PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N

A2

(11) International Publication Number: WO 00/50566

(43) International Publication Date: 31 August 2000 (31.08.00)

(21) International Application Number: PCT/US00/04995

(22) International Filing Date: 25 February 2000 (25.02.00)

(30) Priority Data:
09/257,706 25 February 1999 (25.02.99) US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application
US 09/257,706 (CIP)
Filed on 25 February 1999 (25.02.99)

(71) Applicant (*for all designated States except US*): THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK [US/US]; West 116th Street and Broadway, New York, NY 10027 (US).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): VOSSHALL, Leslie, B. [US/US]; 601 West 113th Street, Apt. 3H, New York, NY 10025 (US). AMREIN, Hubert, O. [CH/US]; 604 Carved Oak Drive, Durham, NC 27707 (US). AXEL, Richard [US/US]; 435 Riverside Drive, Apt. 62, New York, NY 10027 (US).

(74) Agent: WHITE, John, P.; Cooper & Dunham LLP, 1185 Avenue of the Americas, New York, NY 10036 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: GENES ENCODING INSECT ODORANT RECEPTORS AND USES THEREOF

(57) Abstract

This invention provides an isolated nucleic acid molecule encoding an insect odorant receptor. This invention provides a nucleic acid molecule of at least 12 nucleotides capable of specifically hybridizing with the nucleic acid molecule encoding an insect odorant receptor. This invention also provides a purified, insect odorant receptor. This invention provides an antibody capable of specifically binding to an insect odorant receptor. This invention provides a method for identifying cDNA inserts encoding an insect odorant receptors. This invention provides a method of identifying a compound capable of specifically bind to an insect odorant receptor. This invention also provides a method of identifying a compound capable of activating the activity of an insect odorant receptor.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

GENES ENCODING INSECT ODORANT RECEPTORS AND USES THEREOF

This application claims priority and is a continuation-in-part application of U.S. Serial No. 09/257,706, filed February 25, 1999, the contents of which is hereby incorporated by reference.

The invention disclosed herein was made with Government support under NIH:NIMH, 5P50, MH50733-05 and the NINDS, NS29832-07 from the Department of Health and Human Services. Accordingly, the U.S. Government has certain rights in this invention.

Throughout this application, various publications are referred to by arabic numeral within parentheses. Full citations for these publications are presented immediately before the claims. Disclosures of these publications in their entireties are hereby incorporated by reference into this application in order to more fully describe the state of the art to which this invention pertains.

BACKGROUND OF THE INVENTION

All animals possess a "nose," an olfactory sense organ that allows for the recognition and discrimination of chemosensory information in the environment. Humans, for example, are thought to recognize over 10,000 discrete odors with exquisite discriminatory power such that subtle differences in chemical structure can often lead to profound differences in perceived odor quality. What mechanisms have evolved to allow the recognition and discrimination of complex olfactory information and how is olfactory perception ultimately translated into appropriate behavioral responses? The recognition of odors is accomplished by odorant receptors that reside on olfactory cilia, a specialization of the dendrite of the olfactory sensory neuron. The odorant receptor genes encode novel serpentine receptors that traverse the membrane seven times. In several vertebrate species, and in the invertebrate Caenorhabditis elegans, as many as 1000 genes encode odorant receptors, suggesting that 1-5% of the coding potential of the genome in these organisms is devoted to the recognition of olfactory sensory stimuli (Buck and Axel, 1991; Levy et al., 1991; Parmentier et al., 1992; Ben-Arie et al., 1994; Troemel et al., 1995; Sengupta et al., 1996; Robertson, 1998). Thus, unlike color vision in which three photoreceptors can absorb light across the entire visible spectrum, these data suggest that a small number of odorant receptors are insufficient to recognize the full spectrum of distinct molecular structures perceived by the olfactory system. Rather, the olfactory sensory system employs an extremely large number of receptors, each capable of recognizing a small number of odorous ligands.

The discrimination of olfactory information requires that the brain discern which of the numerous receptors have been activated by an odorant. In mammals, individual olfactory sensory neurons express only one of a thousand receptor genes such that the neurons are functionally distinct (Ngai et al., 1993; Ressler et al., 1993; Vassar et al., 1993; Chess et al., 1994; Dulac and Axel, unpublished). The axons from olfactory neurons expressing a specific receptor converge upon two spatially invariant glomeruli among the 1800 glomeruli within the olfactory bulb (Ressler et al., 1994; Vassar et al., 1994; Mombaerts et al., 1996; Wang et al., 1998). The bulb therefore provides a spatial map that identifies which of the numerous receptors has been activated within the sensory epithelium. The quality of an olfactory stimulus would therefore be encoded by specific combinations of glomeruli activated by a given odorant.

The logic of olfactory discrimination is quite different in the nematode, C. elegans. Despite the large size of the odorant receptor gene family, volatile odorants are recognized by only three pairs of chemosensory cells each likely to express a large number of receptor genes (Bargmann and Horvitz, 1991; Colbert and Bargmann, 1995; Troemel et al., 1995). Activation of any one of the multiple receptors in one cell will lead to chemoattraction, whereas activation of receptors in a second cell will result in chemorepulsion (Troemel et al., 1997). The specific neural circuit activated by a given sensory neuron is therefore the determinant of the behavioral response. Thus, this invertebrate olfactory sensory system retains the ability to recognize a vast array of odorants but has only limited discriminatory power.

Vertebrates create an internal representation of the external olfactory world that must translate stimulus features into neural information. Despite the elucidation of a precise spatial map, it has been difficult in vertebrates to discern how this information is decoded to relate the recognition of odors to specific behavioral responses. Genetic analysis of olfactory-driven behavior in invertebrates may ultimately afford a system to understand the mechanistic link between odor recognition and behavior. Insects provide an attractive model system for studying the peripheral and central events in olfaction because they exhibit sophisticated olfactory-driven behaviors under control of an olfactory sensory system that is significantly simpler anatomically than that of vertebrates (Siddiqi, 1987; Carlson, 1996). Olfactory-based associative learning, for example, is robust in insects and results in discernible modifications in the neural representation of odors in the brain (Faber et al., 1998). It may therefore be possible to associate modifications in defined olfactory connections with *in vivo* paradigms for learning and memory.

Olfactory recognition in the fruit fly Drosophila is accomplished by sensory hairs distributed over the surface of the third antennal segment and the maxillary palp. Olfactory neurons within sensory hairs send projections to one of 43 glomeruli within the antennal lobe of the brain (Stocker, 1994; Laissue et al, 1999). The glomeruli are innervated by dendrites of the projection neurons, the insect equivalent of the mitral cells in the vertebrate olfactory bulb, whose cell bodies surround the glomeruli. These antennal lobe neurons in turn project to the mushroom body and lateral horn of the protocerebrum (reviewed in Stocker, 1994). 2-deoxyglucose mapping in the fruit fly (Rodrigues, 1988) and calcium imaging in the honeybee (Joerges et al., 1997; Faber et al., 1998) demonstrate that different odorants elicit defined patterns of glomerular activity, suggesting that in insects as in vertebrates, a topographic map of odor quality is represented in the antennal lobe. However, in the absence of the genes encoding the receptor molecules, it has not been possible to define a physical basis for this spatial map.

In this study, we identify a large family of genes that are likely to encode theodorant receptors of Drosophila melanogaster. Difference cloning, along with analysis of Drosophila genomic sequences, has led to the identification of a novel family of putative seven transmembrane domain receptors likely to be encoded by 100 to 200 genes within the Drosophila genome. Each receptor is expressed in a small subset of sensory cells (0.5-1.5%) that is spatially defined within the antenna and maxillary palp. Moreover, different neurons express distinct complements of receptor genes such that individual neurons are functionally distinct. Identification of a large family of putative odorant receptors in insects indicates that, as in other species, the diversity and specificity of odor recognition is accommodated by a large family of receptor genes. The identification of the family of putative odorant receptor genes may afford insight into the logic of olfactory perception in Drosophila.

Insects provide an attractive system for the study of olfactory sensory perception. We have identified a novel family of seven transmembrane domain proteins, encoded by 100 to 200 genes, that is likely to represent the family of Drosophila odorant receptors. Members of this gene family are expressed in topographically defined subpopulations of olfactory sensory neurons in either the antenna or the maxillary palp. Sensory neurons express different complements of receptor genes, such that individual neurons are functionally distinct. The isolation of candidate odorant receptor genes along with a genetic analysis of olfactory-driven behavior in insects may ultimately afford a system to understand the mechanistic link between odor recognition and behavior.

SUMMARY OF THE INVENTION

This invention provides an isolated nucleic acid molecule encoding an insect odorant receptor. In an embodiment, the isolated nucleic acid molecule comprise: (a) one of the nucleic acid sequences as set forth in Figure 8, (b) a sequence being degenerated to a sequence of (a) as a result of the genetic code; or (c) a sequence encoding one of the amino acid sequences as set forth in Figure 8.

This invention provides a nucleic acid molecule of at least 12 nucleotides capable of specifically hybridizing with the sequence of the above-described nucleic acid molecule. This invention provides a vector which comprises the above-described isolated nucleic acid molecule. In another embodiment, the vector is a plasmid.

This invention also provides a host vector system for the production of a polypeptide having the biological activity of an insect odorant receptor which comprises the above described vector and a suitable host.

This invention provides a method of producing a polypeptide having the biological activity of an insect odorant receptor which comprising growing the above described host vector system under conditions permitting production of the polypeptide and recovering the polypeptide so produced.

This invention also provides a purified, insect odorant receptor. This invention further provides a polypeptide encoded by the above-described isolated nucleic acid molecule.

This invention provides an antibody capable of specifically binding to an insect odorant receptor. This invention also provides an antibody capable of competitively inhibiting the binding of the antibody capable of specifically binding to an insect odorant receptor.

This invention provides a method for identifying cDNA inserts encoding an insect odorant receptors comprising:(a)generating a cDNA library which contains clones carrying cDNA inserts from antennal or maxillary palp sensory neurons; (b) hybridizing nucleic acid molecules of the clones from the cDNA libraries generated in step (a) with probes prepared from the antenna or maxillary palp neurons and probes from heads lacking antenna or maxillary palp neurons or from virgin female body tissue; (c) selecting clones which hybridized with probes from the antenna or maxillary palp neurons but not from head lacking antenna or maxillary palp neurons or virgin female body tissue; and (d) isolating clones which carry the hybridized inserts, thereby identifying the inserts encoding odorant receptors.

This invention also provides cDNA inserts identified by the above method.

This invention further provides a method for identifying DNA inserts encoding an insect odorant receptors comprising:(a) generating DNA libraries which contain clones carrying inserts from a sample which contains at least one antennal or maxillary palp neuron; (b) contacting clones from the cDNA libraries generated in step (a) with nucleic acid molecule capable of specifically hybridizing with the sequence which encodes an insect odorant receptor in appropriate conditions permitting the hybridization of the nucleic acid molecules of the clones and the nucleic acid molecule; (c)selecting clones which hybridized with the nucleic acid molecule; and (d) isolating the clones which carry the hybridized inserts, thereby identifying the inserts encoding the odorant receptors.

This invention also provides a method to identify DNA inserts encoding an insect odorant receptors comprising:
(a) generating DNA libraries which contain clones with inserts from a sample which contains at least one antenna or maxillary palp sensory neuron; (b) contacting the clones from

the DNA libraries generated in step (a) with appropriate polymerase chain reaction primers capable of specifically binding to nucleic acid molecules encoding odorant receptors in appropriate conditions permitting the amplification of the hybridized inserts by polymerase chain reaction; (c) selecting the amplified inserts; and (d) isolating the amplified inserts, thereby identifying the inserts encoding the odorant receptors.

This invention also provides a method to isolate DNA molecules encoding insect odorant receptors comprising:(a) contacting a biological sample known to contain nucleic acids with appropriate polymerase chain reaction primers capable of specifically binding to nucleic acid molecules encoding insect odorant receptors in appropriate conditions permitting the amplification of the hybridized molecules by polymerase chain reaction; (b) isolating the amplified molecules, thereby identifying the DNA molecules encoding the insect odorant receptors.

This invention also provides a method of transforming cells which comprises transfecting a host cell with a suitable vector described above. This invention also provides transformed cells produced by the above method.

This invention provides a method of identifying a compound capable of specifically bind to an insect odorant receptor which comprises contacting a transfected cells or membrane fractions of the above described transfected cells with an appropriate amount of the compound under conditions permitting binding of the compound to such receptor, detecting the presence of any such compound specifically bound to the receptor, and thereby determining whether the compound specifically binds to the receptor.

This invention provides a method of identifying a compound capable of specifically binding to an insect odorant receptor which comprises contacting an appropriate amount of the purified insect odorant receptor with an appropriate amount of the compound under conditions permitting binding of the compound to such purified receptor, detecting the presence of any such compound specifically bound to the receptor, and thereby determining whether the compound specifically binds to the receptor.

This invention also provides a method of identifying a compound capable of activating the activity of an insect odorant receptor which comprises contacting the transfected cells or membrane fractions of the above-described transfected cells with the compound under conditions permitting the activation of a functional odorant receptor response, the activation of the receptor indicating that the compound is capable of activating the activity of a odorant receptor.

This invention also provides a method of identifying a compound capable of activating the activity of an odorant receptor which comprises contacting a purified insect odorant receptor with the compound under conditions permitting the activation of a functional odorant receptor response, the activation of the receptor indicating that the compound is capable of activating the activity of a odorant receptor. In an embodiment, the purified receptor is embedded in a lipid bilayer.

This invention also provides a method of identifying a compound capable of inhibiting the activity of a odorant receptor which comprises contacting the transfected cells or membrane fractions of the above-described transfected cells with an appropriate amount of the compound under conditions permitting the inhibition of a functional odorant receptor response, the inhibition of the receptor response indicating that the compound is capable of inhibiting the activity of a odorant receptor.

This invention provides a method of identifying a compound capable of inhibiting the activity of a odorant receptor which comprises contacting an appropriate amount of the purified insect odorant receptor with an appropriate amount of the compound under conditions permitting the inhibition of a functional odorant receptor response, the inhibition of the receptor response indicating that the compound is capable of activating the activity of a odorant receptor. In an embodiment, the purified receptor is embedded in a lipid bilayer.

This invention also provides the compound identified by the above-described methods.

This invention provides a method of controlling pest populations which comprises identifying odorant ligands by the above-described method which are alarm odorant ligands and spraying the desired area with the identified odorant ligands.

Finally, this invention provides a method of controlling a pest population which comprises identifying odorant ligands by the above-described method which interfere with the interaction between the odorant ligands and the odorant receptors which are associated with fertility.

BRIEF DESCRIPTION OF FIGURES

FIGURE 1 **Identification of Rare Antennal- and Maxillary Palp-Specific Genes**

Candidate antennal/maxillary palp-specific phage were subjected to *in vivo* excision, digestion of resulting pBLUESCRIPT plasmid DNAs with BamHI/Asp718, and electrophoresis on 1.5% agarose gels. Southern blots were hybridized with ^{32}P-labeled cDNA probes generated from antennal/maxillary palp mRNA (Panel A), head minus antennal/maxillary palp mRNA (Panel B), or virgin female body mRNA (Panel C). The ethidium bromide stained gel is shown in Panel D. Of the thirteen clones displayed in this figure, four appear to be antennal/maxillary palp specific (lanes 5, 7, 9, and 11). However, only two are selectively expressed in subsets of cells in chemosensory organs of the adult fly. DOR104, a putative maxillary palp odorant receptor, is in Lane 9. The clone in Lane 11 (RN106) is homologous to lipoprotein and triglyceride lipases and is expressed in a restricted domain in the antenna (data not shown).

FIGURE 2 **Expression of DOR104 in a Subset of Maxillary Palp Neurons**

(A) A frontal section of an adult maxillary palp was hybridized with a digoxigenin-labeled antisense RNA probe and visualized with anti-digoxigenin conjugated to alkaline phosphatase. Seven cells expressing DOR104 are visible in this 15 μm section, which represents about one third of the diameter of the maxillary palp. Serial sections of multiple maxillary palps were scored for DOR104 expression and on average 20 cells per maxillary palp are positive for this receptor.

(B) Transgenic flies carrying a DOR104-lacZ reporter transgene were stained with X-GAL in a whole mount preparation. Maxillary palps were dissected from the head and viewed in a flattened cover slipped preparation under Nomarski optics, which allows the visualization of all 20 cells expressing DOR104-lacZ.

(C) Dendrites and axons of neurons expressing DOR104-lacZ are visible in this horizontal section of a maxillary palp. LacZ expression was visualized with a polyclonal anti-ß-galactosidase primary antibody and a CY3-conjugated secondary antibody. Sections were viewed under epifluorescence and photographed on black and white film.

__FIGURE 3__ **Predicted Amino Acid Sequences of Drosophila Odorant Receptor Genes**

Deduced amino acid sequences of 12 DOR genes are aligned using ClustalW (MacVector, Oxford Molecular). Predicted positions of transmembrane regions (I-VII) are indicated by bars above the alignment. Amino acids identities are marked with dark shading and similarities are indicated with light shading. Protein sequences of DOR87, 53, 67, 104, and 64 were derived from cDNA clones. All others were derived from GENSCAN predictions of intron-exon arrangements in genomic DNA, as indicated by the letter "g" after the gene name. We obtained a partial cDNA clone for DOR62 and found it to be 100% identical to the GENSCAN protein in the region of amino acids 245-381. A 40 amino acid extension for DOR 19 was predicted by GENSCAN analysis. This has been replaced with an asterisk in the alignment, and isolation of cDNA clones for this receptor will resolve whether this extension is physically present in the protein.

FIGURE 4 Receptor Gene Expression in Spatially Restricted Regions of the Antenna

Digoxigenin-labeled antisense RNA probes against 8 DOR genes each hybridize to a small number of cells distributed in distinct regions in the antenna. The total number of cells per antenna expressing a given receptor was obtained by counting positive cells in serial sections of multiple antennae. There are approximately 20 positive cells per antenna for DOR67 (A), 53 (B), and 24 (data not shown); 15 positive cells for DOR62 (C) and 87 (D); and 10 positive cells for DOR64 (E). The actual number of cells staining in these sections is a subset of this total number. With the exception of DOR53 and DOR67, which strongly cross-hybridize, the receptor genes likely identify different olfactory neurons, such that the number of cells staining with a mixed probe (F) is equal to the sum of those staining with the individual probes (A-E). The mixture of DOR53, 67, 62, 87 and 64 labels a total of about 60 cells per antenna. A total of 34 cells stain with the mixed probe in this 15 μm section. Expression of the linked genes DOR71, DOR72, and DOR73 is shown in panels (G), (H), and (I), respectively. DOR71 is expressed in approximately 10 cells in the maxillary palp. Five positive cells are seen in the horizontal section in panel (G). We also examined the expression of the other members of this linkage group and found DOR72 in approximately 15 cells (of which 3 label in this section) (H) and DOR73 in 1 to 2 cells per antenna (I).

