Spring 2007

Analysis of OS-E and OS-F expression in Drosophila melanogaster using GAL4 transgenes with green fluorescent protein

David Anderson
Carroll College, Helena, MT

Follow this and additional works at: https://scholars.carroll.edu/lifesci_theses

Part of the Molecular Genetics Commons

Recommended Citation
https://scholars.carroll.edu/lifesci_theses/166

This Thesis is brought to you for free and open access by the Life and Environmental Sciences at Carroll Scholars. It has been accepted for inclusion in Life and Environmental Sciences Undergraduate Theses by an authorized administrator of Carroll Scholars. For more information, please contact tkratz@carroll.edu.
Analysis of OS-E and OS-F expression in *Drosophila melanogaster* using GAL4 transgenes with green fluorescent protein.

Submitted in partial fulfillment of the requirements for graduation with honors to the Department of Natural Sciences at Carroll College, Helena, Montana

David Anderson

2 April 2007
This thesis for honors recognition has been approved for the Department of Natural Sciences by:

Dr. Marilyn Schendel, Thesis Director

Dr. John Addis, Reader

Dr. Ron Stottlemeyer, Reader
Abstract:

The genes OS-E and OS-F are expressed in the sensory nerve network of *D. melanogaster*’s antennae and maxillary palps and are putative odorant-binding proteins. A transgenic construct used GAL4 to drive production of GFP in tissues where the gene of interest was being transcribed. Each gene drove production of GFP at a different time in pupal development. OS-E was found to fluoresce in sensory structures at 60 hours after pupation and OS-F fluoresced at 90 hours. Real-time quantitative PCR was used to study flies without gene insertions to determine the molecular nature of the expression of the two genes and consequently whether the developmental model shown by GFP was indicative of a natural process. PCR data were taken from a developmental time period when OS-E/GFP production was active and OS-F/GFP was not; these were compared to samples from adult flies. Concentrations of OS-E mRNA and OS-F mRNA were found to be surprisingly similar, indicating either that the GFP model is inaccurate, or that problems with the PCR reactions caused misrepresentative results.
Introduction:

Much about the genetic and molecular nature of olfaction is still not understood. Traditionally, insects have proven useful models for the study of receptor-ligand relationships, pheromone effects, neural development, and many studies of insect olfaction have prompted similar research in vertebrates (Hansson, 2002). Insects have been favored largely for the sensitivity and simplicity of their olfactory systems, especially D. melanogaster which has $10^3$ receptor neurons in its antennae, compared to $10^8$ in the human nose (Carlson, 1991).

D. melanogaster has olfactory organs on its antennae and maxillary palps that are divided into three groups based on morphology: coeloconic sensilla, which are short, conical structures used to detect food and plant odors; basiconic sensilla, which project a short distance from the sensory organs and process the same odors as coeloconic sensilla; and trichoid sensilla, the longest of the three, implicated in pheromone detection (Vosshall, 2000). Olfactory sensilla are perforated with pore tubuoles through which odorant molecules can gain access to the sensillum lumen where odorant-binding proteins transport them to sensory neurons (Vogt, 1985).

McKenna et al. (1994) identified five genes using subtractive hybridization, naming them OS-B, C, D, E, and F. Heads from wild type D. melanogaster from the CS5 breeding stock were frozen, and the chemosensory organs were removed. Messenger RNA (mRNA) was extracted from the antennae and hybridized against mRNA isolated from heads with the olfactory and chemosensory organs removed. A library of 45,000 clones was identified from the antennal sequences that did not hybridize with sequences from the rest of the head. These clones were assumed to be uniquely
expressed in antennae. The library was divided into five groups by the propensity of a group to hybridize with itself, but not with any of the other four, indicating that the sequences in a given class were more similar to one another than to sequences in the other classes. Groups OS-E and OS-F were sequenced and identified as a tightly linked pair of protein-encoding genes with similar sequences to known olfactory proteins. The two genes are 72% identical to each other and OS-F is 29% similar to the pheromone-binding protein APR1 in the moth *A. pernyi*. These two genes were hypothesized to be the result of a duplication due to their proximity and high sequence identity.

The Vogt lab, in which this research was conducted, is largely focused on the study of the tobacco hawk moth, *Manducca sexta*; OS-E and OS-F were chosen in part for their close resemblance to moth odorant-binding proteins. *M. sexta* is a popular subject for histological studies, and it has been shown to carry homologs of the genes studied here (Vogt, unpublished), but *D. melanogaster* is much more conducive to genetic manipulation, namely GAL4/UAS-driven gene expression (Duffy, 2002), and has long been the only reliable candidate for transgene addition (Atkinson et al., 2001). Though workers have recently developed viable transgenic architectures in the mosquito *A. gambiae* (O'Brochta, 2003) among others, the *Drosophila* system was the most reliable and cost-effective solution at the time.

