



Characterization of UDP-glycosyltransferase genes in *Drosophila pseudoobscura*.

Bourgeois, Gregory, Laurie S. Steverson, Daniel Ortiz-Barrientos, and Mohamed A.F. Noor. Department of Biological Sciences, Louisiana State University, Baton Rouge, LA 70803, USA; Corresponding author: E-mail: NOOR@DUKE.EDU.

Abstract

UDP-glycosyltransferases (UGTs) are enzymes used in olfaction that may play a role in mating discrimination in *Drosophila pseudoobscura*. Here, we investigate regions containing putative UGTs on the fourth chromosome of *D. pseudoobscura* by testing for expression and analyzing the sequences. We were able to confirm expression in two of three putative UGT regions, and we identified an intron and localized the end of transcription in one putative UGT.

Introduction

UDP-glycosyltransferases (UGT's) are biotransformation enzymes that have been implicated in olfaction and detoxification in vertebrates and invertebrates (Radominska-Pandya *et al.*, 1999; Wang *et al.*, 1999). UGTs transfer UDP-glucose, a sugar molecule, to hydrophobic substrates. This process makes them hydrophilic so that they can be eliminated by excretion. They are primarily expressed within olfactory organs in *Drosophila* (Wang *et al.*, 1999).

Recently, Ortiz-Barrientos *et al.* (2004) mapped the genetic basis of a behavioral discrimination polymorphism in *D. pseudoobscura* and found putative UGT sequences on the fourth chromosome that may contribute to this phenotype. The study suggested that female mating discrimination may involve traits that respond to olfactory signals. One of the regions identified by Ortiz-Barrientos *et al.* (2004), *Coy-4*, is the region of interest in the present study. A sequence within this region bore similarity to five UGTs from *D. melanogaster*. We looked for all long open reading frames within this region and found three distinct putative protein-coding regions bearing amino acid sequence similarity to UGTs; we have dubbed these UGT region 1, UGT region 2, UGT region 3 (see Figure 1). Regions 2 and 3 were oriented in one direction while region 1 is oriented in the opposite direction. This study was designed to characterize each of these regions by first identifying RNA expression then sequencing cDNA to determine the location of transcript ends and possible introns in *D. pseudoobscura*.

Materials and Methods

The fly strains used were *D. pseudoobscura* Mather, CA 17 (collected 1997) and Flagstaff, AZ (collected 1993). We extracted DNA using the protocol of Gloor and Engels (1992) and RNA using the RNeasy Protect Mini and QIAshredder Kit (Qiagen).

The primers for each region were designed using the *D. pseudoobscura* genome sequence (Richards *et al.*, 2005). The primers were used for reverse transcription of RNA, followed by PCR amplification of the resulting cDNA. The sizes of the cDNA were checked via agarose gel electrophoresis and compared with genomic DNA PCRs using the same primers. We selected products of the expected sizes and purified them using Qiaquick Gel Extraction Kit (Qiagen). We

then sequenced the products in both directions using ABI Big Dye version 3.1 Dye Terminator on an ABI 3100 DNA sequencer (Perkin-Elmer).

To identify the 3' end of one of the UGT transcripts, we used 3' RACE using the Ambion RLM-RACE Kit. A subset of these products was cloned using the TOPO TA Cloning Kit (Invitrogen). These cloned products were also sequenced and analyzed.