FIGURE 5 Odorant Receptors are Restricted to Distinct Populations of Olfactory Neurons

(A-C) Flies of the C155 *elav*-GAL4; UAS-lacZ genotype express cytoplasmic lacZ in all neuronal cells. Panels (A-C) show confocal images of a horizontal maxillary palp section from such a fly incubated with an antisense RNA probe against DOR104 (red) and anti-ß-galactosidase antibody (green). DOR104 recognizes five cells in this maxillary palp section (A), all of which also express *elav*-lacZ (B), as demonstrated by the yellow cells in the merged image in panel ©).

(D, E) DOR64 and DOR87 are expressed in non-overlapping neurons at the tip of the antenna. Antisense RNA probes for DOR64 (digoxigenin-RNA; red) and DOR87 (FITC-RNA; green) were annealed to the same antennal sections and viewed by confocal microscopy. Panel (D) is a digital superimposition of confocal images taken at 0.5 μm intervals through a 10 μm section of the antenna. Cells at different focal planes express both receptors, but no double labeled cells are found.

(F, G) Two color RNA *in situ* hybridization with odorant receptors and odorant binding proteins demonstrates that these proteins are expressed in different populations of cells. DOR53 (FITC-RNA; green) labels a few cells internal to the cuticle at the proximal-medial edge, while PBPRP2 (digoxigenin-RNA; red) labels a large number of cells apposed to the cuticle throughout the antenna (F). The more restricted odorant binding protein OS-F (digoxigenin-RNA; red) also stains cells distinct from those expressing DOR67 (FITC-RNA; green)(G).

FIGURE 6 **Receptor Expression is Conserved Between Individuals**

Frontal sections of antennae from six different individuals were hybridized with digoxigenin-labeled antisense RNA probes against DOR53 (A-C) or DOR87 (D-F). DOR53 labels approximately 20 cells on the proximal-medial edge of the antenna, of which approximately 5 are shown labeling in these sections. DOR87 is expressed in about the same number of cells at the distal tip. Both the position and number of staining cells is conserved between different individuals and is not sexually dimorphic.

<u>**FIGURE 7**</u> **Drosophila Odorant Receptors are Highly Divergent**

Oregon R genomic DNA isolated from whole flies was digested with BamHI (B), EcoRI (E), or HindIII (H), electrophoresed on 0.8% agarose gels, and blotted to nitrocellulose membranes. Blots were annealed with ^{32}P-labeled probes derived from DOR53 cDNA (A), DOR67 cDNA (B), or DNA fragments generated by RT-PCR from antennal mRNA for DOR 24 (C), DOR62 (D), and DOR72 (E). Strong crosshybridization of DOR53 and DOR67 is seen at both high and low stringency (A, B), while DOR24, 62, and 72 reveal only a single hybridizing band in each lane at both low stringency (C-E) and high stringency (data not shown).

<u>**FIGURE 8**</u> **DOR 62, 104, 87, 53, 67, 64, 71g, 72g, 73g, 46, 19g, and 24g**

Both nucleic acid sequence of each DOR and its encoded amino acid sequence are described.

<u>**FIGURE 9**</u> Analysis of axonal projections of olfactory receptor neurons expressing a given *Drosophila* odorant receptor. Result: all neurons expressing a given receptor send their axons to a single glomerulus, or discrete synaptic structure, in the olfactory processing center of the fly brain. This result is identical to that obtained with

mouse odorant receptors: each glomerulus is dedicated to receiving axonal input from neurons expressing a given odorant receptor. Therefore, this result strengthens the argument that these genes indeed function as odorant receptors in *Drosophila*.

<u>FIGURE 10</u> **ClustalW alignments of two subfamilies of the *Drosophila* odorant receptors, the DOR53 (A-1 and A-2) and DOR64 (B) families.** This figure highlights sequence similarities between DOR genes, that are diagnostic hallmarks of the proteins. Residues that are identical in different DOR genes are highlighted in black, while residues that are similar are highlighted in gray.

DETAILED DESCRIPTION OF THE INVENTION

In order to facilitate an understanding of the Experimental Procedures section which follow, certain frequently occurring methods and/or terms are described in Sambrook, et al. (1989).

Throughout this application, the following standard abbreviations are used throughout the specification to indicate specific nucleotides:

C=cytosine	A=adenosine
T=thymidine	G=guanosine

This invention provides an isolated nucleic acid molecule encoding an insect odorant receptor. The nucleic acid includes but is not limited to DNA, cDNA, genomic DNA, synthetic DNA or RNA. In an embodiment, the nucleic acid molecule encodes a Drosophila odorant receptor.

In a further embodiment, the isolated nucleic acid molecule comprise: (a) one of the nucleic acid sequences as set forth in Figure 8, (b) a sequence being degenerated to a sequence of (a) as a result of the genetic code; or (c) a sequence encoding one of the amino acid sequences as set forth in Figure 8.

The nucleic acid molecules encoding a insect receptor includes molecules coding for polypeptide analogs, fragments or derivatives of antigenic polypeptides which differ from naturally-occurring forms in terms of the identity or location of one or more amino acid residues (deletion analogs containing less than all of the residues specified for the protein, substitution analogs wherein one or more residues specified are replaced by other residues and addition analogs where in one or more amino acid residues is added to a terminal or medial portion of the polypeptides) and which share some or all properties of naturally-occurring forms.

These molecules include but not limited to: the incorporation of codons "preferred" for expression by selected non-mammalian hosts; the provision of sites for cleavage by restriction endonuclease enzymes; and the provision of additional initial, terminal or intermediate sequences that facilitate construction of readily expressed vectors. Accordingly, these changes may result in a modified insect odorant receptor. It is the intent of this invention to include nucleic acid molecules which encodes modified insect odorant receptor. Also, to facilitate the expression of receptor in different host cells, it may be necessary to modify the molecule such that the expressed receptors may reach the surface of the host cells. The modified insect odorant receptor should have biological activities similar to the unmodified insect odorant receptor. The molecules may also be modified to increase the biological activity of the expressed receptor.

This invention provides a nucleic acid molecule of at least 12 nucleotides capable of specifically hybridizing with the sequence of the above-described nucleic acid molecule. In an embodiment, the nucleic acid molecule hybridizes with a unique sequence within the sequence of the above-described nucleic acid molecule. This nucleic acid molecule may be DNA, cDNA, genomic DNA, synthetic DNA or RNA.

This invention provides a vector which comprises the above-described isolated nucleic acid molecule. In another embodiment, the vector is a plasmid.

In an embodiment, the above described isolated nucleic acid molecule is operatively linked to a regulatory element.

Regulatory elements required for expression include promoter sequences to bind RNA polymerase and transcription initiation sequences for ribosome binding. For example, a bacterial expression vector includes a promoter such as the lac promoter and for transcription initiation the Shine-Dalgarno sequence and the start codon AUG. Similarly, a eukaryotic expression vector includes a heterologous or homologous promoter for RNA polymerase II, a downstream polyadenylation signal, the start codon AUG, and a termination codon for detachment of the ribosome. Such vectors may be obtained commercially or assembled from the sequences described by methods well-known in the art, for example the methods described above for constructing vectors in general.

This invention also provides a host vector system for the production of a polypeptide having the biological activity of an insect odorant receptor which comprises the above described vector and a suitable host.

This invention also provides a host vector system, wherein the suitable host is a bacterial cell, yeast cell, insect cell, or animal cell. The host cell of the above expression system may be selected from the group consisting of the cells where the protein of interest is normally expressed, or foreign cells such as bacterial cells (such as *E. coli*), yeast cells, fungal cells, insect cells, nematode cells, plant or animal cells, where the protein of interest is not normally expressed. Suitable animal cells include, but are not limited to Vero cells, HeLa cells, Cos cells, CV1 cells and various primary mammalian cells.

This invention provides a method of producing a polypeptide having the biological activity of an insect odorant receptor which comprising growing the above described host vector system under conditions permitting production of the polypeptide and recovering the polypeptide so produced.

This invention also provides a purified, insect odorant receptor. This invention further provides a polypeptide encoded by the above-described isolated nucleic acid molecule.

This invention provides an antibody capable of specifically binding to an insect odorant receptor. This invention also provides an antibody capable of competitively inhibiting the binding of the antibody capable of specifically binding to an insect odorant receptor. In an embodiment, the antibody is monoclonal. In another embodiment, the antibody is polyclonal.

Monoclonal antibody directed to an insect odorant receptor may comprise, for example, a monoclonal antibody directed to an epitope of an insect odorant receptor present on the surface of a cell. Amino acid sequences may be analyzed by methods well known to those skilled in the art to determine whether they produce hydrophobic or hydrophilic regions in the proteins which they build. In the case of cell membrane proteins, hydrophobic regions are well known to form the part of the protein that is inserted into the lipid bilayer which forms the cell membrane, while hydrophilic regions are located on the cell surface, in an aqueous environment.

Antibodies directed to an insect odorant receptor may be serum-derived or monoclonal and are prepared using methods well known in the art. For example, monoclonal antibodies are prepared using hybridoma technology by fusing antibody producing B cells from immunized animals with myeloma cells and selecting the resulting hybridoma cell line producing the desired antibody. Cells such as NIH3T3 cells or 293 cells which express the receptor may be used as immunogens to raise such an antibody. Alternatively, synthetic peptides may be prepared using commercially available machines.

As a still further alternative, DNA, such as a cDNA or a fragment thereof, encoding the receptor or a portion of the receptor may be cloned and expressed. The expressed polypeptide recovered and used as an immunogen.

The resulting antibodies are useful to detect the presence of insect odorant receptors or to inhibit the function of the receptor in living animals, in humans, or in biological tissues or fluids isolated from animals or humans.

This antibodies may also be useful for identifying or isolating other insect odorant receptors. For example, antibodies against the Drosophila odorant receptor may be used to screen an cockroach expression library for a cockroach odorant receptor. Such antibodies may be monoclonal or monospecific polyclonal antibody against a selected insect odorant receptor. Different insect expression libraries are readily available and may be made using technologies well-known in the art.

One means of isolating a nucleic acid molecule which encodes an insect odorant receptor is to probe a libraries with a natural or artificially designed probes, using methods well known in the art. The probes may be DNA or RNA. The library may be cDNA or genomic DNA.

This invention provides a method for identifying cDNA inserts encoding an insect odorant receptors comprising: (a)generating a cDNA library which contains clones carrying cDNA inserts from antennal or maxillary palp sensory neurons; (b) hybridizing nucleic acid molecules of the clones from the cDNA libraries generated in step (a) with probes prepared from the antenna or maxillary palp neurons and probes from heads lacking antenna or maxillary palp neurons or from virgin female body tissue; (c) selecting clones which hybridized with probes from the antenna or maxillary palp neurons but not from head lacking antenna or maxillary palp neurons or virgin female body tissue; and (d) isolating clones which carry the hybridized inserts, thereby identifying the inserts encoding odorant receptors.

In an embodiment of the above method, after step (c), it further comprises: (a) amplifying the inserts from the selected clones by polymerase chain reaction; (b) hybridizing the amplified inserts with probes from the antennal or

maxillary palp neurons; and (c) isolating the clones which carry the hybridized inserts, thereby identifying the inserts encoding the odorant receptors.
In an embodiment, the probes are cDNA probes.

The appropriate polymerase chain reaction primers may be chosen from the conserved regions of the known insect odorant receptor sequences. Alternatively, the primers may be chosen from the regions which are the active sites for the binding of ligands.

This invention also provides cDNA inserts identified by the above method.

This invention further provides amethod for identifying DNA inserts encoding an insect odorant receptors comprising: (a) generating DNA libraries which contain clones carrying inserts from a sample which contains at least one antennal or maxillary palp neuron; (b) contacting clones from the cDNA libraries generated in step (a) with nucleic acid molecule capable of specifically hybridizing with the sequence which encodes an insect odorant receptor in appropriate conditions permitting the hybridization of the nucleic acid molecules of the clones and the nucleic acid molecule; (c) selecting clones which hybridized with the nucleic acid molecule; and (d) isolating the clones which carry the hybridized inserts, thereby identifying the inserts encoding the odorant receptors.

This invention also provides a method to identify DNA inserts encoding an insect odorant receptors comprising:
(a) generating DNA libraries which contain clones with inserts from a sample which contains at least one antenna or maxillary palp sensory neuron; (b) contacting the clones from the DNA libraries generated in step (a) with appropriate polymerase chain reaction primers capable of specifically binding to nucleic acid molecules encoding odorant receptors in appropriate conditions permitting the amplification of the

hybridized inserts by polymerase chain reaction; (c) selecting the amplified inserts; and (d) isolating the amplified inserts, thereby identifying the inserts encoding the odorant receptors.

This invention also provides a method to isolate DNA molecules encoding insect odorant receptors comprising:(a) contacting a biological sample known to contain nucleic acids with appropriate polymerase chain reaction primers capable of specifically binding to nucleic acid molecules encoding insect odorant receptors in appropriate conditions permitting the amplification of the hybridized molecules by polymerase chain reaction; (b) isolating the amplified molecules, thereby identifying the DNA molecules encoding the insect odorant receptors.

This invention also provides a method of transforming cells which comprises transfecting a host cell with a suitable vector described above.

This invention also provides transformed cells produced by the above method. In an embodiment, the host cells are not usually expressing odorant receptors. In another embodiment, the host cells are expressing odorant receptors.

This invention provides a method of identifying a compound capable of specifically binding to an insect odorant receptor which comprises contacting a transfected cells or membrane fractions of the above described transfected cells with an appropriate amount of the compound under conditions permitting binding of the compound to such receptor, detecting the presence of any such compound specifically bound to the receptor, and thereby determining whether the compound specifically binds to the receptor.

This invention provides a method of identifying a compound capable of specifically bind to an insect odorant receptor which comprises contacting an appropriate amount of the purified insect odorant receptor with an appropriate amount of the compound under conditions permitting binding of the compound to such purified receptor, detecting the presence of any such compound specifically bound to the receptor, and thereby determining whether the compound specifically binds to the receptor. In an embodiment, the purified receptor is embedded in a lipid bilayer.The purified receptor may be embedded in the liposomes with proper orientation to carry out normal functions. Liposome technology is well-known in the art.

This invention also provides a method of identifying a compound capable of activating the activity of an insect odorant receptor which comprises contacting the transfected cells or membrane fractions of the above-described transfected cells with the compound under conditions permitting the activation of a functional odorant receptor response, the activation of the receptor indicating that the compound is capable of activating the activity of a odorant receptor.

This invention also provides a method of identifying a compound capable of activating the activity of an odorant receptor which comprises contacting a purified insect odorant receptor with the compound under conditions permitting the activation of a functional odorant receptor response, the activation of the receptor indicating that the compound is capable of activating the activity of a odorant receptor. In an embodiment, the purified receptor is embedded in a lipid bilayer.

This invention also provides a method of identifying a compound capable of inhibiting the activity of a odorant receptor which comprises contacting the transfected cells or membrane fractions of the above-described transfected cells with an appropriate amount of the compound under conditions permitting the inhibition of a functional odorant receptor response, the inhibition of the receptor response indicating that the compound is capable of inhibiting the activity of a odorant receptor.

This invention provides a method of identifying a compound capable of inhibiting the activity of a odorant receptor which comprises contacting an appropriate amount of the purified insect odorant receptor with an appropriated amount of the compound under conditions permitting the inhibition of a functional odorant receptor response, the inhibition of the receptor response indicating that the compound is capable of activating the activity of a odorant receptor. In an embodiment, the purified receptor is embedded in a lipid bilayer.

In a separate embodiment of the above method, the compound is not previously known. This invention also provides the compound identified by the above-described methods.

This invention provides a method of controlling pest populations which comprises identifying odorant ligands by the above-described method which are alarm odorant ligands and spraying the desired area with the identified odorant ligands.

Finally, this invention provides a method of controlling a pest population which comprises identifying odorant ligands by the above-described method which interfere with the interaction between the odorant ligands and the odorant receptors which are associated with fertility.

This invention will be better understood from the Experimental Procedures which follow. However, one skilled in the art will readily appreciate that the specific methods and results discussed are merely illustrative of the invention as described more fully in the claims which follow thereafter.

EXPERIMENTAL PROCEDURES

Experimental Animals

Oregon R flies (Drosophila melanogaster) were raised on standard cornmeal-agar-molasses medium at 25°C. Transgenic constructs were injected into *yw* embryos. C155 *elav*-GAL4 flies were obtained from Corey Goodman (Lin and Goodman, 1994) and Gary Struhl provided the UAS- (cytoplasmic) lacZ stock.

Preparation and differential screening of a Drosophila antennal/maxillary palp cDNA library

Drosophila antennae and maxillary palps were obtained by manually decapitating and freezing 5000 adult flies and shaking antennae and maxillary palps through a fine metal sieve. mRNA was prepared using a polyA+ RNA Purification Kit (Stratagene). An antennal/maxillary palp cDNA library was made from 0.5 μg mRNA using the LambdaZAPIIXR kit from Stratagene.

Briefly, phage were plated at low density (500-1000 pfu/150mm plate) and UV-crosslinked after lifting in triplicate to Hybond-N+ (Amersham). Complex probes were generated by random primed labeling (PrimeItII, Stratagene) of reverse transcribed mRNA (RT-PCR kit, Stratagene) from virgin adult female body mRNA and duplicate lifts hybridized at high stringency for 36 hours (65°C in 0.5M Sodium Phosphate buffer [pH7.3] containing 1% bovine serum albumin, 4% SDS, and 0.5 mg/ml herring sperm DNA). We prescreened the third lift with a mix of all previously cloned OBPs/PBPs (McKenna et al., 1994; Pikielny et al., 1994; Kim et al., 1998) remove a source of abundant but undesired olfactory-specific clones. Approximately 5000 individual OBP/PBP and virgin female body negative phage clones were isolated, their inserts amplified by PCR with T3 and T7 primers, and approximately 3 μg of DNA were electrophoresed on 1.5% agarose gels. Gels were blotted in duplicate to Hybond-N+ (Amersham), filters were

UV-crosslinked, and the resulting Southern blots were subjected to reverse Northern analysis using complex probes generated from virgin female body mRNA. Approximately 500 clones not hybridizing with virgin female body probes were identified and consolidated onto secondary Southern blots in triplicate. These blots were probed with complex probes derived from antennal/maxillary palp, head-minus-antenna/maxillary palp, and virgin female body mRNA. A total of 210 clones negative with head-minus-antenna/maxillary palp and virgin female body probes and strongly positive, weakly positive, or negative with antennal/maxillary palp probes were further analyzed by sequencing and *in situ* hybridization.

