ODORANT BINDING PROTEIN HOMOLOGUES OF THE MALARIA MOSQUITO Anopheles gambiae; POSSIBLE ORTHOLOGUES OF THE OS-E AND OS-F OBPS OF Drosophila melanogaster

RICHARD G. VOGT

Department of Biological Sciences
University of South Carolina
Columbia South Carolina 29208

(Received August 12, 2002; accepted August 27, 2002)

Abstract—Twenty-nine Anopheles gambiae candidate Odorant Binding Proteins (OBPs) were characterized for similarity to OBPs of Drosophila melanogaster and other insects. Twenty-five of these sequences were identified by BLAST searching the A. gambiae genome database. Several A. gambiae sequences were significantly similar to the D. melanogaster OBPs OS-E/OS-F, LUSH and PBPRP2/PBPRP5. Exon boundary comparisons suggest that two A. gambiae genes are orthologues of OS-E and OS-F, justifying the names AgamOS-E (EAA01090, AF437886) and AgamOS-F (EAA14641, AF437884). If these are orthologues, then the gene duplication establishing the OS-E and OS-F lineages predated the divergence of mosquitoes and flies. The identification of orthologous OBPs and other chemosensory genes between D. melanogaster and A. gambiae should accelerate the transfer of physiological and behavioral information between these two species.

Key Words—Odorant Binding Protein, OBP, evolution, odor detection, pheromone detection, chemosensory proteins, Anopheles gambiae, mosquito, olfaction.

INTRODUCTION

Odorant Binding Proteins (OBPs) are small water-soluble proteins thought to transport odors to receptor proteins in olfactory neuron membranes (Vogt et al., 1999). OBPs comprise a multigene family; thirteen have been identified in Manduca sexta through direct sequencing, and 38 in Drosophila melanogaster through direct sequencing and genome analysis (Robertson et al., 1999; Galindo and Smith, 2001;
Graham and Davies, 2002; Vogt et al., 2002). OBPs are differentially expressed among olfactory sensilla, contributing to the unique phenotypes and odor sensitivities of these sensilla (de Bruyne et al., 2001; Shanbhag et al., 2001). In this study, OBP genes of D. melanogaster and other insect species were used to identify OBPs from the A. gambiae genome, and the likely A. gambiae orthologues of the D. melanogaster OBPs OS-E and OS-F (also known as PBPRP3) are identified.

METHODS

Amino acid sequences of previously reported OBPs were used in a BLAST-P search of the A. gambiae genome through the NCBI web site (8/2002). This database identifies hypothetical gene coding regions and the corresponding amino acid sequences. Sequences with an E-value of 0.05 or lower were accepted. Sequences were aligned using ClustalX, and a neighbor joining tree was constructed using MEGA2. No effort was made to modify sequences, although previously reported versions of these sequences were used if available. Exon boundaries were determined by translating the genomic sequences into multiple reading frames to identify the amino acid sequences bounding exon/intron transitions.

RESULTS AND DISCUSSION

Table 1 lists 29 candidate OBP genes from A. gambiae. Twenty-five unique genes were identified through BLAST search analysis of the A. gambiae genome and are listed with the protein accession number (EAA#) and the corresponding gene scaffold (AAA#). Four pairs of identified sequences were considered to represent the same genes; small variations in nucleic acid and amino acid sequences and very low variation in third base positions of codons suggested that differences between members of each pair was either due to allelic variation or sequencing error. Six of the 25 sequences were previously reported from direct sequence analysis (Biessmann et al., 2002). Three additional sequences were also previously reported as A. gambiae OBP. The AgamD7 gene, encoding a salivery protein, was also identified as a candidate OBP homologue.

Figure 1A shows a sequence similarity tree assembled from an alignment of the 29 candidate A. gambiae OBP sequences, 38 D. melanogaster OBP sequences (Galindo and Smith, 2001; Graham and Davies, 2002; Vogt et al., 2002) and OBP related sequences from diverse other species. The long-branch lengths in this tree indicate the considerable sequence divergence between most of these sequences. Most sequences are in branches specific to a given insect group, though several similarity groups do include multiple insect groups. Three similarity groups include both A. gambiae and D. melanogaster sequences: OS-E/OS-F, LUSH, and PBPRP2/PBPRP5. The OS-E/OS-F group includes 1 sequence from Culex