highly expressed in adult than larvae *D. melanogaster*. Of the total of 94 genes, 59 and 52 possessed some acknowledged molecular function and biological process, respectively, as determined by the GO classification system. The distributions of molecular functions of genes upregulated in larvae and adults differed significantly ($G_7 = 29.54$; $P < 0.0001$). We found that genes

TABLE 5. Comparison of expression patterns between *Drosophila melanogaster* and *Camponotus festinatus*. The larval/adult ratio of expression in *C. festinatus* is given for each clone. In addition, the mean expression of all time points from *D. melanogaster* larval and adult stages and the ratio of expression for *D. melanogaster* late larval and late adult time points are provided for comparison (see text for details). Ratios for which larval and adult expression differed significantly between *C. festinatus* life stages after Bonferroni corrections are marked by an asterisk.

C. festinatus clone	*D. melanogaster* homologue	*C. festinatus* ratio	*D. melanogaster* mean ratio	*D. melanogaster* late ratio
A-D-C6	CG1548	0.08*	0.44	1.28
A-C-E3	CG5413	0.01*	0.52	1.43
A-D-E6	CG14207	0.22*	1.86	1.98
A-D-F5	CG7178	0.25*	2.73	1.67
A-C-H4	CG9027	0.26*	1.15	1.67
A-C-B8	CG11715	0.30	1.84	2.32
A-C-F12	CG2145	0.49	1.76	4.24
A-C-F5	CG1915	0.52	1.05	1.75
A-C-E11	CG11390	0.54	1.71	2.64
A-C-G10	CG18143	0.60	0.73	0.76
A-C-G6	CG5939	0.63	2.46	6.58
A-C-G5	CG3246	0.79	1.72	0.86
A-C-D8	CG6838	0.81	0.70	0.62
L-A-C3	CG7070	0.88	0.71	1.09
A-C-C2	CG7490	1.17	0.96	0.46
A-C-B7	CG14792	1.17	0.77	0.35
L-A-A2	CG1646	1.15	1.04	0.59
L-A-F2	CG5920	1.21	0.74	0.48
A-C-D1	CG3314	1.31	0.90	0.49
L-B-E7	CG1263	1.32*	1.33	0.66
L-A-B4	CG9762	1.38	1.21	0.87
L-A-H2	CG7939	1.53	2.14	1.16
L-A-H6	CG8327	1.54	1.30	0.81
L-A-E10	CG2070	1.68	2.12	2.14
L-A-D7	CG11808	1.71	1.73	1.59
L-B-E10	CG3922	2.63*	1.50	0.76
L-B-D5	CG8505	7.25	0.63	3.48

more highly expressed in larvae showed less diversity in molecular function than the genes more highly expressed in adults. Specifically, the 38 genes upregulated in larvae fell into five major functional categories and displayed an associated diversity of 0.504. In contrast, the 21 genes upregulated in adults fell into seven major functional categories and exhibited a higher diversity of 0.795. Similar results were observed for the GO biological processes associated with genes upregulated in larval and adult *D. melanogaster*. The diversity measure for genes more highly expressed in larvae and adults equaled 0.428 and 0.686, respectively, reflecting the more specialized nature of highly expressed genes in larvae and the more diverse nature of highly expressed genes in adults.

DISCUSSION

Holometabolous insects are an ancient and highly successful group of organisms. In part, the success of the Holometabola stems from their ability to occupy different ecological niches in their immature (larval) and mature (adult) developmental stages (Sehnal et al. 1996). A better understanding of the molecular basis of insect development may help us comprehend why these organisms have become so successful. As a first attempt at addressing this question, we examined variation in gene expression as ants developed from larvae into adults and compared these patterns of expression to comparable patterns observed in the well-studied dipteran *D. melanogaster*.

Our primary analysis of sequenced clones from cDNA libraries obtained from the ant *C. festinatus* revealed expected homologies to other taxa (Fig. 1). Many of the *C. festinatus* genes showed strongest similarity to sequences obtained from the honeybee *A. mellifera*, as anticipated given that both *A. mellifera* and *C. festinatus* are within the order Hymenoptera. Most other clones matched sequences from other insects, such as *D. melanogaster* and *Anopheles gambiae*, whose genomes have been completely sequenced. A small proportion of clones exhibited strongest similarity to vertebrate, bacterial, plant, or fungal sequences. Some of these sequences may represent *C. festinatus* genes that are highly diverged from other insects. Another plausible explanation, however, is that these divergent sequences originate from commensal, parasitic, or symbiotic species that exist with or within the ants.