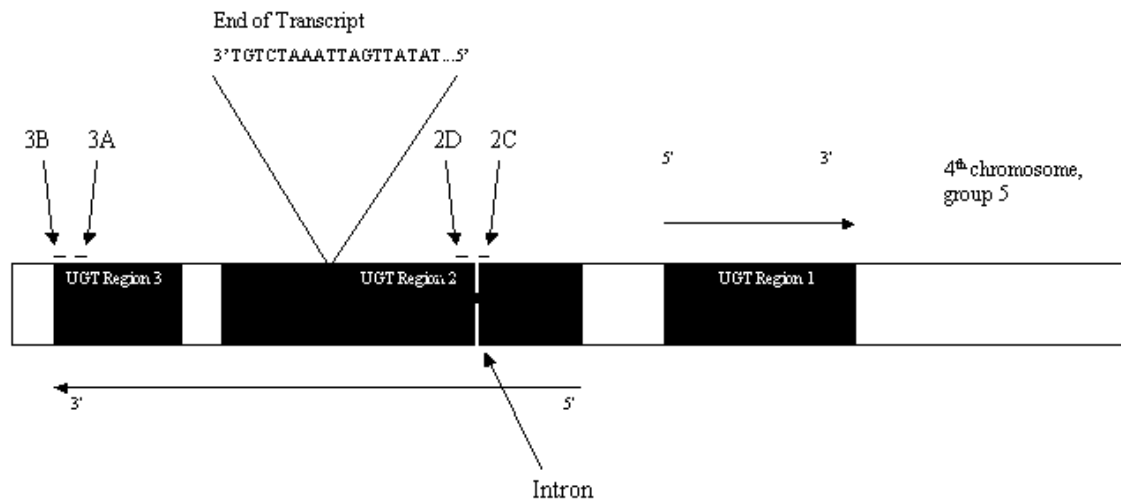


Figure 1: Three putative UGT regions are on the fourth chromosome. The primers used successfully for genomic DNA and cDNA analysis are shown in their approximate annealing areas. The 3' end of the UGT region 2 transcript is shown 38 bp after the putative polyadenylation signal. The putative direction of transcription of each region is also noted. UGT regions 2 and 3 are transcribed in the opposite direction of UGT region 1.

Results

UGT region 1

We successfully amplified and sequenced genomic DNA from within the UGT region 1 sequence. Primer combinations that successfully amplified the correct sequence (as defined by the closest BLAST match to the corresponding region of the *D. pseudoobscura* genome sequence) were 1A + 1B and 1C + 1D (see Table 1). However, despite repeated attempts using the same primers, no RT-PCRs successfully recovered the sequence of this region. The 3' RACE was also unsuccessful.

UGT region 2

We amplified and sequenced both genomic DNA and cDNA from region 2 using one primer pair, 2C + 2D. Using this pair, we were able to determine the presence of an intron in this region. The cDNA sequence bearing the spliced intron was submitted as partial CDS to GenBank (Accession AY880679). There were a couple of other primer pairs used that worked with genomic DNA only (2A + 2B and 2E + 2F). We also successfully performed a 3' RACE on this region to determine the end of the transcript using the primer RACE. The transcript ends with the sequence ...TATATTGATTAAATCTGT. A putative polyadenylation signal (ACUAAA) was noted 38 bp before the 3' end of the transcript (see GenBank accession DQ058143).

UGT region 3

We were able to amplify and sequence both genomic DNA and mRNA from region 3 using one primer pair.

Discussion

Our aim in this study was to confirm and characterize the three putative UGT regions from the *Coy-4* region. We were able to confirm transcription of UGT regions 2 and 3 through RT-PCR and further characterize UGT region 2 using 3' RACE. Despite several attempts using different primer pairs, we were unsuccessful in finding conclusive evidence for the transcription of UGT region 1.

UGT region 2 yielded the most conclusive evidence of formation of a UGT protein. We tried five different primer pairs, confirming genomic DNA with three of these and confirming expression through RT-PCR with one of these pairs and with 3' RACE. This was the only region that yielded successful 3' RACE results. Using sequences from the 3' RACE, we identified the end of the transcript. The putative polyadenylation signal at the end of the UGT region 2 was ACUAAA, a single-base variant of the typical polyadenylation signal AAUAAA, and has been found in the human genome (Beaudoing *et al.*, 2000).

UGT region 3 was less characterized than region 2, yet we accumulated data that confirmed its transcription. Through RT-PCR, we confirmed expression from this region and obtained sequence. We were also able to conclude that this region expresses a separate transcript from region 2 despite their close proximity. We initially hypothesized that UGT regions 2 and 3 could have been a hybrid protein. This was shown to be incorrect based on the 3' RACE sequences from region 2 and the sequence from region 3 which was outside of the end of the region 2 transcript. Further investigation in this region should attempt to characterize this transcript more using expression data from RT-PCR and 3' RACE.

We were able to confirm the sequence of UGT region 1 genomic DNA through PCR. All attempts at finding expression, including 3' RACE, were unsuccessful. More primers could be designed for this region to determine expression.

Some of the difficulty involved in this study involved the design of the primers because UGT sequences are very conserved. This was certainly a problem that was encountered in all regions, especially region 1. Through the use of more methods of looking at expression, like RT-PCR, 3' and 5' RACE, these regions may be able to be characterized further and perhaps eventually lend support of their use in mating discrimination by female *D. pseudoobscura*.

Acknowledgments: This work was funded by a Howard Hughes Medical Institute grant through the Undergraduate Biological Sciences Education Program to LSU (to G.P.B.) as well as National Science Foundation grants 0211007 and 0314552 (to M.A.F.N.). National Institutes of Health grant P20 RR16456 from the Biomedical Research Infrastructure Network Program of the

Table 1. List of primers used in study that were successful in amplifying *D. pseudoobscura* genome or cDNA.