Analysis of Drosophila Genome Project Sequences for Transmembrane Proteins

All Drosophila genomic sequences were batch downloaded in April 1998 from the Berkeley Drosophila Genome Project (Berkeley Drosophila Genome Project, unpublished). Genomic P1 sequences were first analyzed with the GENSCAN program (Burge and Karlin, 1997; http://CCR-081.mit.edu/GENSCAN.html), which predicts intron-exon structures and generates hypothetical coding sequences (CDS) and open reading frames. GENSCAN predicted proteins shorter than 50 amino acids were discarded. The remaining open reading frames were used to search for putative transmembrane regions greater than 15 amino acids with two programs that were obtained from the authors and used in stand-alone mode locally (see Persson and Argos, 1994; Cserzo et al., 1997). The Dense Surface Alignment (DAS) program is available at http://www.biokemi.su.se/~server/DAS/ or from M. Cserzo (miklos@pugh.bip.bham.ac.uk). TMAP is available at ftp://ftp.ebi.ac.uk/pub/software/unix/, or by contacting the author, Bengt Persson (bpn@mbb.ki.se). Scripts were written to apply the DAS and TMAP programs repeatedly to genome scale sequence sets. Genes showing significant sequence similarity to the NCBI non-redundant protein database using BLAST analysis (Altschul et al., 1990; Altschul et al., 1997) were eliminated. All scripts required for these computations were written in standard ANSI C and run on a SUN Enterprise 3000.

Of 229 novel Drosophila proteins with three or more predicted transmembrane spanning regions, 35 showed no clear sequence similarity to any known protein and were selected for further analysis by *in situ* hybridization. Probes for *in situ* hybridization were generated by RT-PCR using antennal/maxillary palp mRNA as a template.

Map positions of DOR Genes

The chromosome position of DOR104 was determined by *in situ* hybridization of a biotin-labeled probe to salivary gland polytene chromosome squashes as described (Amrein et al., 1988).

Chromosomal positions of all other DOR genes were based on chromosome assignments of the P1 clones to which they map, as determined by the Berkeley Drosophila Genome Project (personal communication; http://www.fruitfly.org; see also Hartl et al., 1994; Kimmerly et al., 1996). DOR62 maps to a cosmid sequenced by the European Drosophila Genome Project (unpublished; http://edgp.ebi.ac.uk/; Siden-Kiamos et al., 1990).

RECEPTOR	MAP POSITION	P1 CLONE ACCESSION NUMBER
DOR62	(X) 2F	62D9 (EDGP cosmid)
DOR67	(2L) 22A3	DS00676
DOR53	(2L) 22A2-3	DS05342
DOR64	(2L) 23A1-2	DS06400
DOR71	(2L) 33B1-2	DS07071
DOR72	(2L) 33B1-2	DS07071
DOR73	(2L) 33B1-2	DS07071
DOR87	(2R) 43B1-2	DS08779
DOR19	(2R) 46F5-6	DS01913
DOR24	(2R) 47D6-E2	DS00724
DOR46	(2R) 59D5-7	DS07462
DOR104	(3L) 85B	not applicable

The Isolation of DOR cDNA Clones and Southern Blotting

We screened $3x10^6$ clones of the antennal/maxillary palp library described above with PCR probes for the genes DOR87, DOR53, DOR67, DOR64, and DOR62. cDNAs were present at a frequency ranging from 1:200,000 (DOR67) to 1:1,000,000 (DOR62) in the library and their sequences were remarkably similar to the hypothetical CDS predicted by the GENSCAN program. The frequency of these genes is similar to that of DOR104, which is present at 1:125,000 in the antennal/maxillary palp library. All sequencing was with ABI cycle sequencing kits and reactions were run on an ABI 310 or 377 sequencing system.

Five μg of Oregon R genomic DNA isolated from whole flies were digested with BamHI, EcoRI, or HindIII, electrophoresed on 0.8% agarose gels, and blotted to Nitropure nitrocellulose membranes (Micron Separations Inc.). Blots were baked and annealed with ^{32}P-labeled probes derived from cDNA probes of DOR53 and DOR67, or PCR fragments from DOR24, DOR62, and DOR72. Hybridization was at 42°C for 36 hours in 5XSSCP, 10X Denhardts, 500 μg/ml herring sperm DNA, and either 50% (high stringency) or 25% (low stringency) formamide (Sambrook et al., 1989). Blots were washed for 1 hour in 0.2X SSC, 0.5% SDS at 65°C (high stringency) or 1XSSC, 0.5% SDS at 42°C (low stringency).

***In situ* Hybridization**

RNA *in situ* hybridization was carried out essentially as described (Schaeren-Wiemers and Gerfin-Moser, 1993). This protocol was modified to include detergents in most steps to increase sensitivity and reduce background. The hybridization buffer contained 50% formamide, 5X SSC, 5X Denhardts, 250 μg/ml yeast tRNA, 500 μg/ml herring sperm DNA, 50 μg/ml Heparin, 2.5 mM EDTA, 0.1% Tween-20, 0.25% CHAPS. All antibody steps were in the presence of 0.1% Triton X-100, and the reaction was developed in buffer containing 0.1%

Tween-20. Slides were mounted in Glycergel (DAKO) and viewed with Nomarski optics.

Fluorescent *in situ* hybridization was carried out as above with either digoxigenin or FITC labeled RNA probes. The digoxigenin probe was visualized with sheep anti-digoxigenin (Boehringer) followed by donkey anti-sheep CY3 (Jackson). FITC probes were visualized with mouse anti-FITC (Boehringer) and goat anti-mouse Alexa 488 (Molecular Probes) following preincubation with normal goat serum. Sections were mounted in Vectashield reagent (Vector Labs) and viewed on a Biorad 1024 Confocal Microscope.

For double labeling with a neural marker, animals of the genotype C155 *elav*-Gal4; UAS-lacZ were sectioned and first hybridized with a digoxigenin labeled antisense DOR104 RNA probe and developed as described above. Neuron-specific expression of lacZ driven by the *elav*-Gal4 enhancer trap was visualized with a polyclonal rabbit anti-ß-galactosidase antibody (Organon-Technika/Cappel), visualized by a goat anti-rabbit Alexa488 conjugated secondary antibody (Molecular Probes) following preincubation with normal goat serum.

The proportion of neurons in the third antennal segment was calculated by comparing the number of nuclei staining with the 44C11 ELAV monoclonal (kindly provided by Lily Jan) and those staining with TOTO-3 (Molecular Probes), a nucleic acid counterstain, in several confocal sections of multiple antennae. On average, 36% of the nuclei in the antenna were ELAV positive.

<u>DOR104-lacZ Transgene Construction and Histochemical Staining</u>

A genomic clone containing the DOR104 coding region and several kb of upstream sequence was isolated from a genomic library prepared from flies isogenic for the third chromosome (a gift of Kevin Moses and Gerry Rubin). Approximately 3 kb of DNA immediately upstream of the putative translation start site of DOR104 were isolated by PCR and subcloned into the pCasperAUGßGal vector (Thummel et al., 1988). ß-galactosidase activity staining was carried out with whole mount head preparations essentially as described in Wang et al. (1998). Frozen sections of DOR104-lacZ maxillary palps were incubated with a polyclonal rabbit anti-ß-galactosidase antibody and as described above.

EXPERIMENTAL RESULTS

Cloning Candidate Odorant Receptors

In initial experiments, we isolated a cDNA encoding a putative odorant receptor by a difference cloning strategy designed to detect cDNA copies of mRNA present at extremely low frequencies in an mRNA population. In the antenna and maxillary palp, about 30% of the cells are olfactory neurons. If each neuron expressed only one of a possible 100 different odorant receptor genes at a level of 0.1% of the mRNA in a sensory neuron, then a given receptor mRNA would be encountered at a frequency of one in 300,000 in antennal mRNA. If 100 different receptor genes were expressed, then the entire family of receptor genes would be represented at a frequency of one in 3,000 mRNAs. We therefore introduced experimental modifications into standard difference cloning to allow for the identification of extremely rare mRNAs whose expression is restricted to either the antenna or the maxillary palp.

Briefly, 5000 insets from an antennal/maxillary palp cDNA library were prescreened (see Experimental Procedures) and then subjected to Southern blot hybridization with cDNA probes from antennal/maxillary palp, head minus antenna/maxillary palp, or virgin female body mRNA (see Figure 1). This Southern blot hybridization (or reverse Northern) to candidate cDNAs allows for the detection of sequences present at a frequency of 1 in 100,000 in the probe, a sensitivity about one hundred-fold greater than that of plaque screening (see Experimental Procedures). This procedure led to the identification of multiple antennal/maxillary palp-specific cDNAs that were analyzed by DNA sequencing and *in situ* hybridization. One cDNA, DOR104 (for Drosophila Odorant Receptor) (Figure 1, Lane 9), encodes a putative seven-transmembrane domain protein with no obvious sequence similarity to known serpentine receptors (Figure 3). *In situ* hybridization revealed that this cDNA anneals to about 15% of the 120 sensory neurons within the maxillary palp but does not anneal with neurons in either the brain or antenna. Seven cells expressing DOR104 are shown in the frontal maxillary palp section in Figure 2A.

These observations suggested that DOR104 might be one member of a larger family of odorant receptor genes within the Drosophila genome. However, we were unable to identify additional genes homologous to DOR104 by low stringency hybridization to genomic DNA and cDNA libraries or upon analysis of linked genes in a genomic walk. We therefore analyzed the Drosophila genome database for families of multiple transmembrane domain proteins that share sequence similarity with DOR104. Sequences representing about 10% of the Drosophila genome were downloaded (Berkeley Drosophila Genome Project) and subjected to GENSCAN analysis (Burge and Karlin, 1997) to predict the intron-exon structure of all sequences within the database. Open reading frames greater than 50 amino acids were searched for proteins with three or more predicted transmembrane-spanning regions using the dense alignment surface (DAS) and TMAP algorithms (Persson and Argos, 1994; Cserzo et al., 1997; also see Experimental Procedures). Of 229 candidate genes identified in this manner, 11 encoded proteins that define a novel divergent family of presumed seven transmembrane domain proteins with sequence similarity to the DOR104 sequence. This family of candidate odorant receptors does not share any conserved sequence motifs with previously identified families of seven transmembrane domain receptors. cDNA clones containing the coding regions for 5 of the 11 genes identified by GENSCAN analysis have been isolated from an antennal/maxillary palp cDNA library and their sequences are provided in Figure 3. The remaining 6 protein sequences derive from GENSCAN predictions for intron-exon arrangement. Their organization conforms well to the actual structure determined from the cDNA sequences of other members of the gene family (Figure 3).

The receptors consist of a short extracellular N-terminal domain (usually less than 50 amino acids) and seven presumed membrane-spanning domains. Analysis of presumed transmembrane domains (Kyte and Doolittle, 1982; Persson and Argos, 1994; Cserzo et al., 1997) reveals multiple hydrophobic segments, but it is not possible from this analysis to unequivocally determine either the number or placement of the membrane spanning domains. At present, our assignment of transmembrane domains is therefore tentative.

The individual family members are divergent and most exhibit from 17-26% amino acid identity. Two linked clusters of receptor genes constitute small subfamilies of genes with significantly greater sequence conservation. Two linked genes, DOR53 and DOR67, exhibit 76% amino acid identity, whereas the three linked genes, DOR71, 72 and 73, reveal 30-55% identity (Figure 3; see below). Despite the divergence, each of the genes shares short, common motifs in fixed positions within the putative seven transmembrane domain structure that define these sequences as highly divergent members of a novel family of putative receptor molecules.

Expression of the DOR Gene Family in Olfactory Neurons

If this gene family encodes putative odorant receptors in the fly, we might expect that other members of the family in addition to DOR104 would also be expressed in olfactory sensory neurons. We therefore performed *in situ* hybridization to examine the pattern of receptor expression of each of the 11 additional members of the gene family in adult and developing organisms. In Drosophila, olfactory sensory neurons are restricted to the maxillary palp and third antennal segment. The third antennal segment is covered with approximately 500 fine sensory bristles or sensilla (Stocker, 1994), each containing from one to four neurons (Venkatesh and Singh, 1984). The maxillary palp is covered with approximately 60 sensilla, each of which is innervated by two or three neurons (Singh and Nayak, 1985). Thus, the third antennal segment and maxillary palp contain about 1500 and 120 sensory neurons, respectively.

RNA *in situ* hybridization experiments were performed with digoxigenin-labeled RNA antisense probes to each of the 11 new members of the gene family under conditions of high stringency. One linked pair of homologous genes, DOR53 and DOR67, crosshybridizes, whereas the remaining 10 genes exhibit no crosshybridization under these conditions (see below). Eight of the 11 genes hybridize to a small subpopulation (0.5-1.5%) of the 1500 olfactory sensory neurons in the third antennal segment (Figure 4). One gene, DOR71, is expressed in about 10% of the sensory neurons in the maxillary palp but not in the antenna (Figure 4G). We have not detected expression of DOR46 or DOR19 in the antenna or the maxillary palp. Expression of this gene family is only observed in cells within the antenna and maxillary palp. No hybridization was observed in neurons of the brain, nor was hybridization observed in any sections elsewhere in the adult fly or in any tissue at any stage during embryonic development. However, we do find hybridization to a small number of cells in the developing antennae in the late pupal stage (data not shown). We have not yet determined whether this family of receptors is expressed in the larval olfactory apparatus.

Only about one third of the cells in the third antennal segment and the maxillary palp are neurons (data not shown), which are interspersed with non-neuronal sensillar support cells and glia. We have performed two experiments to demonstrate that the family of seven transmembrane domain receptor genes is expressed in sensory neurons rather than support cells or glia within the antenna and maxillary palp. First, we developed two-color fluorescent antibody detection schemes to co-localize receptor expression in cells that express the neuron-specific RNA binding protein, ELAV (Robinow and White, 1988). An enhancer trap line carrying an insertion of GAL4 at the *elav* locus expresses high levels of lacZ in neurons when crossed to a transgenic UAS-lacZ responder line (Lin and Goodman, 1994). Fluorescent antibody detection of lacZ identifies the sensory neurons in a horizontal section of the maxillary palp (Figure 5B). Hybridization with the receptor probe DOR104 reveals expression in 5 of the 12 lacZ positive cells in a horizontal section of the maxillary palp (Figure 5A). All cells that express DOR104 are also positive for lacZ (Figure 5C), indicating that this receptor is expressed only in neurons.

In a second experiment we have demonstrated that the receptor genes are not expressed in non-neuronal cells. The support cells of the antenna express different members of a family of odorant binding proteins (McKenna et al., 1994; Pikielny et al., 1994; Kim et al., 1998). These genes encode abundant low molecular weight proteins thought to transport odorants through the sensillar lymph (reviewed in Pelosi, 1994). Two-color *in situ* experiments with a probe for the odorant binding protein, PBPRP2 (Pikielny et al., 1994), reveal hybridization to a large number of cells broadly distributed throughout the antenna (Figure 5F). In the same section, however, the probe DOR53 anneals to a non-overlapping subpopulation of neurons restricted to the medial-proximal domain of the antenna. In a similar experiment, *in situ* hybridization with the odorant binding protein, OS-F (McKenna et al., 1994), identifies a spatially restricted subpopulation of support cells in the antenna, whereas the DOR67 probe identifies a distinct subpopulation of neurons in a medial-proximal domain (Figure 5G). Thus, the putative odorant receptor genes are expressed in a subpopulation of sensory neurons distinct from the support cells that express the odorant binding proteins. Taken together, these data demonstrate that 10 of the 12 family members we have identified are expressed in small subpopulations of olfactory sensory neurons in the antenna and maxillary palp.

Spatially Defined Patterns of Receptor Expression

The *in situ* hybridization experiments reveal that each receptor is expressed in a spatially restricted subpopulation of neurons in the antenna or maxillary palp (Figure 4). The total number of cells expressing each receptor per antenna was obtained by counting the positive cells in serial sections of antennae from multiple flies. These numbers are presented in the legend of Figure 4. DOR67 and 53, for example, anneal to about 20 neurons on the medial proximal edge of the antenna (Figure 4A and B), whereas DOR62 and 87 anneal to subpopulations of 20 cells at the distal edge of the antenna (Figure 4C-D). Approximately 10 cells in the distal domain express DOR64 (Figure 4E). Each of the three linked genes DOR71, 72, and 73 is expressed in different neurons. DOR72 is expressed in approximately 15 antennal cells (Figure 4H), while DOR73 is expressed in 1 to 2 cells at the distal edge of the antenna (Figure 4I). In contrast, DOR71 is expressed in approximately 10 maxillary palp neurons but is not detected in the antenna (Figure 4G). The three sensillar types are represented in a coarse topographic map across the third antennal segment. The proximal-medial region, for example, contains largely basiconic sensilla. Receptors expressed in this region (DOR53 and 67) are therefore likely to be restricted to the large basiconic sensilla. More distal regions contain a mixture of all three sensilla types and it is therefore not possible from these data to assign specific receptors to specific sensillar types.

The spatial pattern of neurons expressing a given receptor is conserved between individuals. *In situ* hybridization with two receptor probes to three individual flies reveals that both the frequency and spatial distributions of the hybridizing neurons is conserved in different individuals (Figure 6). At present, we cannot determine the precision of this topographic map and can only argue that given receptors are expressed in localized domains.

In preliminary experiments, we have demonstrated that the spatial pattern of expression of one receptor, DOR104, can be recapitulated in transgenic flies with a promoter fragment flanking the DOR104 gene. The fusion of the presumed DOR104 promoter (consisting of 3 kb of 5' DNA immediately adjacent to the coding region) to the lacZ reporter gene has allowed us to visualize a subpopulation of neurons expressing DOR104 within the maxillary palp. Whole mount preparations of the heads of transgenic flies reveal a small subpopulation of sensory neurons within the maxillary palp whose cell bodies exhibit blue color after staining with X-gal (Figure 2B). The number of positive cells, approximately 20 per maxillary palp, corresponds well with that seen for DOR104 RNA expression. Immunofluorescent staining of sections with antibodies directed against ß-galactosidase more clearly reveals the dendrites and axons of these bipolar neurons in the maxillary palp (Figure 2C). Levels of lacZ expression in these transgenic lines are low and further amplification will be necessary to allow us to trace the axons to glomeruli in the antennal lobe. Nonetheless, the data suggest that the information governing the spatial pattern of DOR104 expression in a restricted subpopulation of maxillary palp neurons resides within 3 kb of DNA 5' to the DOR104 gene.