The GAL4 system for gene expression was developed by Brand and Perrimon (1993) for use in *Drosophila* because the existing system had inherent limitations that made it valuable only for observing genes that were expressed equally in every tissue. Prior to the GAL4 system, transgenic animals were constructed with promoters from the heat-shock system, which regulates gene transcription in response to extracellular factors,
such as high temperature or presence of pathogens. The problem with this system was that HSF, the system's transcriptional activator, was produced at a constitutive level that, while low, confounded fluorescence-based studies. When HSF was activated, it would transcribe genes all over the body, making it impossible to analyze a single structure or system. The GAL4 gene, isolated from yeast, can be inserted at random in the genome and then screened for successful insertions where the sequence is immediately downstream from the gene of interest and is controlled by the same promoter. GAL4 encodes a protein that binds to an upstream activating sequence (UAS) that needs to be inserted immediately upstream from the gene targeted for ectopic expression which, in this case, is green fluorescent protein (GFP). The UAS/GFP sequence pair can be inserted anywhere in the genome because GAL4 expression is very localized. When the gene of interest is expressed in a transgenic fly with all three gene insertions, the GAL4 protein is also produced. GAL4 then binds to the UAS region on any surrounding DNA and promotes the production of GFP in the immediate area. For example, this system would cause antennae and other sensory structures to fluoresce when olfactory genes began to transcribe in those places during fly development. Unlike the heat-shock system, the UAS-GFP response, which causes a visible fluorescence, can produce a very localized result making it valuable for observing neural development.

R.G. Vogt (unpublished) raised D. melanogaster from the w1118 stock with GAL4-UAS/GFP insertions controlled by the OS-E and OS-F promoters. Flies were raised in jars with food and pupae were collected every twelve hours and analyzed periodically under a fluorescent microscope to determine a time line for the expression of the two genes. Workers observed the fluorescent activity of OS-E at 60 hours after pupation and
OS-F at 90 hours after pupation: these expression times were accurate within 12 hours. This finding was interesting given that the control regions of the two genes are nearly identical.

At this point there was no information to indicate that gene expression patterns in transgenic lines of *D. melanogaster* were also observed in wild-type animals. While transgenic animals are a valuable tool for analyzing gene expression, allowing researchers to observe activity at a single locus in great detail, it is unwise to assume that all observed phenomena are paralleled in organisms that have not been genetically manipulated. It is possible that data obtained using transgenic flies could be unique to those flies. Sanders Williams and Wagner (2000) outline several possible problems with gene insertion, such as difficulty controlling dosage due to the somewhat random distribution of UAS/GFP inserts, insertional mutagenesis caused by inserting a sequence in the middle of a necessary gene, and insertion into a heterochromatic region causing transcriptional silencing, all of which produce inconsistent phenotypic variation that can confound results.

For the purposes of my research, it was necessary to demonstrate that there was a significant difference on a developmental time line in the expression of the two genes in a naturally occurring (non-transgenic) fly. My first objective was to confirm that OS-E would fluoresce in sensory structures at 60% of development (60 hours after pupation) and OS-F would fluoresce at 90% of development. My second objective was to test these observations in flies without gene insertions on the molecular level. To accomplish this, messenger RNA was extracted from adult flies and pupal samples at 60, 75, and 90 hours after pupation, cDNA was made, and quantitative PCR was used to determine the relative
concentrations of each gene product. My hypothesis was that the qPCR data would confirm that the OS-E gene product was present in significant quantity at 60 and 75 hours after pupation, and that the OS-F gene product was not.

**Materials and Methods:**

**Crosses:** Virgin female *D. melanogaster* of the w¹¹¹⁸ breeding stock with UAS gene insertions to drive expression of GFP were crossed with two different groups of males, one had a GAL4 insertion at OS-E and the other had the same insertion at OS-F. Larvae from the crosses were raised in an incubator at 25°C. Every 12 hours all new pupae were removed to separate containers. The approximate observed time between pupation and emergence was 105 hours (+/-12 hours). Fly heads were examined for fluorescence at 50, 60, 75, and 90 hours after pupation to provide a visual time-line for expression of GFP. The heads of pupae were collected at 60, 75, and 90 hours after pupation and frozen at -80°C until RNA extraction was performed.