Approximately 40% of sequences from each library did not show significant similarity to any sequences in GenBank. These results are consistent with at least one other genomic study from an aculeate hymenopteran. An intensive survey of expressed sequence tags from *A. mellifera* found that 61% of clones failed to show significant similarity to protein sequences in the nonredundant protein (nr) database (Whitfield et al. 2002). Many of the *C. festinatus* sequences that failed to show matches to known sequences were relatively short or high in AT content (60–70% AT, data not shown), suggesting that they may be truncated sequences or contain noncoding sequences, such as 3′ untranslated regions. Such regions likely evolve relatively quickly because of lack of functional constraint and, therefore, would be difficult to align to existing sequences found within GenBank (Li 1997).

The distribution of clones showing sequence homology to various taxa differed significantly between the L/P and A libraries. This difference resulted from the proportion of *C. festinatus* clones that showed strongest homology to ‘‘*A. mellifera*’’ versus ‘‘other insect’’ (Fig. 1). The observed variation may have resulted from differential selection acting on genes expressed in distinct life stages. However, we note that many sequences in the other-insect category displayed homology to sequences in other Hymenoptera (typically other ants or wasps). Therefore the significant difference may not reflect differential evolutionary processes in the life stages. Rigorously testing any hypothesis related to differential selection on genes expressed at various developmental stages will necessitate the sequencing of many more genes.

We discovered that many genes show substantial differences in expression between ant adult and larval life stages. In total, 24% of the clones printed on our array showed significant differences in expression using conservative criteria (Fig. 2). Moreover, differences in expression observed using cDNA microarray technology were confirmed in several genes using quantitative real-time PCR (Fig. 3). Consequently, our data suggest that a majority of genes studied will show distinct expression patterns between larval and adult ants. Similar results were found in *D. melanogaster*, where 86% of genes studied showed significant variation during development (White et al. 1999; Arbeitman et al. 2002).

The patterns of gene expression observed in this study

showed some agreement with expectations derived from the proposed functions of the genes. For example, a conspicuous result detailed in Table 4 is that many genes upregulated in larvae are apparently related to protein biosynthesis. Several of the putative *D. melanogaster* homologues, such as *mRpS17, RpS17, RpL32, RpL8, RpL7A, RpL35, RpL13, sop, sta, RpLP0,* and *RpS19,* code for ribosomal proteins. The putative homologue *Ef1α100E,* is also associated with protein production, while the genes *Fbp1* and *Tom40* possess protein transport or translocase activity. In addition, three structural genes were upregulated in larvae, including β*Tub60D* and two other genes that have not been described in detail in *D. melanogaster* (gene symbols CG7658 and CG8505). Upregulation of genes involved in transcription, translation, or protein production in larvae relative to adults was expected, because the larval stage in holometabolous insects is associated with rapid growth (Hölldobler and Wilson 1990; Heming 2003).

In contrast to the genes that were upregulated in larvae, those upregulated in adults tended to have more diverse function. Somewhat surprisingly several genes expected to be involved in *D. melanogaster* development, such as CG11314, *Prm, smp-30,* were upregulated in adult ants. It is possible, however, that these gene products have functions in adults, in addition to those in larvae. Other genes upregulated in adults were associated with adult-specific functions, such as sensory perception (*PebIII*) and gametogenesis (CG2839). In addition, three genes were partly associated with muscle development or function (*Mp20, wupA,* and *Tm2*). Many of the other putative *D. melanogaster* homologues displayed a variety of activities associated with processes ranging from proteolysis to biosynthesis. The diversity of genes upregulated in adult ants may reflect the far more complex environment in which adult ants exist. Whereas ant larvae normally remain sedentary, adult ants must participate in a variety of tasks such as food acquisition, colony defense, and nest construction (Hölldobler and Wilson 1990). Indeed, this hypothesis gains some support from comparison of functions of genes upregulated in larval and adult *D. melanogaster*. Our analyses indicated that genes highly expressed in larval *D. melanogaster* displayed a lower diversity of functions than those more highly expressed in adult.