Individual Neurons Express Different Complements of Receptors

An understanding of the logic of olfactory discrimination in Drosophila will require a determination of the diversity and specificity of receptor expression in individual neurons. In the vertebrate olfactory epithelium, a given neuron is likely to express only one receptor from the family of 1,000 genes (Ngai et al., 1993; Ressler et al., 1993; Vassar et al., 1993; Chess et al., 1994; Dulac and Axel, unpublished). In the nematode C. elegans, however, individual chemosensory neurons are thought to express multiple receptor genes (Troemel et al., 1995). Our observations with the putative Drosophila odorant receptors indicate that a given receptor probe anneals with 0.5-1.5% of antennal neurons, suggesting that each cell expresses only a subset of receptor genes. If we demonstrate that each of the different receptor probes hybridizes with distinct, nonoverlapping subpopulations of neurons, this would provide evidence that neurons differ with respect to the receptors they express.

In situ hybridization was therefore performed with either a mix of five receptor probes (Figure 4F) or individually with each of the five probes (Figure 4A-E). We observe that the number of olfactory neurons identified with the mixed probe (about 60 per antenna) approximates the sum of the positive neurons detected with the five individual probes. These results demonstrate that individual receptors are expressed in distinct nonoverlapping populations of olfactory neurons.

We have performed an additional experiment using two-color RNA *in situ* hybridization to ask whether two receptor genes, DOR64 and DOR87, expressed in interspersed cells in the distal antenna are expressed in different neurons. Antisense RNA probes for the two genes were labeled with either digoxigenin- or FITC-UTP and were used in pairwise combinations in *in situ* hybridization to sections through the Drosophila antenna. Although these two genes are expressed in overlapping lateral-distal domains, two-color *in situ* hybridization reveals that neurons expressing DOR64 do not express DOR87, rather each gene is expressed in distinct cell populations (Figure 5D and E). Taken together, these data suggest that olfactory sensory neurons within the antenna are functionally distinct and express different complements of odorant receptors. At the extreme, the experiments are consistent with a model in which individual neurons express only a single receptor gene.

Our differential cloning procedure identified one additional gene, A45, which shares weak identity (24%) with the DOR gene family over a short region (93 amino acids). This gene, however, does not appear to be a classical member of the DOR family: it is far more divergent and significantly larger than the other family members (486 amino acids). This gene is expressed in all olfactory sensory neurons (data not shown). If A45 does encode a divergent odorant receptor, then it would be present in all sensory neurons along with different complements of the more classical members of the DOR gene family.

The Size and Organization of the Odorant Receptor Gene Family

How large is the family of odorant receptor genes in Drosophila? Unlike vertebrate odorant receptors, which share 40-98% sequence identity at the amino acid level, the fly receptors are extremely divergent. The extent of sequence similarity between receptor subfamilies ranges from 20-30%. The maxillary palp receptor DOR104 is the most distantly related member of the family with about 17% identity to the other receptor genes. Inspection of the receptor sequences suggests that Southern blot hybridizations, even those performed at low stringency, are unlikely to reveal multiple additional members of a gene family. In accord with this, Southern blot hybridization with receptor probes DOR24, 62, and 72, performed at either high or low stringency, reveals only a single hybridizing band following cleavage of genomic DNA with three different restriction endonucleases (Figure 7C-E). The two linked clusters of receptors contain genes with a greater degree of sequence conservation and define small subfamilies of receptor genes. A cluster of three receptors, DOR71, 72, and 73, is located at map position 33B1-2. The antennal receptors DOR72 and 73 are 55% identical and both exhibit about 30% identity to the third gene at the locus, DOR71, which is expressed in the maxillary palp. DOR67 and DOR53, members of a second subfamily, reside within 1 kb of each other at map position 22A2-3 and exhibit 76% sequence identity. Not surprisingly, these two linked genes crosshybridize at low stringency. Southern blots probed with either DOR67 or DOR53 reveal two hybridizing bands corresponding to the two genes within the subfamily but fail to detect additional subfamily members in the chromosome (Figure 7A and B).

The members of the receptor gene family described here are present on all but the small fourth chromosome. No bias is observed toward telomeric or centromeric regions. The map positions, as determined from P1 and cosmid clones (Berkeley Drosophila Genome Project; European Drosophila Genome Project) are provided in Experimental Procedures. A comparatively large number of receptor genes map to chromosome 2 because the Berkeley Drosophila Genome Project has concentrated its efforts on this chromosome. Unlike the distribution of odorant receptors in nematodes and mammals (Ben-Arie et al., 1994; Troemel et al., 1995; Robertson, 1998), only small linked arrays have been identified and the majority of the family members are isolated at multiple, scattered loci in the Drosophila genome.

The high degree of divergence among members of the Drosophila odorant receptor gene family is more reminiscent of the family of chemoreceptors in C. elegans than the more highly conserved odorant receptors of vertebrates. Estimates of the size of the Drosophila receptor gene family, therefore, cannot be obtained by either Southern blot hybridization or PCR analysis of genomic DNA. Rather, our estimates of the gene family derive from the statistics of small numbers. We detect 12 members of the odorant receptor gene family from a Drosophila genome database that includes roughly 10% of the genome. Recognizing a possible bias in our estimate, it seems reasonable at present to estimate that the odorant receptor family is likely to include 100 to 200 genes. This is in accord with independent estimates from *in situ* hybridization experiments that demonstrate that a given receptor probe hybridizes with 0.5-1.5% of the neurons. If we assume that a given neuron expresses only a single receptor gene, these observations suggest that the gene family would include 100 to 200 members.

EXPERIMENTAL DISCUSSION

The Size and Divergence of the Gene Family

We have identified a novel family of seven transmembrane domain proteins that is likely to encode the Drosophila odorant receptors. The number of different receptor genes expressed in the neurons of the antenna and maxillary palp will reflect the diversity and specificity of odor recognition in the fruit fly. How large is the Drosophila odorant receptor gene family? We have identified 11 members of this divergent gene family in the Drosophila DNA database. The potential for bias notwithstanding, it seems reasonable to assume then that since only 10% of genomic sequence has been deposited, this gene family is likely to contain from 100 to 200 genes. However, significant errors in our estimates could result from bias in the nature of the sequences represented in the 10% of the Drosophila genome analyzed to date. *In situ* hybridization experiments demonstrating that each of the receptor genes labels from 0.5-1.5% of the olfactory sensory neurons are in accord with the estimate of 100 to 200 receptor genes.

Several divergent odorant receptor gene families, each encoding seven transmembrane proteins, have been identified in vertebrate and invertebrate species. In mammals, volatile odorants are detected by a family of as many as 1,000 receptors each expressed in the main olfactory epithelium (Buck and Axel, 1991; Levy et al., 1991; Parmentier et al., 1992; Ben-Arie et al., 1994). This gene family shares features with the serpentine neurotransmitter receptors and is conserved in all vertebrates examined. Terrestrial vertebrates have a second anatomically and functionally distinct olfactory system, the vomeronasal organ, dedicated to the detection of pheromones. Vomeronasal sensory neurons express two distinct families of receptors each thought to contain from 100 to 200 genes: one novel family of serpentine receptors (Dulac and Axel, 1995), and a second related to the metabotropic neurotransmitter receptors (Herrada and Dulac, 1997; Matsunami and Buck, 1997; Ryba and Tirindelli, 1997).

In the invertebrate C. elegans, chemosensory receptors are organized into four gene families that share 20-40% sequence similarity within a family and essentially no sequence similarity between families (Troemel et al., 1995; Sengupta et al., 1996; Robertson, 1998). The four gene families in C. elegans together contain about 1,000 genes engaged in the detection of odors. The nematode receptors exhibit no sequence conservation with the three distinct families of vertebrate odorant receptor genes. Our studies reveal that Drosophila has evolved an additional divergent gene family of serpentine receptors comprised of from 100 to 200 genes. The observation that a similar function, chemosensory detection, is accomplished by at least eight highly divergent gene families, sharing little or no sequence similarity, is quite unusual.

Why is the evolutionary requirement for odorant receptors so often met by recruitment of novel gene families rather than exploiting pre-existing odorant receptor families in ancestral genomes? The character of natural odorants along with their physical properties (e.g. aqueous or volatile) represent important selectors governing the evolution of receptor gene families. The use of common “anthropomorphic” odorant sets in the experimental analysis of olfactory specificity has led to the prevailing view that significant overlap exists in the repertoire of perceived odors between different species. Studies of odorant specificity in different species often employ odors at artificially high concentrations and may present an inaccurate image of the natural repertoire of odorants. We simply do not know the nature of the odors that initially led to the ancestral choice of receptor genes during the evolution of the nematode, insect, or vertebrate species. Clearly, vastly different properties in salient odors could dictate the recruitment of new gene families to effect an old function, olfaction. The character of the odor is not the only evolutionary selector. Odorant receptors must interact with other components in the signal transduction pathway [G proteins (for review see Buck, 1996; Bargmann and Kaplan, 1998) and perhaps even RAMPs (McLatchie et al, 1998) and rho (Mitchell et al., 1998)] that may govern the choice of one family of serpentine receptors over another. Moreover, mammalian receptors not only recognize odorants in the environment but are likely to recognize guidance cues governing formation of a sensory map in the brain (Wang et al., 1998). Thus, the multiple properties required of the odorant receptors might change vastly over evolutionary time and this might underlie the independent origins of the multiple chemosensory receptor gene families.

Establishing a Topographic Map in the Antenna and the Brain

We observe that individual receptor genes in the fly are expressed in topographically conserved domains within the antenna. This highly ordered spatial distribution of receptor expression differs from that observed in the mammalian olfactory epithelium. In mammals, a given receptor can be expressed in one of four broad but circumscribed zones in the main olfactory epithelium (Ressler et al., 1993; Vassar et al., 1993). A given zone can express up to 250 different receptors and neurons expressing a given receptor within a zone appear to be randomly dispersed (Ressler et al., 1993; Vassar et al., 1993). The highly ordered pattern of expression observed in the Drosophila antenna might have important implications for patterning the projections to the antennal lobe. In visual, somatosensory, and auditory systems the peripheral receptor sheet is highly ordered and neighbor relations in the periphery are maintained in the projections to the brain. These observations suggest that the relative position of the sensory neuron in the periphery will determine the pattern of projections to the brain.

Our data on the spatial conservation of receptor expression in the antenna suggest that superimposed upon coarse spatial patterning of olfactory sensilla (Venkatesh and Singh, 1984; Ray and Rodrigues, 1995; Reddy et al., 1997) must be more precise positional information governing the choice of receptor expression. This spatial information might dictate the fixed topographic pattern of receptor expression in the peripheral receptor sheet and at the same time govern the ordered sensory projections to the brain. This relationship between positional identity and the pattern of neuronal projections has been suggested for both peripheral sensory neurons (Merritt and Whitington, 1995; Grillenzoni et al., 1998) and neurons in the embryonic central nervous system of Drosophila (Doe and Skeath, 1996).

Implications for Sensory Processing

In mammals, olfactory neurons express only one of the thousand odorant receptor genes. Neurons expressing a given receptor project with precision to 2 of the 1800 glomeruli in the mouse olfactory bulb. Odorants will therefore elicit spatially defined patterns of glomerular activity such that the quality of an olfactory stimulus is encoded by the activation of a specific combination of glomeruli (Stewart et al., 1979; Lancet et al., 1982; Kauer et al., 1987; Imamura et al., 1992; Mori et al., 1992; Katoh et al., 1993; Friedrich and Korsching, 1997). Moreover, the ability of an odorant to activate a combination of glomeruli allows for the discrimination of a diverse array of odors far exceeding the number of receptors and their associated glomeruli. In the nematode, an equally large family of receptor genes is expressed in 16 pairs of chemosensory cells, only three of which respond to volatile odorants (Bargmann and Horvitz, 1991; Bargmann et al., 1993). This immediately implies that a given chemosensory neuron will express multiple receptors and that the diversity of odors recognized by the nematode might approach that of mammals, but the discriminatory power is necessarily dramatically reduced.

What does the character of the gene family we have identified in Drosophila tell us about the logic of olfactory processing in this organism? We estimate that the Drosophila odorant receptors comprise a family of from 100 to 200 genes. Moreover, the pattern of expression of these genes in the third antennal segment suggests that individual sensory neurons express a different complement of receptors and, at the extreme, our data are consistent with the suggestion that individual neurons express one or a small number of receptors. As in the case of mammals, the problem of odor discrimination therefore reduces to a problem of the brain discerning which receptors have been activated by a given odorant. If the number of different types of neurons exceeds the number of glomeruli (43) (Stocker, 1994; Laissue et al., 1999), it immediately follows that a given glomerulus must receive input from more than one kind of sensory neuron. This implies that a single glomerulus will integrate multiple olfactory stimuli. One possible consequence of this model would be a loss of discriminatory power while maintaining the ability to recognize a vast array of odors. Alternatively, significant processing of sensory input may occur in the fly antennal lobe to afford discrimination commensurate with the large number of receptors.

This model of olfactory coding is in sharp contrast with the main olfactory system of vertebrates in which sensory neurons express only a single receptor and converge on only a single pair of spatially fixed glomeruli in the olfactory bulb. Moreover, each projection neuron in the mammalian bulb extends its dendrite to only a single glomerulus. Thus the integration and decoding of spatial patterns of glomerular activity, in vertebrates, must occur largely in the olfactory cortex. In the fruit fly, the observation that the number of receptors may exceed the number of glomeruli suggests that individual glomeruli will receive input from more than one type of sensory neuron. A second level of integration in the antennal lobe is afforded by subsets of projection neurons that elaborate extensive dendritic arbors that synapse with multiple glomeruli. Thus, the Drosophila olfactory system reveals levels of processing and integration of sensory input in the antennal lobe that is likely to be restricted to higher cortical centers in the main olfactory system of vertebrates.

Protein and Nucleic Acid (nt) Sequences of 55 Drosophila Odorant Receptor Genes

The following includes those genes first identified in 1998-1999. Protein sequences used single letter amino acid codes.

DOR10

MEKLRSYEDFIFMANMMFKTLGYDLFHTPKPWWRYLLVRGYFVLCTISNFYEASMVTT
RIIEWESLAGSPSKIMRQGLHFFYMLSSQLKFITFMINRKRLLQLSHRLKELYPHKEQ
NQRKYEVNKYYLSCSTRNVLYVYYFVMVVMALEPLVQSQFIVNVSLGTDLWMMCVSSQ
ISMHLGYLANMLASIRPSPETEQQDCDFLASIIKRHQLMIRLQKDVNYVFGLLLASNL
FTTSCLLCCMAYYTVVEGFNWEGISYMMLFASVAAQFYVVSSHGQMLIDLLMTITYRF
FAVIRQTVEK

DOR10nt

ATGGAAAAACTACGTTCCTATGAGGATTTCATCTTCATGGCCAACATGATGTTCAAGA
CCCTTGGCTACGATCTATTCCATACACCCAAACCCTGGTGGCGCTATCTGCTTGTGCG
AGGATACTTCGTTTTGTGCACGATCAGCAACTTTTACGAGGCTTCCATGGTGACGACA
AGGATAATTGAGTGGGAATCCTTGGCCGGAAGTCCCTCCAAAATAATGCGACAGGGTC
TGCACTTCTTTTACATGTTGAGTAGCCAATTGAAATTTATCACATTCATGATAAATCG
CAAACGCCTACTGCAGCTGAGCCATCGTTTGAAAGAGTTGTATCCTCATAAAGAGCAA
AATCAAAGGAAGTACGAGGTGAATAAATACTACCTATCCTGTTCCACGCGCAATGTTT
TGTACGTGTACTACTTTGTAATGGTCGTCATGGCACTGGAACCCCTCGTTCAGTCCCA
GTTCATAGTGAATGTGAGCCTGGGCACAGATCTGTGGATGATGTGCGTCTCAAGCCAA
ATATCGATGCACTTGGGCTATCTGGCCAATATGTTGGCCTCCATTCGACCAAGTCCAG
AAACGGAACAACAAGACTGTGACTTCTTGGCCAGCATTATAAAGAGACATCAACTAAT
GATCAGGCTTCAAAAGGACGTGAACTATGTTTTTGGACTCTTATTGGCATCTAATCTG
TTTACCACATCCTGTTTACTTTGCTGCATGGCGTACTATACCGTCGTCGAAGGTTTCA
ATTGGGAGGGCATTTCCTATATGATGCTCTTTGCTAGTGTAGCTGCCCAGTTCTACGT
TGTCAGCTCACACGGACAAATGTTAATAGATTTGTTGATGACCATCACATACAGATTT
TTCGCGGTTATACGACAAACTGTAGAAAAG

DOR104

MASLQFHGNVDADIRYDISLDPARESNLFRLLMGLQLANGTKPSPRLPKWWPKRLEMI
GKVLPKAYCSMVIFTSLHLGVLFTKTTLDVLPTGELQAITDALTMTIIYFFTGYGTIY
WCLRSRRLLAYMEHMNREYRHHSLAGVTFVSSHAAFRMSRNFTVVWIMSCLLGVISWG
VSPLMLGIRMLPLQCWYPFDALGPGTYTAVYATQLFGQIMVGMTFGFGGSLFVTLSLL
LLGQFDVLYCSLKNLDAHTKLLGGESVNGLSSLQEELLLGDSKRELNQYVLLQEHPTD
LLRLSAGRKCPDQGNAFHNALVECIRLHRFILHCSQELENLFSPYCLVKSLQITFQLC
LLVFVGVSGTREVLRIVNQLQYLGLTIFELLMFTYCGELLSRHSIRSGDAFWRGAWWK
HAHFIRQDILIFLVNSRRAVHVTAGKFYVMDVNRLRSVITQAFSFLTLLQKLAAKKTE
SEL