**RNA Extraction:** The collected heads were kept frozen with liquid nitrogen and homogenized with a mortar and pestle. A mixture of 500μL guanidinium thiocyanate, 50μL sodium acetate, 100μL chloroform isoamyl, and 500μL water-saturated phenol was added to the homogenate and centrifuged for 20 minutes at 20,000Xg at 4°C. The top (aqueous) layer was added to 500μL isopropanol and stored at -20°C for one hour. The solution was then centrifuged at 20,000Xg at 4°C for 40 minutes, and the resultant pellet was washed with 500μL of 75% ethanol and dried under a vacuum for five minutes. The pellet, which contained the extracted RNA, was resuspended in 50μL distilled water that
had been treated with diethylpyrocarbonate (DEPC water) to remove RNase and stored at -20°C.

Synthesis of cDNA (using Invitrogen Cat# 11904-018): A 5μL aliquot of the RNA suspension was mixed with 1μL dNTP mix, 1μL Oligo(dT)12-18(0.5g/L), and 7μL DEPC water, and this solution was incubated at 65°C for five minutes, then iced for one minute. Two μL 10X RT buffer, 4μL MgCl₂ (25mM), 2μL DTT, and 1μL RNase OUT were then added, and the reaction mixture was incubated at 42°C for 2 minutes. One μL of Taq polymerase was added, and the mixture was incubated at 42°C for 50 minutes. DNA synthesis was terminated by heat inactivation at 70°C for 15 minutes, and the addition of 1μL RNase H was followed by a 20 minute incubation at 37°C. The cDNA was stored at -80°C.

Real-time qPCR: One μL of cDNA (either OS-E, OS-F, or ribosomal) was added to a mix containing 10μL SyberGreen, 2μL gene specific primers (1μL sense, 1μL antisense), and 7μL DEPC water. This mixture was loaded into a 96-well plate and assayed in triplicate using the BioRad iCycler iQ Real-Time PCR System which optically analyzes each reaction well after every temperature cycle and records the data automatically. The SyberGreen solution shows a visible response to DNA synthesis that can be analyzed quantitatively.

Results:

When transgenic flies were observed to verify the existing data (Vogt,
unpublished), the OS-E/GFP flies began to fluoresce at 60 hours after pupation, to increase in fluorescence until 80 hours after pupation, and to maintain the same appearance (suggesting a stable amount of GFP) into adulthood. The OS-F/GFP samples, however, showed weak fluorescence beginning at 90 hours and grew steadily brighter throughout their lifespan.

To confirm that the fluorescence was indicative of gene expression in the absence of genetic manipulation, molecular analysis of w^{1118} flies without the GAL4-UAS insert was attempted via RNA extraction and cDNA synthesis followed by quantitative PCR (qPCR). Successful qPCR analysis would determine whether the OS-E and OS-F genes were being transcribed in significant quantity at the above indicated times.

![Agarose gel electrophoresis](image)

*Figure 1: Agarose gel electrophoresis of cDNA from 60-hour OS-E and OS-F samples made with full length primers for each gene to confirm successful RNA extraction and cDNA synthesis. Adult samples are not pictured.*

Heads were removed from wild type *Drosophila* adults and pupae at 60, 75, and 90 hours after pupation, and RNA was extracted. For reliability during PCR, cDNA was
made from the putative RNA extractions, and the four samples were tested for quantity and purity in an agarose gel (Figure 1). The 60-hour, 75-hour, and adult samples (adult sample absent from Figure 1) indicated successful cDNA synthesis, so those three extractions were analyzed with qPCR.

Real-time quantitative PCR was used to determine the relative concentration of each cDNA sample. After each amplification cycle, the samples were automatically measured for a concentration-dependent visual indicator, and data were analyzed based on the number of amplification cycles each sample underwent before achieving the threshold concentration for visual response (Figure 2).

The concentrations of OS-E and OS-F mRNA were compared to the concentration of ribosomal RNA in each sample. The rRNA was used to normalize measurements because the RNA extraction process can yield varying amounts of product, and rRNA was assumed to be in constant concentration since ribosomes

![Figure 2: Relative amounts of cDNA](image)

*Figure 2: Amounts of cDNA measured in number of PCR amplification cycles required for each sample to reach the concentration of unamplified cDNA from rRNA.*
are synthesized at the same rate in the cell throughout development. Since qPCR is sensitive to the different initial concentrations of the samples, data are measured in units “less than rRNA” in an attempt to correct for the significant difference in sample concentration visible where the OS-E band is much brighter than the OS-F band (Figure 1). The y-axis in Figure 2 is consequently inverted with respect to concentration, i.e. the point at y=11 is read: “60-hour OS-E cDNA underwent 11 PCR cycles more than rRNA to achieve the same concentration,” or “there is $2^{11}$ times less 60-hour OS-E cDNA than rRNA in the given sample.” Distribution on the x-axis has no representational significance.