UGT	Primer Abbreviation	Primer Sequence
1	1A	GGATACGGAGCACATGAGCAGAG
1	1B	GATGCTAGGGACACAAGACGCC
1	1C	ACGGTCATCTACTGGTCCGAATATGTT
1	1D	GACTTGTTGAATCACTTCTCTGTTTCAGGTT
2	2A	CGGCAAGCATGAGTGCCA
2	2B	ATTAATGCTGGGCACATAGGCGATT
2	2C	TGGAGTACCGCAACAAAACTCCTAC
2	2D	CATTCCCCAAGAACTCCGCCATAT
2	2E	CCACGGATGGAGCTATACTCCT
2	2F	CGATCAAACCCAATGCTCTAAGGC
2	RACE	ATATGGCGGAGTTCTTGGGGAATG
3	3A	CATTGGTCACCGAGGGCTTTG
3	3B	GTGCGGCACCATGGTGG

National Center for Research Resources provided support for L.S.S. as an undergraduate fellow. The authors thank L. Burk and A. Chang for technical support.

References: Gloor, G.B., and W.R. Engels 1992, *Dros. Inf. Serv.* 71: 148-149; Ortiz-Barrientos, D., B.A. Counterman, and M.A.F. Noor 2004, *PLoS Biol.* 2: 2256-2263; Radominska-Pandya, A., P.J. Czernik, J.M. Little, E. Battaglia, and P.I. Mackenzie 1999, *Drug Metab. Rev.* 4: 817-899; Richards, S., Y. Liu, B.R. Bettencourt, P. Hradecky, S. Letovsky *et al.* (50 authors) 2005, *Genome Res.* 1: 1-18; Wang, Q., G. Hasan, and C.W. Pikielny 1999, *J. Biol. Chem.* 15: 10309-10315.



Study of *Drosophila* association with certain plant species in Islamabad, Pakistan.

Amin ud Din M^{1,3,4}, S.M.N. Khan², A. Bakhsh¹, R. Aleem³, S. Haque³, and A. Salam⁴.

¹Department of Biology, G. C. DG Khan, Pakistan; ²Department of Zoology, AJK Univ., Muzafarabad; ³Department of Biology, QAU, Islamabad, Pakistan; ⁴Inst. of P&A Bio., BZU, Multan, Pakistan; For correspondence, [email: amin1158@hotmail.com](mailto:amin1158@hotmail.com).

The contributions of *Drosophila* as a model system for understanding basic biological mechanisms are even more evident today than in the previous years. It's why a number of workers have been busy exploring the various ecological aspects of *Drosophila* fauna like association with different plant species, because understanding of the pattern in which different species of *Drosophila* are distributed across and within different vegetation types is necessary for accurate interpretation of their local ecology and biodiversity (Van Klinken and Walter, 2001).

According to various studies, the restricted geographical distributions of many native and some cosmopolitan species of *Drosophila* have suggested that they may have specific habitat preferences, such as traps placed among oak trees, which attract five times as many *Drosophila pseudoobscura* as do those in a meadow or shady ravine (Dobzhansky and Epling, 1944) and *D. occidentalis* comes in greater numbers to traps placed near a stream than to those in drier areas (Cooper and Dobzhansky, 1956). Similarly, Montgomery (1975) found 77% of the pictured-winged *Drosophila* species to be specific to a single host family of plants. Similarly *D. repleta* has found to be associated with Cactaceae and *D. subobscura* restricted to fruit bearing plants (Monclus, 1978). Van Klinken and Walter (2001) and O'Grady *et al.* (2003) have also studied the ecological association of various *Drosophila* species and discussed the possible reasons of such associations with plants.

In Pakistan, a similar study was conducted during the favorable season (September to April) to know the association of *Drosophila* species with plants in Islamabad. Collections of *Drosophila* species like *D. immigrans* (D1), *D. hydei* (D2), *D. takahashii* (D3), *D. leontia* (D4), *D. melanogaster* (D5), *D. malerkotliana* (D6), *D. Suzuki* (D7), and *D. nepalensis* (D8) were made by using ripe fermented fruits from the plants *Cassia fistula*, *Callistemon citrinus*, *Dodonaea viscosa*, *Thevetia peruvirana*, *Eucalyptus lanceolatus*, *Bougainvillea spectabilis*, *Sambucus nigra*, *Punica granatum*, *Ficus carica*, *Psidium guava*, *Carissa carandas*, and *Lantana camara*, and information with respect to the plant species from which the flies were collected was recorded that is presented in Table 1.

The results indicated that the traps from two plant species (*Thevetia pervirana* and *Sambucus nigra*) remained without flies. Actually these plants are poisonous and insecticidal. So these are not suitable for *Drosophila* collection. Three plant species (*Bougainvillea spectabilis*, *Ficus carica*,