DOR104nt

GAATTCGGCACGAGCAGTCGATGGCCAGTCTTCAGTTCCACGGCAACGTCGATGCGGA
CATCAGGTATGATATTAGCCTGGATCCGGCTAGGGAATCGAATCTCTTCCGTCTGCTA
ATGGGACTCCAGTTGGCGAATGGCACGAAGCCATCGCCGCGGTTACCCAAATGGTGGC
CAAAGCGGCTGGAAATGATTGGTAAAGTGCTGCCCAAAGCCTATTGTTCCATGGTGAT
TTTCACCTCCCTGCATTTGGGTGTCCTGTTCACGAAAACCACACTGGATGTCCTGCCG
ACGGGGGAGCTGCAGGCCATAACGGATGCCCTCACCATGACCATAATATACTTTTTCA
CGGGCTACGGCACCATCTACTGGTGCCTGCGCTCCCGGCGCCTCTTGGCCTACATGGA
GCACATGAACCGGGAGTATCGCCATCATTCGCTGGCCGGGGTGACCTTTGTGAGTAGC
CATGCGGCCTTTAGGATGTCCAGAAACTTCACGGTGGTGTGGATAATGTCCTGCCTGC
TGGGCGTGATTTCCTGGGGCGTTTCGCCACTGATGCTGGGCATCCGGATGCTGCCGCT
CCAATGTTGGTATCCCTTCGACGCCCTGGGTCCCGGCACATATACGGCGGTCTATGCT
ACACAACTTTTCGGTCAGATCATGGTGGGCATGACCTTTGGATTCGGGGGATCACTGT
TTGTCACCCTGAGCCTGCTACTCCTGGGACAATTCGATGTGCTCTACTGCAGCCTGAA
GAACCTGGATGCCCATACCAAGTTGCTGGGCGGGGAGTCTGTAAATGGCCTGAGTTCG
CTGCAAGAGGAGTTGCTGCTGGGGGACTCGAAGAGGGAATTAAATCAGTACGTTTTGC
TCCAGGAGCATCCGACGGATCTGCTGAGATTGTCGGCAGGACGAAAATGTCCTGACCA
AGGAAATGCGTTTCACAACGCCTTGGTGGAATGCATTCGCTTGCATCGCTTCATTCTG
CACTGCTCACAGGAGTTGGAGAATCTATTCAGTCCATATTGTCTGGTCAAGTCACTGC
AGATCACCTTTCAGCTTTGCCTGCTGGTCTTTGTGGGCGTTTCGGGTACTCGAGAGGT
CCTGCGGATTGTCAACCAGCTACAGTACTTGGGACTGACCATCTTCGAGCTCCTAATG
TTCACCTATTGTGGCGAACTCCTCAGTCGGCATAGTATTCGATCTGGCGACGCCTTTT
GGAGGGGTGCGTGGTGGAAGCACGCCCATTTCATCCGCCAGGACATCCTCATCTTTCT
GGTCAATAGTAGACGTGCAGTTCACGTGACTGCCGGCAAGTTTTATGTGATGGATGTG
AATCGTCTAAGATCGGTTATAACGCAGGCGTTCAGCTTCTTGACTTTGCTGCAAAAGT
TGGCTGCCAAGAAGACGGAATCGGAGCTCTAAACTGGTACCACGCATCGATATTTATT
TAGCGCATTAAAAAAAAGTCGAGTAAAAGCAAAAAAAAAAAAAAAAAA

DOR105

MFEDIQLIYMNIKILRFWALLYDKNLRRYVCIGLASFHIFTQIVYMMSTNEGLTGIIR
NSYMLVLWINTVLRAYLLLADHDRYLALIQKLTEAYYDLLNLNDSYISEILDQVNKVG
KLMARGNLFFGMLTSMGFGLYPLSSSERVLPFGSKIPGLNEYESPYYEMWYIFQMLIT
PMGCCMYIPYTSLIVGLIMFGIVRCKALQHRLRQVALKHPYGDRDPRELREEIIACIR
YQQSIIEYMDHINELTTMMFLFELMAFSALLCALLFMLIIVSGTSQLIIVCMYINMIL
AQILALYWYANELREQNLAVATAAYETEWFTFDVPLRKNILFMMMRAQRPAAILLGNI
RPITLELFQNLLNTTYTFFTVLKRVYG

DOR105nt

ATGTTTGAAGACATTCAGCTAATCTACATGAATATCAAGATATTGCGATTCTGGGCCC
TGCTCTATGACAAAAACTTGAGGCGTTATGTGTGCATTGGACTGGCCTCATTCCACAT
CTTCACCCAAATCGTCTACATGATGAGTACCAATGAAGGACTAACCGGGATAATTCGT
AACTCATATATGCTCGTCCTTTGGATTAATACGGTGCTGCGAGCTTATCTCTTGCTGG
CGGATCACGACAGATATTTGGCTTTGATCCAAAAACTAACTGAGGCCTATTACGATTT
ACTGAATCTGAACGATTCGTATATATCGGAAATATTGGACCAGGTGAACAAGGTGGGA
AAGTTGATGGCTAGGGGCAATCTGTTCTTTGGCATGCTCACATCCATGGGATTCGGTC
TGTACCCATTGTCCTCCAGCGAAAGAGTCCTGCCATTTGGCAGCAAAATTCCTGGTCT
AAATGAGTACGAGAGTCCGTACTATGAGATGTGGTACATCTTTCAGATGCTCATCACC
CCGATGGGCTGTTGCATGTACATTCCGTACACCAGTCTGATTGTGGGCTTGATAATGT
TCGGCATTGTGAGGTGCAAGGCTTTGCAGCATCGCCTCCGCCAGGTGGCGCTTAAGCA
TCCGTACGGAGATCGCGATCCCCGTGAACTGAGGGAGGAGATCATAGCCTGCATACGT
TACCAGCAGAGCATTATCGAGTACATGGATCACATAAACGAGCTGACCACCATGATGT
TCCTATTCGAACTGATGGCCTTTTCGGCGCTGCTCTGTGCGCTGCTCTTTATGCTGAT
TATCGTCAGCGGCACCAGTCAGCTGATAATTGTTTGCATGTACATTAACATGATTCTG
GCCCAAATACTGGCCCTCTATTGGTATGCAAATGAGTTAAGGGAACAGAATCTGGCGG
TGGCCACCGCAGCCTACGAAACGGAGTGGTTCACCTTCGACGTTCCACTGCGCAAAAA
CATCCTGTTCATGATGATGAGGGCACAGCGGCCAGCTGCAATACTACTGGGCAATATA
CGCCCCATCACTTTGGAACTGTTCCAAAACCTACTGAACACAACCTATACATTTTTTA
CGGTTCTCAAGCGAGTCTACGGA

DOR107

MYPRFLSRNYPLAKHLFFVTRYSFGLLGLRFGKEQSWLHLLWLVFNFVNLAHCCQAEF
VFGWSHLRTSPVDAMDAFCPLACSFTTLFKLGWMWWRRQEVADLMDRIRLLIGEQEKR
EDSRRKVAQRSYYLMVTRCGMLVFTLGSITTGAFVLRSLWEMWVRRHQEFKFDMPFRM
LFHDFAHRMPWFPVFYLYSTWSGQVTVYAFAGTDGFFFGFTLYMAFLLQALRYDIQDA
LKPIRDPSLRESKICCQRLADIVDRHNEIEKIVKEFSGIMAAPTFVHFVSASLVIATS

VIDILLYSGYNIIRYVVYTFTVSSAIFLYCYGGTEMSTESLSLGEAAYSSAWYTWDRE
TRRRVFLIILRAQRPITVRVPFFAPSLPVFTSVIKFTGSIVALAKTIL

DOR107nt

ATGTATCCGCGATTCCTCAGCCGTAACTATCCGCTGGCCAAGCATTTGTTCTTCGTCA
CCAGATACTCCTTTGGCCTGCTGGGCCTGAGATTTGGCAAAGAGCAATCGTGGCTTCA
CCTCTTGTGGCTGGTGTTCAATTTCGTTAACCTGGCGCACTGCTGCCAGGCGGAGTTC
GTCTTCGGCTGGAGTCACTTGCGCACCAGTCCCGTGGATGCCATGGACGCCTTTTGTC
CTCTGGCCTGCAGTTTCACCACGCTCTTCAAGCTGGGATGGATGTGGTGGCGTCGCCA
GGAAGTAGCTGATCTAATGGACCGCATCCGCTTGCTCATCGGGGAGCAGGAGAAGAGG
GAGGACTCCCGGAGAAAGGTGGCTCAAAGGAGCTACTATCTCATGGTCACCAGGTGCG
GTATGCTGGTCTTCACCCTGGGCAGCATTACCACTGGAGCCTTCGTTCTGCGTTCCCT
TTGGGAAATGTGGGTGCGTCGTCATCAGGAGTTCAAATTCGATATGCCCTTTCGCATG
CTGTTCCACGACTTTGCGCATCGCATGCCCTGGTTTCCAGTTTTCTATCTCTACTCCA
CATGGAGTGGCCAGGTCACTGTGTACGCCTTTGCTGGTACAGATGGTTTCTTCTTTGG
CTTTACCCTCTACATGGCCTTCTTGCTGCAGGCCTTAAGATACGATATCCAGGATGCC
CTCAAGCCAATAAGAGATCCCTCGCTTAGGGAATCCAAAATCTGCTGTCAGCGATTGG
CGGACATCGTGGATCGCCACAATGAGATAGAGAAGATAGTCAAGGAATTTTCTGGAAT
TATGGCTGCTCCAACTTTTGTTCACTTCGTATCAGCCAGCTTAGTGATAGCCACCAGC
GTCATTGATATACTATTGTATTCCGGCTATAACATCATCCGTTACGTGGTGTACACCT
TCACGGTTTCCTCGGCCATCTTCCTCTATTGCTACGGAGGCACAGAAATGTCAACTGA
GAGCCTTTCCTTGGGAGAAGCAGCCTACAGCAGTGCCTGGTATACTTGGGATCGAGAG
ACCCGCAGGCGGGTCTTTCTCATTATCCTGCGTGCTCAACGACCCATTACGGTGAGGG
TGCCCTTTTTTGCACCATCGTTACCAGTCTTCACATCGGTCATCAAGTTTACAGGTTC
GATTGTGGCACTGGCTAAGACGATACTG

DOR108

MDKHKDRIESMRLILQVMQLFGLWPWSLKSEEEWTFTGFVKRNYRFLLHLPITFTFIG
LMWLEAFISSNLEQAGQVLYMSITEMALVVKILSIWHYRTEAWRLMYELQHAPDYQLH
NQEEVDFWRREQRFFKWFFYIYILISLGVVYSGCTGVLFLEGYELPFAYYVPFEWQNE
RRYWFAYGYDMAGMTLTCISNITLDTLGCYFLFHISLLYRLLGLRLRETKNMKNDTIF
GQQLRAIFIMHQRIRSLTLTCQRIVSPYILSQIILSALIICFSGYRLQHVGIRDNPGQ
FISMLQFVSVMILQIYLPCYYGNEITVYANQLTNEVYHTNWLECRPPIRKLLNAYMEH
LKKPVTIRAGNSFAVGLPIFVKTINNAYSFLALLLNVSN

DOR108nt

ATGGATAAACACAAGGATCGCATTGAATCCATGCGCCTAATTCTTCAGGTCATGCAAC
TATTTGGCCTCTGGCCGTGGTCCTTGAAATCGGAAGAGGAGTGGACTTTCACCGGTTT
TGTAAAGCGCAACTATCGCTTCCTGCTCCATCTGCCCATTACCTTCACCTTTATTGGA
CTCATGTGGCTGGAGGCCTTCATCTCGAGCAATCTGGAGCAGGCTGGCCAGGTTCTGT
ACATGTCCATCACCGAGATGGCTTTGGTGGTGAAAATCCTGAGCATTTGGCACTATCG
CACCGAAGCTTGGCGGCTGATGTACGAACTCCAACATGCTCCGGACTACCAACTCCAC
AACCAGGAGGAGGTAGACTTTTGGCGCCGGGAGCAACGATTCTTCAAGTGGTTCTTCT
ACATCTACATTCTGATTAGCTTGGGCGTGGTATATAGTGGCTGCACTGGAGTACTTTT
TCTGGAGGGCTACGAACTGCCCTTTGCCTACTACGTGCCCTTCGAATGGCAGAACGAG
AGAAGGTACTGGTTCGCCTATGGTTACGATATGGCGGGCATGACGCTGACCTGCATCT
CAAACATTACCCTGGACACCCTGGGTTGCTATTTCCTGTTCCATATCTCTCTTTTGTA
CCGACTGCTTGGTCTGCGATTGAGGGAAACGAAGAATATGAAGAATGATACCATTTTT
GGCCAGCAGTTGCGTGCCATCTTCATTATGCATCAGAGGATTAGAAGCCTAACCCTGA
CCTGCCAGAGAATCGTATCTCCCTATATCCTATCTCAGATCATTTTGAGTGCCCTGAT
CATCTGCTTTAGTGGATACCGCTTGCAGCATGTGGGAATTCGCGATAATCCCGGCCAG
TTTATATCCATGTTGCAGTTTGTCAGTGTGATGATCCTGCAGATTTACTTGCCCTGCT
ACTATGGAAACGAGATAACCGTGTATGCCAATCAGCTGACCAACGAGGTTTACCATAC
CAATTGGCTGGAATGTCGGCCACCGATTCGAAAGTTACTCAATGCCTACATGGAGCAC
CTGAAGAAACCGGTGACCATCCGGGCTGGCAACTCCTTCGCCGTGGGACTACCAATTT
TTGTTAAGACCATCAACAACGCCTACAGTTTCTTGGCTTTATTACTAAATGTATCGAA
T

DOR109

MESTNRLSAIQTLLVIQRWIGLLKWENEGEDGVLTWLKRIYPFVLHLPLTFTYIALMW
YEAITSSDFEEAGQVLYMSITELALVTKLLNIWYRRHEAASLIHELQHDPAFNLRNSE
EIKFWQQNQRNFKRIFYWYIWGSLFVAVMGYISVFFQEDYELPFGYYVPFEWRTRERY
FYAWGYNVVAMTLCCLSNILLDTLGCYFMFHIASLFRLLGMRLEALKNAAEEKARPEL
RRIFQLHTKVRRLTRECEVLVSPYVLSQVVFSAFIICFSAYRLVHMGFKQRPGLFVTT
VQFVAVMIVQIFLPCYYGNELTFHANALTNSVFGTNWLEYSVGTRKLLNCYMEFLKRP
VKVRAGVFFEIGLPIFVKTINNAYSFFALLLKISK

DOR109nt

ATGGAGTCTACAAATCGCCTAAGTGCCATCCAAACACTTTTAGTAATCCAACGTTGGA
TAGGACTTCTTAAATGGGAAAACGAGGGCGAGGATGGAGTATTAACCTGGCTAAAACG
AATATATCCTTTTGTACTGCACCTTCCACTGACCTTCACGTATATTGCCTTAATGTGG
TATGAAGCTATTACATCGTCAGATTTTGAGGAAGCTGGTCAAGTTCTGTACATGTCCA

TCACCGAACTGGCATTGGTCACTAAACTGCTGAATATTTGGTATCGTCGTCATGAAGC
TGCTAGTCTAATCCACGAATTGCAACACGATCCCGCATTTAATCTGCGCAATTCGGAG
GAAATCAAATTCTGGCAGCAAAATCAGAGGAACTTTAAGAGAATATTTTACTGGTACA
TCTGGGGCAGCCTTTTCGTGGCTGTAATGGGTTATATAAGCGTGTTTTTCCAGGAGGA
TTACGAGCTGCCCTTTGGCTACTACGTGCCATTCGAGTGGCGCACCAGGGAACGATAC
TTCTACGCTTGGGGCTATAATGTGGTGGCCATGACCCTGTGCTGTCTATCCAACATCC
TACTGGACACACTAGGCTGTTATTTCATGTTCCACATCGCCTCGCTTTTCAGGCTTTT
GGGAATGCGACTGGAGGCCTTGAAAAATGCAGCCGAAGAGAAAGCCAGACCGGAGTTG
CGCCGCATTTTCCAACTGCACACTAAAGTCCGCCGATTGACGAGGGAATGCGAAGTGT
TAGTTTCACCCTATGTTCTATCCCAAGTGGTCTTCAGTGCCTTCATCATCTGCTTCAG
TGCCTATCGACTGGTGCACATGGGCTTCAAGCAGCGACCTGGACTCTTCGTGACCACC
GTGCAATTCGTGGCCGTCATGATCGTCCAGATTTTCTTGCCCTGTTACTACGGCAATG
AGTTGACCTTTCATGCCAATGCACTCACTAATAGTGTCTTCGGTACCAATTGGCTGGA
GTACTCCGTGGGCACTCGCAAGCTGCTTAACTGCTACATGGAGTTCCTCAAGCGACCG
GTTAAAGTGCGAGCTGGGGTGTTCTTTGAAATAGGACTACCCATCTTTGTGAAGACCA
TCAACAATGCCTACAGTTTCTTCGCCCTGCTGCTAAAGATATCCAAG

DOR110

MLFNYLRKPNPTNLLTSPDSFRYFEYGMFCMGWHTPATHKIIYYITSCLIFAWCAVYL
PIGIIISFKTDINTFTPNELLTVMQLFFNSVGMPFKVLFFNLYISGFYKAKKLLSEMD
KRCTTLKERVEVHQGVVRCNKAYLIYQFIYTAYTISTFLSAALSGKLPWRIYNPFVDF
RESRSSFWKAALNETALMLFAVTQTLMSDIYPLLYGLILRVHLKLLRLRVESLCTDSG
KSDAENEQDLINYAAAIRPAVTRTIFVQFLLIGICLGLSMINLLFFADIWTGLATVAY
INGLMVQTFPFCFVCDLLKKDCELLVSAIFHSNWINSSRSYKSSLRYFLKNAQKSIAF
TAGSIFPISTGSNIKVAKLAFSVVTFVNQLNIADRLTKN

DOR110nt

ATGTTGTTCAACTATCTGCGAAAGCCGAATCCCACAAACCTTTTGACTTCTCCGGACT
CATTTAGATACTTTGAGTATGGAATGTTTTGCATGGGATGGCACACACCAGCAACGCA
TAAGATAATCTACTATATAACATCCTGTTTGATTTTTGCTTGGTGTGCCGTATACTTG
CCAATCGGAATCATCATTAGTTTCAAAACGGATATTAACACATTCACACCGAATGAAC
TGTTGACAGTTATGCAATTATTTTTCAATTCAGTGGGAATGCCATTCAAGGTTCTGTT
CTTCAATTTGTATATTTCTGGATTTTACAAGGCCAAAAAGCTCCTTAGCGAAATGGAC
AAACGTTGCACCACTTTGAAGGAGCGAGTGGAAGTGCACCAAGGTGTGGTCCGTTGCA
ACAAGGCCTACCTCATTTACCAGTTCATTTATACCGCGTACACTATTTCAACATTTCT
ATCGGCGGCTCTTAGTGGAAAATTGCCATGGCGCATCTATAATCCTTTTGTGGATTTT
CGAGAAAGTAGATCCAGTTTTTGGAAAGCTGCCCTCAACGAGACAGCACTTATGCTAT
TTGCTGTGACTCAAACCCTAATGAGTGATATATATCCACTGCTTTATGGTTTGATCCT
GAGAGTTCACCTCAAACTTTTGCGACTAAGAGTGGAGAGCCTGTGCACAGATTCTGGA