Real-time qPCR amplification of w^{1118} flies without gene insertions indicated that the 60-hour OS-E and OS-F samples had relatively similar amounts of mRNA at $2^{11}$ and $2^{10.26}$ times fewer copies than rRNA, respectively (Figure 2). The 75-hour samples did not yield results. The 60-hour OS-F sample was 1.7 times more abundant than OS-E at 60 hours while the adult samples showed a more significant difference: OS-F was about 7.4 times more abundant than OS-E, reflecting a rate of increase that was faster than OS-E by a factor of 6.7.

**Discussion:**

The purpose of this research was to determine if a correlation existed between antennal fluorescence in transgenic *D. melanogaster* and the presence of OS-E and OS-F mRNA in the heads of wild-type animals. The qPCR data did not appear to support my hypothesis that OS-E would be present in head tissues at 60 hours after pupation and OS-F would not. The molecular data failed to confirm the absence of OS-F gene product in
60-hour samples implied by observed fluorescence in transgenic animals. On the contrary, the qPCR results actually indicated that OS-F mRNA was twice as abundant as OS-E mRNA. The OS-E qPCR data did, however, show an increase in the amount of OS-E mRNA between 60 hours and adulthood, supporting the observation that GFP increased over time in transgenic individuals. The data were unable to show whether mRNA stopped increasing when fluorescence peaked because no data from 80 hours or later were available. The qPCR data also showed that the rate of OS-F mRNA production increased almost seven times faster than that of OS-E.

The amount of OS-F mRNA in the adult was within one standard deviation of the concentration of ribosomal RNA, indicating that the two were likely in similar concentrations. This result was unexpected because OS-E and OS-F are confined only to the sensory organs while synthesis of ribosomes occurs in every cell in the sample. The measured concentrations of mRNA might be inaccurate. There is, however, the possibility that OS-F mRNA is at an exceedingly high concentration in the olfactory structures of adult flies, but experimental confirmation of this possibility would be necessary before further speculation.

Some of these results might be explained by defective equipment. The Bio-Rad qPCR kit that was used relied on a clear sticker to seal the tops of all 96 wells in the plate at the same time. The sticker was found to adhere to the plate insufficiently, leaving some wells exposed to the air and others exposed to adjacent wells. Ineffective capping of the PCR wells may have caused mixing of the contents of adjacent wells, which might account for the apparently similar concentrations of OS-E and OS-F in 60-hour samples since a mismeasurement of one amplification cycle can affect data by a factor of two.
Also, many samples evaporated in the early cycles of PCR and did not show measurable results. Consequently, data were only recorded from wells with no observed loss of liquid, which did not preclude mixing between adjacent wells. This problem has since been solved, but the experiment was not repeated because my access to the equipment was limited.

If the qPCR data are assumed to be accurate, the pattern seen here may be the result of a mechanism that regulates translation. If mRNA to protein translation is the regulated step in this pathway, it would not be surprising to find relatively high levels of mRNA in the qPCR and a significantly lower concentration of protein in the form of GFP fluorescence. Such a process would need to work on the primary mRNA transcript before the GFP sequence is separated from OS-E/OS-F by splicing.

It is also possible that the difference in GFP expression times was a result of genetic manipulation and not indicative of the natural OS-E/OS-F development time line. There are reports of OS-F being expressed without OS-E in *D. virilis*, and it is thought that the two genes might be undergoing an adaptive divergence (Hekmat-Scafe, 2000). If the data collected here are accurate, the transgenic flies may have inadvertently shown a disruption in OS-F expression caused by the plasmid insertion. If so, a knockout experiment, where lines of flies are raised with a loss-of-function mutation in one or another of these genes, might be valuable in clarifying the apparent interdependence between the two genes. If *D. melanogaster* with nonfunctional OS-F proteins are viable, the situation would present an interesting complement to *D. virilis* and perhaps elucidate functional consequences of the base-pair differences between the two genes.
Works Cited


degradating enzyme: the sensillar esterase of Antheraea polyphemus. Proceedings

503.