One of the central objectives of this study was to compare patterns of expression of genes in *C. festinatus* to homologous genes in *D. melanogaster*. We obtained data on the expression patterns of 27 putative *D. melanogater* homologues over the same developmental period considered in this study (Arbeitman et al. 2002). Surprisingly, we found relatively little accord between the observed patterns for homologous genes in these two holometabolous insects, despite the qualitative similarity in the types of genes expressed (see above). It is unlikely that the difference resulted from the small number of genes analyzed. The sample size was large enough to detect a significant correlation if one existed. For example, the power of our test $(1 - \beta)$ for the correlation between the larval to adult ratios of *C. festinatus* genes and time-point-adjusted *D. melanogaster* genes is 0.87 (Zar 1999). It is possible that homologous genes or matching life stages may not have been compared. Alternatively, the genes may be expressed in the same manner in homologous tissues of *C. festinatus* and *D. melanogaster* larvae and adults, but the tissues may make up different proportions of individuals in the two life stages in the two taxa.

However, it is also likely that the genes may have taken on distinct functions over the course of evolution or possess the same function but show differential expression patterns. Although both *D. melanogaster* and *C. festinatus* are holometabolous insects, they are very distantly related. Their most recent common ancestor is believed to have existed approximately 300 million years ago (Labandeira 1998). Indeed, gene expression appears to evolve rapidly in diverse taxa (Gu et al. 2002; Oleksiak et al. 2002; Meiklejohn et al. 2003; Ranz et al. 2003; Rifkin et al. 2003; Townsend et al. 2003; Wray et al. 2003), so it is quite possible that homologous gene products could be expressed in different life stages or that the products have distinct or divergent functions in the two taxa. We anticipate that future comparisons of gene expression across insect taxa will help unravel these puzzling observations and contribute substantially to our understanding of development (Gibson 2002).

Acknowledgments

This research was supported in part by a National Institutes of Health Postdoctoral Excellence in Research and Teaching fellowship administered through the University of Arizona Center for Insect Science (GM00708-02). We thank C. Vanier for statistical assistance and the University of Arizona Genomic Analysis and Technology Core for assistance in array fabrication and analysis.

Literature Cited

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. J. Mol. Biol. 215: 403–410.

Arbeitman, M. N., E. E. M. Furlong, F. Imam, E. Johnson, B. H. Null, B. S. Baker, M. A. Krasnow, M. P. Scott, R. W. Davis, and K. P. White. 2002. Gene expression during the life cycle of *Drosophila melanogaster*. Science 297:2270–2275.

Bolstad, B. M., R. A. Irizarry, M. Astrand, and T. P. Speed. 2003. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 19:185–193.

Bustin, S. A. 2002. Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): trends and problems. J. Mol. Endocrinol. 29:23–39.

Evans, J. D., and D. E. Wheeler. 1999. Differential gene expression between developing queens and workers in the honey bee, *Apis mellifera*. Proc. Natl. Acad. Sci. USA 96:5575–5580.

———. 2000. Expression profiles during honeybee caste determination. Genome Biol. 2:1–6.

Flybase Consortium. 2003. The FlyBase database of the *Drosophila* genome projects and community literature. Nucleic Acids Res. 31:172–175.

Freeman, W. M., S. J. Walker, and K. E. Vrana. 1999. Quantitative RT-PCR: pitfalls and potential. BioTechniques 26:112–125.

Gene Ontology Consortium. 2000. Gene ontology: tool for the unification of biology. Nat. Genet. 25:25–29.

Gibson, G. 2002. Microarrays in ecology and evolution: a preview. Mol. Ecol. 11:17–24.

Gilbert, L. I., R. Rybczynski, and S. S. Tobe. 1996a. Endocrine cascade in insect metamorphosis. Pp. 59–108 *in* L. I. Gilbert, J. R. Tata, and B. G. Atkinson, eds. Metamorphosis: postembryonic reprogramming of gene expression in amphibian and insect cells. Academic Press, New York.

Gilbert, L. I., J. R. Tata, and B. G. Atkinson. 1996b. Metamorphosis:

postembryonic reprogramming of gene expression in amphibian and insect cells. Academic Press, New York.
Gu, Z., D. Nicolae, H. H.-S. Lu, and W.-H. Li. 2002. Rapid divergence in expression between duplicate genes inferred from microarray data. Trends Genet. 18:609–613.
Hartfelder, K., and W. Engels. 1998. Social insect polymorphism: hormonal regulation of plasticity in development and reproduction in the honeybee. Curr. Topics Dev. Biol. 40:45–77.
Heming, B. S. 2003. Insect development and evolution. Comstock Publishing Associates, Ithaca, NY.
Hölldobler, B., and E. O. Wilson. 1990. The ants. The Belknap Press of Harvard Univ. Press, Cambridge, MA.
Hunt, J. H., and R. R. Snelling. 1975. A checklist of the ants of Arizona. J. Ariz. Acad. Sci. 10:20–23.
Kerr, M. K., C. A. Afshari, L. Bennett, P. Bushel, J. Martinez, N. J. Walker, and G. A. Churchill. 2002. Statistical analysis of gene expression microarray experiment with replication. Stat. Sinica 12:203–217.
Klein, D. 2002. Quantification using real-time PCR technology: applications and limitations. Trends Mol. Med. 8:257–260.
Kristensen, N. P. 1991. Phylogeny of extant hexapods. Pp. 125–139 *in* I. D. Naumann, ed. Insects of Australia. Cornell Univ. Press, Ithaca, NY.
Labandeira, C. C. 1998. Early history of arthropod and vascular plant associations. Annu. Rev. Earth Planetary Sci. 26:329–377.
Li, W.-H. 1997. Molecular evolution. Sinauer Associates, Sunderland, MA.
Meiklejohn, C. D., J. Parsch, J. M. Ranz, and D. L. Hartl. 2003. Rapid evolution of male-biased gene expression in *Drosophila*. Proc. Natl. Acad. Sci. USA 100:9894–9899.
Nei, M. 1987. Molecular evolutionary genetics. Columbia Univ. Press, New York.
Oleksiak, M. F., G. A. Churchill, and D. L. Crawford. 2002. Variation in gene expression within and among natural populations. Nat. Genet. 32:261–266.
Ranz, J. M., C. I. Castillo-Davis, C. D. Meiklejohn, and D. L. Hartl. 2003. Sex-dependent gene expression and evolution of the *Drosophila* trascriptome. Science 300:1742–1745.
Rifkin, S. A., J. Kim, and K. P. White. 2003. Evolution of gene expression in the *Drosophila melanogaster* subgroup. Nat. Genet. 33:138–144.
Ross, K. G., and R. W. Matthews. 1991. The social biology of wasps. Comstock Publishing Associates, Ithaca, NY.
Seeley, T. D. 1995. The wisdom of the hive. Harvard Univ. Press, Cambridge, MA.
Sehnal, F., P. Svacha, and J. Zrzavy. 1996. Evolution of insect metamorphosis. Pp. 3–58 *in* L. I. Gilbert, J. R. Tata, and B. G. Atkinson, eds. Metamorphosis: postembryonic reprogramming of gene expression in amphibian and insect cells. Academic Press, New York.
Sokal, R. R., and F. J. Rohlf. 1995. Biometry. W. H. Freeman, New York.
Townsend, J. P., D. Cavalieri, and D. L. Hartl. 2003. Population genetic variation in genome-wide gene expression. Mol. Biol. Evol. 20:955–963.
Truman, J. W., and L. M. Riddiford. 1999. The origin of insect metamorphosis. Nature 401:447–452.
———. 2002. Endocrine insights into the evolution of metamorphosis in insects. Annu. Rev. Entomol. 47:467–500.
White, K. P., S. A. Rifkin, P. Hurban, and D. S. Hogness. 1999. Microarray analysis of *Drosophila* development during metamorphosis. Science 286:2179–2184.
Whitfield, C. W., M. R. Band, M. F. Bonaldo, C. G. Kumar, L. Liu, J. R. Pardinas, H. M. Robertson, M. B. Soares, and G. E. Robinson. 2002. Annotated expressed sequence tags and cDNA microarrays for studies of brain and behavior in the honey bee. Genome Res. 12:555–566.
Whiting, M. F. 2002. Phylogeny of the holometabolous insect orders: molecular evidence. Zool. Scr. 31:3–15.
Wolfinger, R. D., G. Gibson, E. D. Wolfinger, L. Bennett, H. Hamadeh, P. Bushel, C. Afshari, and R. S. Paules. 2001. Assessing gene significance from cDNA microarry expression data via mixed models. J. Comp. Biol. 8:625–637.
Wray, G. A., M. W. Hahn, E. Abouheif, J. P. Balhoff, M. Pizer, M. V. Rockman, and L. A. Romano. 2003. The evolution of trascriptional regulation in eukaryotes. Mol. Biol. Evol. 20: 1377–1419.
Zar, J. H. 1999. Biostatistical analysis. Prentice-Hall, Upper Saddle River, NJ.

Corresponding Editor: K. Shaw