AAAAGCGATGCTGAAAACGAGCAAGATTTGATTAACTATGCTGCAGCAATACGACCAG
CGGTTACCCGCACAATTTTCGTTCAATTCCTCTTGATCGGAATTTGCCTTGGCCTTTC
AATGATCAATCTACTCTTCTTTGCCGACATCTGGACAGGATTGGCCACAGTGGCTTAC
ATCAATGGTCTAATGGTGCAGACATTTCCATTTTGCTTCGTTTGTGATCTACTCAAAA
AGGATTGTGAACTTCTTGTGTCGGCCATATTTCATTCCAACTGGATTAATTCAAGCCG
CAGTTACAAGTCATCTTTGAGATATTTTCTGAAGAACGCCCAGAAATCAATTGCTTTT
ACAGCCGGCTCTATTTTTCCCATTTCTACTGGCTCGAATATTAAGGTGGCTAAGCTGG
CATTTTCGGTGGTTACTTTTGTCAATCAACTTAACATAGCTGACAGATTGACAAAGAA
C

DOR111

MLFRKRKPKSDDEVITFDELTRFPMTFYKTIGEDLYSDRDPNVIRRYLLRFYLVLGFL
NFNAYVVGEIAYFIVHIMSTTTLLEATAVAPCIGFSFMADFKQFGLTVNRKRLVRLLD
DLKEIFPLDLEAQRKYNVSFYRKHMNRVMTLFTILCMTYTSSFSFYPAIKSTIKYYLM
GSEIFERNYGFHILFPYDAETDLTVYWFSYWGLAHCAYVAGVSYVCVDLLLIATITQL
TMHFNFIANDLEAYEGGDHTDEENIKYLHNLVVYHARALDINKKCTFQSSRIGHSAFN
QNWLPCSTKYKRILQFIIARSQKPASIRPPTFPPISFNTFMKVISMSYQFFALLRTTY
YG

DOR111nt

ATGCTGTTCCGCAAACGTAAGCCAAAAAGTGACGATGAAGTCATCACCTTCGACGAAC
TTACCCGGTTTCCGATGACTTTCTACAAGACCATCGGCGAGGATCTGTACTCCGATAG
GGATCCGAATGTGATAAGGCGTTACCTGCTACGTTTTTATCTGGTACTCGGTTTTCTC
AACTTCAATGCCTATGTGGTGGGCGAAATCGCGTACTTTATAGTCCATATAATGTCGA
CGACTACTCTTTTGGAGGCCACTGCAGTGGCACCGTGCATTGGCTTCAGCTTCATGGC
CGACTTTAAGCAGTTCGGTCTCACAGTGAATAGAAAGCGATTGGTCAGATTGCTGGAT
GATCTCAAGGAGATATTTCCTTTAGATTTAGAAGCGCAGCGGAAGTATAACGTATCGT
TTTACCGGAAACACATGAACAGGGTCATGACCCTATTCACCATCCTCTGCATGACCTA
CACCTCGTCATTTAGCTTTTATCCAGCCATCAAGTCGACCATAAAGTATTACCTTATG
GGATCGGAAATCTTTGAGCGCAACTACGGATTTCACATTTTGTTTCCCTACGACGCAG
AAACGGATCTGACGGTCTACTGGTTTTCCTACTGGGGATTGGCTCATTGTGCCTATGT
GGCCGGAGTTTCCTACGTCTGCGTGGATCTCCTGCTGATCGCGACCATAACCCAGCTG
ACCATGCACTTCAACTTTATAGCGAATGATTTGGAGGCCTACGAAGGAGGTGATCATA
CGGATGAAGAAAATATCAAATACCTGCACAACTTGGTCGTCTATCATGCCAGGGCGCT
GGATATTAACAAGAAATGTACATTTCAGAGCTCTCGGATTGGCCATTCGGCATTTAAT
CAGAACTGGTTGCCATGCAGCACCAAATACAAACGCATCCTGCAATTTATTATCGCGC
GCAGCCAGAAGCCCGCCTCTATAAGACCGCCTACCTTTCCACCCATATCTTTTAATAC
CTTTATGAAGGTAATCAGCATGTCGTATCAGTTTTTTGCACTGCTCCGCACCACATAT

TATGGT

DOR114

MLTKKDTQSAKEQEKLKAIPLHSFLKYANVFYLSIGMMAYDHKYSQKWKEVLLHWTFI
AQMVNLNTVLISELIYVFLAIGKGSNFLEATMNLSFIGFVIVGDFKIWNISRQRKRLT
QVVSRLEELHPQGLAQQEPYNIGHHLSGYSRYSKFYFGMHMVLIWTYNLYWAVYYLVC
DFWLGMRQFERMLPYYCWVPWDWSTGYSYYFMYISQNIGGQACLSGQLAADMLMCALV
TLVVMHFIRLSAHIESHVAGIGSFQHDLEFLQATVAYHQSLIHLCQDINEIFGVSLLS
NFVSSSFIICFVGFQMTIGSKIDNLVMLVLFLFCAMVQVFMIATHAQRLVDASEQIGQ
AVYNHDWFRADLRYRKMLILIIKRAQQPSRLKATMFLNISLVTVSDLLQLSYKFFALL
RTMYVN

DOR114nt

ATGTTGACTAAGAAGGATACTCAAAGTGCCAAGGAGCAGGAAAAGTTGAAGGCCATTC
CATTGCACAGCTTTCTGAAATATGCCAACGTGTTCTATTTATCGATTGGAATGATGGC
CTACGATCACAAGTACAGTCAAAAGTGGAAGGAGGTCCTGCTGCACTGGACATTCATT
GCCCAGATGGTCAATCTGAATACAGTGCTCATCTCGGAACTGATTTACGTATTCCTGG
CGATCGGCAAAGGTAGCAATTTTCTGGAGGCCACCATGAATCTGTCTTTCATTGGATT
TGTCATCGTTGGTGACTTCAAAATCTGGAACATTTCGCGGCAGAGAAAGAGACTCACC
CAAGTGGTCAGCCGATTGGAAGAACTGCATCCGCAAGGCTTGGCTCAACAAGAACCCT
ATAATATAGGGCATCATCTGAGCGGCTATAGCCGATATAGCAAATTTTACTTCGGCAT
GCACATGGTGCTGATATGGACGTACAACCTGTATTGGGCCGTTTACTATCTGGTCTGT
GATTTCTGGCTGGGAATGCGTCAATTTGAGAGGATGCTGCCCTACTACTGCTGGGTTC
CCTGGGATTGGAGTACCGGATATAGCTACTATTTCATGTATATCTCACAGAATATCGG
CGGTCAGGCTTGTCTGTCCGGTCAGCTAGCAGCTGACATGTTAATGTGCGCCCTGGTC
ACTTTGGTGGTGATGCACTTCATCCGGCTTTCCGCTCACATCGAGAGTCATGTTGCGG
GCATTGGCTCATTCCAGCACGATTTGGAGTTCCTCCAAGCGACGGTGGCGTATCACCA
GAGCTTGATCCACCTCTGCCAGGATATCAATGAGATATTCGGTGTTTCACTGTTGTCC
AACTTTGTATCCTCGTCGTTTATCATCTGCTTCGTGGGTTTCCAGATGACCATCGGCA
GCAAGATCGACAACCTGGTAATGCTTGTGCTTTTCCTGTTTTGTGCCATGGTTCAGGT
CTTCATGATTGCCACCCATGCTCAGAGGCTCGTTGATGCGAGTGAACAGATTGGTCAA
GCGGTCTATAATCACGACTGGTTCCGTGCTGATCTGCGGTATCGTAAAATGCTGATCC
TGATTATTAAGAGGGCCCAACAGCCGAGTCGACTCAAGGCCACAATGTTCCTGAACAT
CTCACTGGTCACCGTGTCGGATCTCTTGCAACTCTCGTACAAATTCTTTGCCCTTCTG
CGCACAATGTACGTGAAT

DOR115

MEKLMKYASFFYTAVGIRPYTNGEESKMNKLIFHIVFWSNVINLSFVGLFESIYVYSA
FMDNKFLEAVTALSYIGFVTVGMSKMFFIRWKKTAITELINELKEIYPNGLIREERYN
LPMYLGTCSRISLIYSLLYSVLIWTFNLFCVMEYWVYDKWLNIRVVGKQLPYLMYIPW
KWQDNWSYYPLLFSQNFAGYTSAAGQISTDVLLCAVATQLVMHFDFLSNSMERHELSG
DWKKDSRFLVDIVRYHERILRLSDAVNDIFGIPLLLNFMVSSFVICFVGFQMTVGVPP
DIVVKLFLFLVSSMSQVYLICHYGQLVADASYGFSVATYNQKWYKADVRYKRALVIII
ARSQKVTFLKATIFLDITRSTMTDVRNCVLSV

DOR115nt

ATGGAGAAGCTAATGAAGTACGCTAGCTTCTTCTACACAGCAGTGGGCATACGGCCAT
ATACCAATGGTGAAGAATCCAAAATGAACAAACTTATATTTCACATAGTTTTTTGGTC
CAATGTGATTAACCTCAGCTTCGTTGGATTATTTGAGAGCATTTACGTTTACAGTGCC
TTCATGGATAATAAGTTCCTGGAAGCAGTCACTGCGTTGTCCTACATTGGCTTCGTAA
CCGTAGGCATGAGCAAGATGTTCTTCATCCGGTGGAAGAAAACGGCTATAACTGAACT
GATTAATGAATTGAAGGAGATCTATCCGAATGGTTTGATCCGAGAGGAAAGATACAAT
CTGCCGATGTATCTGGGCACCTGCTCCAGAATCAGCCTTATATATTCCTTGCTCTACT
CTGTTCTCATCTGGACATTCAACTTGTTTTGTGTAATGGAGTATTGGGTCTATGACAA
GTGGCTCAACATTCGAGTGGTGGGCAAACAGTTGCCGTACCTCATGTACATTCCTTGG
AAATGGCAGGATAACTGGTCGTACTATCCACTGTTATTCTCCCAGAATTTTGCAGGAT
ACACATCTGCAGCTGGTCAAATTTCAACCGATGTCTTGCTCTGCGCGGTGGCCACTCA
GTTGGTAATGCACTTCGACTTTCTCTCAAATAGTATGGAACGCCACGAATTGAGTGGA
GATTGGAAGAAGGACTCCCGATTTCTGGTGGACATTGTTAGGTATCACGAACGTATAC
TCCGCCTTTCAGATGCAGTGAACGATATATTTGGAATTCCACTACTACTCAACTTCAT
GGTATCCTCGTTCGTCATCTGCTTCGTGGGATTCCAGATGACTGTTGGAGTTCCGCCG
GATATAGTTGTGAAGCTCTTCCTCTTCCTTGTCTCTTCGATGAGTCAGGTCTATTTGA
TTTGTCACTATGGTCAACTGGTGGCCGATGCTAGCTACGGATTTTCGGTTGCCACCTA
CAATCAGAAGTGGTATAAAGCCGATGTGCGCTATAAACGAGCCTTGGTTATTATTATA
GCTAGATCGCAGAAGGTAACTTTTCTAAAGGCCACTATATTCTTGGATATTACCAGGT
CCACTATGACAGATGTACGCAACTGTGTATTGTCAGTG

DOR116

MELLPLAMLMYDGTRVTAMQYLIPGLPLENNYCYVVTYMIQTVTMLVQGVGFYSGDLF
VFLGLTQILTFADMLQVKVKELNDALEQKAEYRALVRVGASIDGAENRQRLLLDVIRW
HQLFTDYCRAINALYYELIATQVLSMALAMMLSFCINLSSFHMPSAIFFVVSAYSMSI
YCILGTILEFAYDQVYESICNVTWYELSGEQRKLFGFLLRESQYPHNIQILGVMSLSV
RTALQIVKLIYSVSMMMMNRA

DOR116nt

ATGGAACTCCTGCCATTGGCCATGCTAATGTACGATGGAACCCGGGTTACTGCGATGC
AGTATTTAATTCCGGGTCTACCGCTTGAGAACAATTATTGCTACGTAGTCACGTACAT
GATTCAGACGGTGACAATGCTCGTGCAAGGAGTCGGATTCTACTCCGGTGATTTGTTC
GTATTTCTCGGCTTAACGCAGATCCTAACTTTCGCCGATATGCTGCAGGTGAAGGTGA
AAGAGCTAAACGATGCCCTGGAACAAAAAGCGGAATACAGAGCTCTAGTCCGAGTTGG
AGCTTCTATTGATGGAGCGGAAAATCGTCAACGCCTTCTCTTGGATGTTATAAGATGG
CATCAATTATTCACGGACTACTGTCGCGCCATAAATGCCCTCTACTACGAATTGATCG
CCACTCAGGTTCTTTCGATGGCTTTGGCCATGATGCTCAGCTTCTGCATTAATTTGAG
CAGCTTTCACATGCCTTCGGCTATCTTTTTCGTGGTTTCTGCCTACAGCATGTCCATC
TATTGCATTCTGGGCACCATTCTTGAGTTTGCATATGACCAGGTGTACGAGAGCATCT
GTAATGTGACCTGGTATGAGTTGAGTGGCGAACAGCGAAAGCTTTTTGGTTTTTTGTT
GCGGGAATCCCAGTATCCGCACAATATTCAGATACTTGGAGTTATGTCGCTTTCCGTG
AGAACGGCTCTGCAGATTGTTAAACTAATTTATAGCGTATCCATGATGATGATGAATC
GGGCG

DOR117

MDLRRWFPTLYTQSKDSPVRSRDATLYLLRCVFLMGVRKPPAKFFVAYVLWSFALNFC
STFYQPIGFLTGYISHLSEFSPGEFLTSLQVAFNAWSCSTKVLIVWALVKRFDEANNL
LDEMDRRITDPGERLQIHRAVSLSNRIFFFFMAVYMVYATNTFLSAIFIGRPPYQNYY
PFLDWRSSTLHLALQAGLEYFAMAGACFQDVCVDCYPVNFVLVLRAHMSIFAERLRRL
GTYPYESQEQKYERLVQCIQDHKVILRFVDCLRPVISGTIFVQFLVVGLVLGFTLINI
VLFANLGSAIAALSFMAAVLLETTPFCILCNYLTEDCYKLADALFQSNWIDEEKRYQK
TLMYFLQKLQQPITFMAMNVFPISVGTNISVSRCAL

DOR117nt

ATGGATCTGCGAAGGTGGTTTCCGACCTTGTACACCCAGTCGAAGGATTCGCCAGTTC
GCTCCCGAGACGCGACCCTGTACCTCCTACGCTGCGTCTTCTTAATGGGCGTCCGCAA
GCCACCTGCCAAGTTTTTCGTGGCCTACGTGCTCTGGTCCTTCGCACTGAATTTCTGC
TCAACATTTTATCAGCCAATTGGCTTTCTCACAGGCTATATAAGCCATTTATCAGAGT
TCTCCCCGGGAGAGTTTCTAACTTCGCTGCAGGTGGCCTTTAATGCTTGGTCCTGCTC
TACAAAAGTCCTGATAGTGTGGGCACTAGTTAAGCGCTTTGACGAGGCTAATAACCTT
CTCGACGAGATGGATAGGCGTATCACAGACCCCGGAGAGCGTCTTCAGATTCATCGCG
CTGTCTCCCTCAGTAACCGTATATTCTTCTTTTTCATGGCAGTCTACATGGTTTATGC
CACTAATACGTTTCTGTCGGCGATCTTCATTGGAAGGCCACCGTACCAAAATTACTAC
CCTTTTCTGGACTGGCGATCTAGCACTCTGCATCTAGCTCTGCAGGCCGGTCTGGAAT

ACTTCGCCATGGCTGGCGCCTGCTTCCAGGACGTTTGCGTTGATTGCTACCCAGTCAA
TTTCGTTTTGGTCCTGCGTGCCCACATGTCGATCTTCGCGGAGCGCCTTCGACGTTTG
GGAACTTATCCTTATGAAAGCCAGGAGCAGAAATATGAACGATTGGTTCAGTGCATAC
AAGATCACAAAGTAATTTTGCGATTTGTTGACTGCCTGCGTCCTGTTATTTCTGGTAC
CATCTTCGTGCAATTCTTGGTTGTGGGGTTGGTGCTGGGCTTTACCCTAATTAACATT
GTCCTGTTCGCCAACTTGGGATCGGCCATCGCAGCGCTCTCGTTTATGGCCGCAGTGC
TTCTAGAGACGACTCCCTTCTGCATATTGTGCAATTATCTCACAGAAGACTGCTACAA
GCTGGCCGATGCCCTGTTTCAGTCAAACTGGATTGATGAGGAGAAACGATACCAAAAG
ACACTCATGTACTTCCTACAGAAACTGCAGCAGCCTATAACCTTCATGGCTATGAACG
TGTTTCCAATATCTGTGGGAACTAACATCAGTGTAAGCAGATGTGCCCTT

DOR118

MKFIGWLPPKQGVLRYVYLTWTLMTFVWCTTYLPLGFLGSYMTQIKSFSPGEFLTSLQ
VCINAYGSSVKVAITYSMLWRLIKAKNILDQLDLRCTAMEEREKIHLVVARSNHAFLI
FTFVYCGYAGSTYLSSVLSGRPPWQLYNPFIDWHDGTLKLWVASTLEYMVMSGAVLQD
QLSDSYPLIYTLILRAHLDMLRERIRRLRSDENLSEAESYEELVKCVMDHKLILRYCA
IIKPVIQGTIFTQFLLIGLVLGFTLINVFFFSDIWTGIASFMFVITILLQTFPFCYTC
NLIMEDCESLTHAIFQSNWVDASRRYKTTLLYFLQNVQQPIVFIAGGIFQISMSSNIS
VAKFAFSVITITKQMNIADKFKTD

DOR118nt

ATGAAGTTTATTGGATGGCTGCCCCCCAAGCAGGGTGTGCTCCGGTATGTGTACCTCA
CCTGGACGCTAATGACGTTCGTGTGGTGTACAACGTACCTGCCGCTTGGCTTCCTTGG
TAGCTACATGACGCAGATCAAGTCCTTCTCCCCTGGAGAGTTTCTCACTTCACTCCAG
GTGTGCATTAATGCCTACGGCTCATCGGTAAAAGTTGCAATCACATACTCCATGCTCT
GGCGCCTTATCAAGGCCAAGAACATTTTGGACCAGCTGGACCTGCGCTGCACCGCCAT
GGAGGAGCGCGAAAAGATCCACCTAGTGGTGGCCCGCAGCAACCATGCCTTTCTCATC
TTCACCTTTGTCTACTGCGGATATGCCGGCTCCACCTACCTGAGCTCGGTTCTCAGCG
GGCGTCCGCCCTGGCAGCTGTACAATCCCTTTATTGATTGGCATGACGGCACACTCAA
GCTCTGGGTGGCCTCCACGTTGGAGTACATGGTGATGTCAGGCGCCGTTCTGCAGGAT
CAACTCTCGGACTCTTACCCATTGATCTATACCCTCATCCTTCGTGCTCACTTGGACA
TGCTAAGGGAGCGCATCCGACGCCTCCGTTCCGATGAGAACCTGAGCGAGGCCGAGAG
CTATGAAGAGCTGGTCAAATGTGTGATGGACCACAAGCTCATTCTAAGATACTGCGCG
ATTATTAAACCAGTAATCCAGGGGACCATCTTCACACAGTTTCTGCTGATCGGCCTGG
TTCTGGGCTTCACGCTGATCAACGTGTTTTTCTTCTCAGACATCTGGACGGGCATCGC
ATCATTTATGTTTGTTATAACCATTTTGCTGCAGACCTTCCCCTTCTGCTACACATGC

AACCTCATCATGGAGGACTGCGAGTCCTTGACCCATGCTATTTTCCAGTCCAACTGGG
TGGATGCCAGTCGTCGCTACAAAACAACACTACTGTATTTTCTCCAAAACGTGCAGCA
GCCTATCGTTTTCATTGCAGGCGGTATCTTTCAGATATCCATGAGCAGCAACATAAGT
GTGGCAAAGTTTGCTTTCTCCGTGATAACCATTACCAAGCAAATGAATATAGCTGACA
AATTTAAGACGGAC

DOR119

MAVFKLIKPAPLTEKVQSRQGNIYLYRAMWLIGWIPPKEGVLRYVYLFWTCVPFAFGV
FYLPVGFIISYVQEFKNFTPGEFLTSLQVCINVYGASVKSTITYLFLWRLRKTEILLD
SLDKRLANDSDRERIHNMVARCNYAFLIYSFIYCGYAGSTFLSYALSGRPPWSVYNPF
IDWRDGMGSLWIQAIFEYITMSFAVLQDQLSDTYPLMFTIMFRAHMEVLKDHVRSLRM
DPERSEADNYQDLVNCVLDHKTILKCCDMIRPMISRTIFVQFALIGSVLGLTLVNVFF
FSNFWKGVASLLFVITILLQTFPFCYTCNMLIDDAQDLSNEIFQSNWVDAEPRYKATL
VLFMHHVQQPIIFIAGGIFPISMNSNITVAKFAFSIITIVRQMNLAEQFQATGGCGGT
GTTCAAGCTAATCAAACCGGCTCCGTTGACCGAGAAGGTGCAGTCCCGCCAGGGGAAT
ATATATCTGTACCGTGCCATGTGGCTCATCGGATGGATTCCGCCGAAGGAGGGAGTCC
TGCGCTACGTGTATCTCTTCTGGACCTGCGTGCCCTTCGCCTTCGGGGTGTTTTACCT
GCCCGTGGGCTTCATCATCAGCTACGTGCAGGAGTTCAAGAACTTCACGCCGGGCGAG
TTCCTTACCTCGCTGCAGGTGTGCATCAATGTGTATGGCGCCTCGGTGAAGTCCACCA
TCACCTACCTCTTCCTCTGGCGACTGCGCAAGACGGAGATCCTTCTGGACTCCCTGGA
CAAGAGGCTGGCGAACGACAGCGATCGCGAGAGGATCCACAATATGGTGGCGCGCTGC
AACTACGCCTTTCTCATCTACAGCTTCATCTACTGCGGATACGCGGGTTCCACTTTCC
TGTCCTACGCCCTCAGTGGTCGTCCTCCGTGGTCCGTCTACAATCCCTTCATCGATTG
GCGCGATGGCATGGGCAGCCTGTGGATCCAGGCCATATTCGAGTACATCACCATGTCC
TTCGCCGTGCTGCAGGACCAGCTATCCGACACGTATCCCCTGATGTTCACCATTATGT
TCCGGGCCCACATGGAGGTCCTCAAGGATCACGTGCGGAGCCTGCGCATGGATCCCGA
GCGCAGTGAGGCAGACAACTATCAGGATCTGGTGAACTGCGTGCTGGACCACAAGACT
ATACTGAAATGCTGTGACATGATTCGCCCCATGATATCCCGCACCATCTTCGTGCAAT
TCGCGCTGATTGGTTCCGTTTTGGGCCTGACCCTGGTGAACGTGTTCTTCTTCTCGAA
CTTCTGGAAGGGCGTGGCCTCGCTCCTGTTCGTCATCACCATCCTGCTGCAGACCTTC
CCGTTCTGCTACACCTGCAACATGCTGATCGACGATGCCCAGGATCTGTCCAACGAGA
TTTTCCAGTCCAACTGGGTGGACGCGGAGCCGCGCTACAAGGCGACGCTGGTGCTCTT
CATGCACCATGTTCAGCAGCCCATAATCTTCATTGCCGGAGGCATCTTTCCCATCTCT
ATGAACAGCAACATAACCGTGGCCAAGTTCGCCTTCAGCATCATTACAATAGTGCGAC
AAATGAATCTGGCCGAGCAGTTCCAG

DOR120

MTKFFFKRLQTAPLDQEVSSLDASDYYYRIAFFLGWTPPKGALLRWIYSLWTLTTMWL
GIVYLPLGLSLTYVKHFDRFTPTEFLTSLQVDINCIGNVIKSCVTYSQMWRFRRMNEL
ISSLDKRCVTTTQRRIFHKMVARVNLIVILFLSTYLGFCFLTLFTSVFAGKAPWQLYN
PLVDWRKGHWQLWIASILEYCVVSIGTMQELMSDTYAIVFISLFRCHLAILRDRIANL
RQDPKLSEMEHYEQMVACIQDHRTIIQCSQIIRPILSITIFAQFMLVGIDLGLAAISI
LFFPNTIWTIMANVSFIVAICTESFPCCMLCEHLIEDSVHVSNALFHSNWITADRSYK
SAVLYFLHRAQQPIQFTAGSTFPISVQSNIAVAKFAFTIITIVNQMNLGEKFFSDRSN
GDINP

DOR120nt

ATGACCAAGTTCTTCTTCAAGCGCCTGCAAACTGCTCCACTTGATCAGGAGGTGAGTT
CCCTTGATGCCAGCGACTACTACTACCGCATCGCATTTTTCCTGGGCTGGACCCCGCC
CAAGGGGGCTCTGCTCCGATGGATCTACTCCCTGTGGACTCTGACCACGATGTGGCTG
GGTATCGTGTACCTGCCGCTCGGACTGAGCCTCACCTATGTGAAGCACTTCGATAGAT
TCACGCCGACGGAGTTCCTGACCTCCCTGCAGGTGGATATCAACTGCATCGGGAACGT
GATCAAGTCATGCGTAACTTATTCCCAGATGTGGCGTTTTCGCCGGATGAATGAGCTT
ATCTCGTCCCTGGACAAGAGATGTGTGACTACGACACAGCGTCGAATTTTCCATAAGA
TGGTGGCACGGGTTAATCTCATCGTGATTCTGTTCTTGTCCACGTACTTGGGCTTCTG
CTTTCTAACTCTGTTCACTTCGGTTTTCGCTGGCAAAGCTCCTTGGCAGCTGTACAAC
CCACTGGTGGACTGGCGGAAAGGCCATTGGCAGCTATGGATTGCCTCCATCCTGGAGT
ACTGTGTGGTCTCCATTGGCACCATGCAGGAGTTGATGTCCGACACCTACGCCATAGT
GTTCATCTCCTTGTTCCGCTGCCACCTGGCTATTCTCAGAGATCGCATAGCTAATCTG
CGGCAGGATCCGAAACTCAGTGAGATGGAACACTATGAGCAGATGGTGGCCTGCATTC
AGGATCATCGAACCATCATACAGTGCTCCCAGATTATTCGACCCATCCTGTCGATCAC
TATCTTTGCCCAGTTCATGCTGGTTGGCATTGACTTGGGTCTGGCGGCCATCAGCATC
CTCTTCTTTCCGAACACCATTTGGACGATCATGGCAAACGTGTCGTTCATCGTGGCCA
TCTGTACAGAGTCCTTTCCATGCTGCATGCTCTGCGAGCATCTGATCGAGGACTCCGT
CCATGTGAGCAACGCCCTGTTCCACTCAAACTGGATAACCGCGGACAGGAGCTACAAG
TCGGCGGTTCTGTATTTCCTGCACCGGGCTCAGCAACCCATTCAATTCACGGCCGGCT
CCATATTTCCCATTTCGGTGCAGAGCAACATAGCCGTGGCCAAGTTCGCGTTCACAAT
CATCACAATCGTGAACCAAATGAATCTGGGCGAGAAGTTCTTCAGTGACAGGAGCAAT
GGCGATATAAATCCT

DOR121

MLTDKFLRLQSALFRLLGLELLHEQDVGHRYPWRSICCILSVASFMPLTIAFGLQNVQ
NVEQLTDSLCSVLVDLLALCKIGLFLWLYKDFKFLIGQFYCVLQTETHTAVAEMIVTR

ESRRDQFISAMYAYCFITAGLSACLMSPLSMLISYHEQVNCSRNFHFPVCKKKYCLIS
RILRYSFCRYPWDNMKLSNYIISYFWNVCAALGVALPTVCVDTLFCSLSHNLCALFQI
ARHKMMHFEGRNTKETHENLKHVFQLYALCLNLGHFLNEYFRPLICQFVAASLHLCVL
CYQLSANILQPALLFYAAFTAAVVGQVSIYCFCGSSIHSECQLFGQAIYESSWPHLLQ
ENLQLVSSLKIAMMRSSLGCPIDGYFFEANRETLITVSKAFIKVSKKTPQVND

DOR121

ATGCTGACGGACAAGTTCCTCCGACTGCAGTCCGCTTTATTTCGCCTTCTCGGACTCG
AATTGTTGCACGAGCAGGATGTTGGCCATCGATATCCTTGGCGCAGCATCTGCTGCAT
TCTCTCGGTGGCCAGTTTCATGCCCCTGACCATTGCGTTTGGCCTGCAAAACGTCCAA
AATGTGGAGCAATTAACCGACTCACTCTGCTCGGTTCTCGTGGATTTGCTGGCCCTGT
GCAAAATCGGGCTTTTCCTTTGGCTTTACAAGGACTTCAAGTTCCTAATAGGGCAGTT
CTATTGTGTTTTGCAAACGGAAACCCACACCGCTGTCGCTGAAATGATAGTGACCAGG
GAAAGTCGTCGGGATCAGTTCATCAGTGCTATGTATGCCTACTGTTTCATTACGGCTG
GCCTTTCGGCCTGCCTGATGTCCCCTCTATCCATGCTGATTAGCTACCACGAACAGGT
GAATTGCAGCCGAAATTTCCATTTCCCAGTGTGTAAGAAAAAGTACTGCTTAATATCC
AGAATATTAAGATACAGTTTCTGCAGATATCCCTGGGACAATATGAAGCTGTCCAACT
ACATCATTTCCTATTTCTGGAATGTGTGTGCTGCATTGGGCGTGGCACTGCCCACCGT
TTGTGTGGACACACTGTTCTGTTCTCTGAGCCATAATCTCTGTGCCCTATTCCAGATT
GCCAGGCACAAAATGATGCACTTTGAGGGCAGAAATACCAAAGAGACTCATGAGAACT
TAAAGCACGTGTTTCAACTATATGCGTTGTGTTTGAACCTGGGCCATTTCTTAAACGA
ATATTTCAGACCGCTCATCTGCCAGTTTGTGGCAGCCTCACTGCACTTGTGTGTCCTG
TGCTACCAACTGTCTGCCAATATCCTGCAGCCAGCGTTACTCTTCTATGCCGCATTTA
CGGCAGCAGTTGTTGGCCAGGTGTCTATATACTGCTTCTGCGGATCGAGCATCCATTC
GGAGTGTCAGCTATTTGGCCAGGCCATCTACGAGTCCAGCTGGCCCCATCTGCTGCAG
GAAAACCTGCAGCTTGTAAGCTCCTTAAAAATTGCCATGATGCGATCGAGTTTGGGAT
GTCCCATCGATGGTTACTTCTTCGAGGCCAATCGGGAGACGCTCATCACGGTGAGTAA
AGCGTTTATAAAAGTGTCCAAAAAGACACCTCAAGTGAATGAT

DOR14

MDYDRIRPVRFLTGVLKWWRLWPRKESVSTPDWTNWQAYALHVPFTFLFVLLLWLEAI
KSRDIQHTADVLLICLTTTALGGKVINIWKYAHVAQGILSEWSTWDLFELRSKQEVDM
WRFEHRRFNRVFMFYCLCSAGVIPFIVIQPLFDIPNRLPFWMWTPFDWQQPVLFWYAF
IYQATTIPIACACNVTMDAVNWYLMLHLSLCLRMLGQRLSKLQHDDKDLREKFLELIH
LHQRLKQQALSIEIFISKSTFTQILVSSLIICFTIYSMQMDLPGFAAMMQYLVAMIMQ
VMLPTIYGNAVIDSANMLTDSMYNSDWPDMNCRMRRLVLMFMVYLNRPVTLKAGGFFH
IGLPLFTKVVFSTLENPCISYLYFRP

DOR14nt

ATGGACTACGATCGAATTCGACCGGTGCGATTTTTGACGGGAGTGCTGAAATGGTGGC
GTCTCTGGCCGAGGAAGGAATCGGTGTCCACACCGGACTGGACTAACTGGCAGGCATA
TGCCTTGCACGTTCCATTTACATTCTTGTTTGTGTTGCTTTTGTGGTTGGAGGCAATC
AAGAGCAGGGATATACAGCATACCGCCGATGTCCTTTTGATTTGCCTAACCACCACTG
CCTTGGGAGGTAAAGTTATCAATATCTGGAAGTATGCCCATGTGGCCCAAGGCATTTT
GTCCGAGTGGAGCACGTGGGATCTTTTCGAGCTGAGGAGCAAACAGGAAGTGGATATG
TGGCGATTCGAGCATCGACGTTTCAATCGTGTTTTTATGTTTTACTGTTTGTGCAGTG
CTGGTGTAATCCCATTTATTGTGATTCAACCGTTGTTTGATATCCCAAATCGATTGCC
CTTCTGGATGTGGACACCATTCGATTGGCAGCAGCCTGTTCTCTTCTGGTATGCATTC
ATCTATCAGGCCACAACCATTCCTATTGCCTGTGCTTGCAACGTAACCATGGACGCTG
TTAATTGGTACTTGATGCTGCATCTGTCCTTGTGTTTGCGTATGTTGGGCCAGCGATT
GAGTAAGCTTCAGCATGATGACAAGGATCTGAGGGAGAAGTTCCTGGAACTGATCCAT
CTGCACCAGCGACTCAAGCAACAGGCCTTGAGCATTGAAATCTTTATTTCGAAGAGCA
CGTTCACCCAAATTCTGGTCAGTTCCCTTATCATTTGCTTCACCATTTACAGCATGCA
GATGGACTTGCCAGGATTTGCCGCCATGATGCAGTACCTAGTGGCCATGATCATGCAG
GTCATGCTGCCCACCATATATGGTAACGCCGTCATCGATTCTGCAAATATGTTGACCG
ATTCCATGTACAATTCGGATTGGCCGGATATGAATTGCCGAATGCGTCGCCTAGTTTT
AATGTTTATGGTGTACTTAAATCGACCGGTGACCTTAAAAGCCGGTGGCTTTTTTCAT
ATTGGTTTACCTCTGTTTACCAAGGTTGTATTTTCTACTCTGGAAAATCCTTGTATAA
GTTATCTTTATTTCAGACCA

DOR16

MTDSGQPAIADHFYRIPRISGLIVGLWPQRIRGGGGRPWHAHLLFVFAFAMVVVGAVG
EVSYGCVHLDNLVVALEAFCPGTTKAVCVLKLWVFFRSNRRWAELVQRLRAILWESRR
QEAQRMLVGLATTANRLSLLLLSSGTATNAAFTLQPLIMGLYRWIVQLPGQTELPFNI
ILPSFAVQPGVFPLTYVLLTASGACTVFAFSFVDGFFICSCLYICGAFRLVQQDIRRI
FADLHGDSVDVFTEEMNAEVRHRLAQVVERHNAIIDFCTDLTRQFTVIVLMHFLSAAF
VLCSTILDIMLVSPFSEAFLWGGYPWVCRATGFSHRLHSAAVLKVFPCFHCLLFFPGF
SSRSVLIRFSRFVCLLCGCGCGSLRWQFISA

DOR16nt

ATGACTGACAGCGGGCAGCCTGCCATTGCCGACCACTTTTATCGGATTCCCCGCATCT
CCGGCCTCATTGTCGGCCTCTGGCCGCAAAGGATAAGGGGCGGGGGCGGTCGTCCTTG
GCACGCCCATCTGCTCTTCGTGTTCGCCTTCGCCATGGTGGTGGTGGGTGCGGTGGGC
GAGGTGTCGTACGGCTGTGTCCACCTGGACAACCTGGTGGTGGCGCTGGAGGCCTTCT
GCCCCGGAACCACCAAGGCGGTCTGCGTTTTGAAGCTGTGGGTCTTCTTCCGCTCCAA

TCGCCGGTGGGCGGAGTTGGTCCAGCGCCTGCGGGCTATTTGTGGGAATCGCGGCGG
CAGGAGGCCCAGAGGATGCTGGTCGGACTGGCCACCACGGCCAACAGGCTCAGCCTGT
TGTTGCTCAGCTCTGGCACGGCGACAAATGCCGCCTTCACCTTGCAACCGCTGATTAT
GGGTCTCTACCGCTGGATTGTGCAGCTGCCAGGTCAAACCGAGCTGCCCTTTAATATC
ATACTGCCCTCGTTTGCCGTGCAGCCAGGAGTCTTTCCGCTCACCTACGTGCTGCTGA
CCGCTTCCGGTGCCTGCACCGTTTTCGCCTTCAGCTTCGTGGACGGATTCTTCATTTG
CTCGTGCCTCTACATCTGCGGCGCTTTCCGGCTGGTGCAGCAGGACATTCGCAGGATA
TTTGCCGATTTGCATGGCGACTCAGTGGATGTGTTCACCGAGGAGATGAACGCGGAGG
TGCGGCACAGACTGGCCCAAGTTGTCGAGCGGCACAATGCGATTATCGATTTCTGCAC
GGACCTAACACGCCAGTTCACCGTTATCGTTTTAATGCATTTCCTGTCCGCCGCCTTC
GTCCTCTGCTCGACCATCCTGGACATCATGTTGGTGAGCCCCTTTTCAGAGGCCTTCC
TTTGGGGCGGGTATCCTTGGGTTTGTCGCGCCACTGGCTTTTCGCATCGCCTGCATTC
GGCGGCTGTTTTAAAAGTTTTTCCCTGTTTTCACTGTTTGCTGTTTTTCCCTGGCTTT
TCCAGCCGCTCCGTTCTGATTCGGTTTTCCCGATTTGTTTGTTTGCTTTGTGGCTGCG
GCTGCGGCTCTCTCCGGTGGCAATTTATAAGCGCATGA

DOR19

MVTEDFYKYQVWYFQILGVWQLPTWAADHQRRFQSMRFGFILVILFIMLLLFSFEMLN
NISQVREILKVFFMFATEISCMAKLLHLKLKSRKLAGLVDAMLSPEFGVKSEQEMQML
ELDRVAVVRMRNSYGIMSLGAASLILIVPCFDNFGELPLAMLEVCSIEGWICYWSQYL
FHSICLLPTCVLNITYDSVAYSLLCFLKVQLQMLVLRLEKLGPVIEPQDNEKIAMELR
ECAAYYNRIVRFKDLVELFIKGPGSVQLMCSVLVLVSNLYDMSTMSIANGDAIFMLKT
CIYQLVMLWQIFIICYASNEVTVQSSRLCHSIYSSQWTGWNRANRRIVLLMMQRFNSP
MLLSTFNPTFAFSLEAFGSVGQQKFLYISFITGYALLLSDRQLLLQLLRTAEARQQLN
FETPQHLKIFKPIFKSTQNVMHVH

DOR19nt

ATGGTTACGGAGGACTTTTATAAGTACCAGGTGTGGTACTTCCAAATCCTTGGTGTTT
GGCAGCTCCCCACTTGGGCCGCAGACCACCAGCGTCGTTTTCAGTCCATGAGGTTTGG
CTTCATCCTGGTCATCCTGTTCATCATGCTGCTGCTTTTCTCCTTCGAAATGTTGAAC
AACATTTCCCAAGTTAGGGAGATCCTAAAGGTATTCTTCATGTTCGCCACGGAAATAT
CCTGCATGGCCAAATTATTGCATTTGAAGTTGAAGAGCCGCAAACTCGCTGGCTTGGT
TGATGCGATGTTGTCCCCAGAGTTCGGCGTTAAAAGTGAACAGGAAATGCAGATGCTG
GAATTGGATAGAGTGGCGGTTGTCCGCATGAGGAACTCCTACGGCATCATGTCCCTGG
GCGCGGCTTCCCTGATCCTTATAGTTCCCTGTTTCGACAACTTTGGCGAGCTACCACT
GGCCATGTTGGAGGTATGCAGCATCGAGGGATGGATCTGCTATTGGTCGCAGTACCTT
TTCCACTCGATTTGCCTGCTGCCCACTTGTGTGCTGAATATAACCTACGACTCGGTGG
CCTACTCGTTGCTCTGTTTCTTGAAGGTTCAGCTACAAATGCTGGTCCTGCGATTAGA

AAAGTTGGGTCCTGTGATCGAACCCCAGGATAATGAGAAAATCGCAATGGAACTGCGT
GAGTGTGCCGCCTACTACAACAGGATTGTTCGTTTCAAGGACCTGGTGGAGCTGTTCA
TAAAGGGGCCAGGATCTGTGCAGCTCATGTGTTCTGTTCTGGTGCTGGTGTCCAACCT
GTACGACATGTCCACCATGTCCATTGCAAACGGCGATGCCATCTTTATGCTCAAGACC
TGTATCTATCAGCTGGTGATGCTCTGGCAGATCTTCATCATTTGCTACGCCTCCAACG
AGGTAACTGTCCAGAGCTCTAGGTTGTGTCACAGCATCTACAGCTCCCAATGGACGGG
ATGGAACAGGGCAAACCGCCGGATTGTCCTTCTCATGATGCAGCGCTTTAATTCCCCG
ATGCTCCTGAGCACCTTTAACCCCACCTTTGCTTTCAGCTTGGAGGCCTTTGGTTCTG
TAGGGCAGCAGAAATTCCTTTATATATCATTTATTACTGGTTATGCTCTTCTCCTTTC
AGATCGTCAACTGCTCCTACAGCTACTTCGCACTGCTGAAGCGCGTCAACAGTTAAAT
TTCGAAACACCGCAGCACCTAAAGATTTTCAAGCCGATTTTTAAAAGCACTCAAAACG
TTATGCACGTACAT

DOR20

MSKGVEIFYKGQKAFLNILSLWPQIERRWRIIHQVNYVHVIVFWVLLFDLLLVLHVMA
NLSYMSEVVKAIFILATSAGHTTKLLSIKANNVQMEELFRRLDNEEFRPRGANEELIF
AAACERSRKLRDFYGALSFAALSMILIPQFALDWSHLPLKTYNPLGENTGSPAYWLLY
CYQCLALSVSCITNIGFDSLCSSLFIFLKCQLDILAVRLDKIGRLITTSGGTVEQQLK
ENIRYHMTIVELSKTVERLLCKPISVQIFCSVLVLTANFYAIAVVSCEFATRRLSVCD
LSGVHVDSDFYIVLLCRVGIPYPKCLPRPVMNFIVSEVTQRSLDLPHELYKTSWVDWD
YRSRRIALLFMQRLHSTLRIRTLNPSLGFDLMLFSSVSSFRVLTFLCTVANFHNEAH

DOR20nt

ATGAGCAAAGGAGTAGAAATCTTTTACAAGGGCCAGAAGGCATTCTTGAACATCCTCT
CGTTGTGGCCTCAGATAGAACGCCGGTGGAGAATCATCCACCAGGTGAACTATGTCCA
CGTAATTGTGTTTTGGGTGCTGCTCTTTGATCTCCTCTTGGTGCTCCATGTGATGGCT
AATTTGAGCTACATGTCCGAGGTTGTGAAAGCCATCTTTATCCTGGCCACCAGTGCAG
GGCACACCACCAAGCTGCTGTCCATAAAGGCGAACAATGTGCAGATGGAGGAGCTCTT
TAGGAGATTGGATAACGAAGAGTTCCGTCCTAGAGGCGCCAACGAAGAGTTGATCTTT
GCAGCAGCCTGTGAAAGAAGTAGGAAGCTTCGGGACTTCTATGGAGCGCTTTCGTTTG
CCGCCTTGAGCATGATTCTCATACCCCAGTTCGCCTTGGACTGGTCCCACCTTCCGCT
CAAAACATACAATCCGCTTGGCGAGAATACCGGCTCACCTGCTTATTGGCTCCTCTAC
TGCTATCAGTGTCTGGCCTTGTCCGTATCCTGCATCACCAACATAGGATTCGACTCAC
TCTGCTCCTCACTGTTCATCTTCCTCAAGTGCCAGCTGGACATTCTGGCCGTGCGACT
GGACAAGATCGGTCGGTTAATCACTACTTCTGGTGGCACTGTGGAACAGCAACTTAAG
GAAAATATCCGCTATCACATGACCATCGTTGAACTGTCGAAAACCGTGGAGCGTCTAC
TTTGCAAGCCGATTTCGGTGCAGATCTTCTGCTCGGTTTTGGTGCTGACTGCCAATTT
CTATGCCATTGCTGTGGTGAGCTGTGAATTCGCAACAAGAAGACTATCAGTATGTGAC
CTATCAGGCGTGCATGTTGATTCAGATTTTTATATTGTGCTACTATGCCGGGTGGGTA

TTCCATATCCGAAATGCCTCCCCAGGCCAGTAATGAATTTCATCGTCAGTGAGGTAAC
CCAGCGCAGCCTGGACCTTCCGCACGAGCTGTACAAGACCTCCTGGGTGGACTGGGAC
TACAGGAGCCGAAGGATTGCGCTCCTCTTTATGCAACGCCTTCACTCGACCTTGAGGA
TTAGGACACTTAATCCAAGTCTTGGTTTTGACTTAATGCTCTTCAGCTCGGTGAGTTC
TTTCCGTGTTTTGACTTTTTTGTGCACTGTAGCCAATTTCCATAATGAGGCTCAT

DOR24

MDSFLQVQKSTIALLGFDLFSENREMWKRPYRAMNVFSIAAIFPFILAAVLHNWKNVL
LLADAMVALLITILGLFKFSMILYLRRDFKRLIDKFRLLMSNEAEQGEEYAEILNAAN
KQDQRMCTLFRTCFLLAWALNSVLPLVRMGLSYWLAGHAEPELPFPCLFPWNIHIIRN
YVLSFIWSAFASTGVVLPAVSLDTIFCSFTSNLCAFFKIAQYKVVRFKGGSLKESQAT
LNKVFALYQTSLDMCNDLNQCYQPIICAQFFISSLQLCMLGYLFSITFAQTEGVYYAS
FIATIIIQAYIYCYCGENLKTESASFEWAIYDSPWHESLGAGGASTSICRSLLISMMR
AHRGFRITGYFFEANMEAFSSIVRTAMSYITMLRSFS

DOR24nt

GGCACGAGCCTTGTCGACATGGACAGTTTTCTGCAAGTACAGAAGAGCACCATTGCTC
TTCTGGGCTTTGATCTCTTTAGTGAAAATCGAGAAATGTGGAAACGCCCCTATAGAGC
AATGAATGTGTTTAGCATAGCTGCCATTTTTCCCTTTATCCTGGCAGCTGTGCTCCAT
AATTGGAAGAATGTATTGCTGCTGGCCGATGCCATGGTGGCCCTACTAATAACCATTC
TGGGCCTATTCAAGTTTAGCATGATACTTTACTTACGTCGCGATTTCAAGCGACTGAT
TGACAAATTTCGTTTGCTCATGTCGAATGAGGCGGAACAGGGCGAGGAATACGCCGAG
ATTCTCAACGCAGCAAACAAGCAGGATCAACGAATGTGCACTCTGTTTAGGACTTGTT
TCCTCCTCGCCTGGGCCTTGAATAGTGTTCTGCCCCTCGTGAGAATGGGTCTCAGCTA
TTGGTTAGCAGGTCATGCAGAGCCCGAGTTGCCTTTTCCCTGTCTTTTTCCCTGGAAT
ATCCACATCATTCGCAATTATGTTTTGAGCTTCATCTGGAGCGCTTTCGCCTCGACAG
GTGTGGTTTTACCTGCTGTCAGCTTGGATACCATATTCTGTTCCTTCACCAGCAACCT
GTGCGCCTTCTTCAAAATTGCGCAGTACAAGGTGGTTAGATTTAAGGGCGGATCCCTT
AAAGAATCACAGGCCACATTGAACAAAGTCTTTGCCCTGTACCAGACCAGCTTGGATA
TGTGCAACGATCTGAATCAGTGCTACCAACCGATTATCTGCGCCCAGTTCTTCATTTC
ATCTCTGCAACTCTGCATGCTGGGATATCTGTTCTCCATTACTTTTGCCCAGACAGAG
GGCGTGTACTATGCCTCTTTCATAGCCACCATCATTATACAAGCCTATATCTACTGCT
ACTGCGGGGAGAACCTGAAGACGGAGAGTGCCAGCTTCGAGTGGGCCATCTACGACAG
TCCGTGGCACGAGAGTTTGGGTGCTGGTGGAGCCTCTACCTCGATCTGCCGATCCTTG
CTGATCAGCATGATGCGGGCTCATCGGGGATTCCGCATTACGGGATACTTCTTCGAGG
CAAACATGGAGGCCTTCTCATCGATTGTTCGCACGGCTATGTCCTACATCACAATGCT
GAGATCATTCTCCTAAATGTGGTTTGACCACAAGGCTTTGGATTGATTTTTGTGCAAT

TTTTGTTTTATTGCTGAGCATGCGTTGCCGTACGACATTTAACAATCGATCTTACGTA
ATTTACATATGATAATCTCACATATTGTTCGTTAAGCACTAAGTAGAATGTAGAATGT
GAATTGGCTGTAGAAATGCACAGATGAAGCACGAAAAAAAAAAAAAAAAAAAAAA

DOR25

MNDSGYQSNLSLLRVFLDEFRSVLRQESPGLIPRLAFYYVRAFLSLPLYRWINLFIMC
NVMTIFWTMFVALPESKNVIEMGDDLVWISGMALVFTKIFYMHLRCDEIDELISDFEY
YNRELRPHNIDEEVLGWQRLCYVIESGLYINCFCLVNFFSAAIFLQPLLGEGKLPFHS
VYPFQWHRLDLHPYTFWFLYIWQSLTSQHNLMSILMVDMVGISTFLQTALNLKLLCIE
IRKLGDMEVSDKRFHEEFCRVVRFHQHIIKLVGKANRAFNGAFNAQLMASFSLISIST
FETMAAAAVDPKMAAKFVLLMLVAFIQLSLWCVSGTLVYTQSVEVAQAAFDINDWHTK
SPGIQRDISFVILRAQKPLMYVAEPFLPFTLGTYMLVLKNCYRLLALMQESM

DOR25nt

ATGAACGACTCGGGTTATCAATCAAATCTCAGCCTTCTGCGGGTTTTTCTCGACGAGT
TCCGATCGGTTCTGCGGCAGGAAAGTCCCGGTCTCATCCCACGCCTGGCTTTTTACTA
TGTTCGCGCCTTTCTGAGCTTGCCCCTGTACCGATGGATCAACTTGTTCATCATGTGC
AATGTGATGACCATTTTCTGGACCATGTTCGTGGCCCTGCCCGAGTCGAAGAACGTGA
TCGAAATGGGCGACGACTTGGTTTGGATTTCGGGGATGGCACTGGTGTTCACCAAGAT
CTTTTACATGCATTTGCGTTGCGACGAGATCGATGAACTTATTTCGGATTTTGAATAC
TACAACCGGGAGCTGAGACCCCATAATATCGATGAGGAGGTGTTGGGTTGGCAGAGAC
TGTGCTACGTGATAGAATCGGGTCTATATATCAACTGCTTTTGCCTGGTCAACTTCTT
CAGTGCCGCTATTTTCCTGCAACCTCTGTTGGGCGAGGGAAAGCTGCCCTTCCACAGC
GTCTATCCGTTTCAATGGCATCGCTTGGATCTGCATCCCTACACGTTCTGGTTCCTCT
ACATCTGGCAGAGTCTGACCTCGCAGCACAACCTAATGAGCATTCTAATGGTGGATAT
GGTAGGCATTTCCACGTTCCTCCAGACGGCGCTCAATCTCAAGTTGCTTTGCATCGAG
ATAAGGAAACTGGGGGACATGGAGGTCAGTGATAAGAGGTTCCACGAGGAGTTTTGTC
GTGTGGTTCGCTTCCACCAGCACATTATCAAGTTGGTGGGGAAAGCCAATAGAGCTTT
CAATGGCGCCTTCAATGCACAATTAATGGCCAGTTTCTCCCTGATTTCCATATCCACT
TTCGAGACCATGGCTGCAGCGGCTGTGGATCCCAAAATGGCCGCCAAGTTCGTGCTTC
TCATGCTGGTGGCATTCATTCAACTGTCGCTTTGGTGCGTCTCTGGAACTTTGGTTTA
TACTCAGTCAGTGGAGGTGGCTCAGGCTGCTTTTGATATCAACGATTGGCACACCAAA
TCGCCAGGCATCCAGAGGGATATATCCTTTGTGATACTACGAGCCCAGAAACCCCTGA
TGTATGTGGCCGAACCATTTCTGCCCTTCACCCTGGGAACCTATATGCTTGTACTGAA
GAACTGCTATCGTTTGCTGGCCCTGATGCAAGAATCGATGTAG

DOR28

MYSPEEAAELKRRNYRSIREMIRLSYTVGFNLLDPSRCGQVLRIWTIVLSVSSLASLY
GHWQMLARYIHDIPRIGETAGTALQFLTSIAKMWYFLFAHRQIYELLRKARCHELLQK
CELFERMSDLPVIKEIRQQVESTMNRYWASTRRQILIYLYSCICITTNYFINSFVINL
YRYFTKPKGSYDIMLPLPSLYPAWEHKGLEFPYYHIQMYLETCSLYICGMCAVSFDGV
FIVLCLHSVGLMRSLNQMVEQATSELVPPDRRVEYLRCCIYQYQRVANFATEVNNCFR
HITFTQFLLSLFNWGLALFQMSVGLGNNSSITMIRMTMYLVAAGYQIVVYCYNGQRFA
TASEEIANAFYQVRWYGESREFRHLIRMMLMRTNRGFRLDVSWFMQMSLPTLMAVSSG
AEQSRGPAGPAGPAGPPPRVPSYSQFHLIDSQMVRTSGQYFLLLQNVNQK

DOR28nt

ATGTACTCACCGGAAGAGGCGGCCGAACTGAAGAGGCGCAACTATCGCAGCATCAGG
GAGATGATCCGACTCTCCTATACGGTGGGCTTCAACCTGTTGGATCCTTCCCGATGCG
GACAGGTGCTCAGAATCTGGACAATTGTCCTTAGCGTGAGTAGCTTGGCATCGCTTTA
TGGGCACTGGCAAATGTTAGCCAGGTACATTCATGATATTCCACGCATTGGAGAGACC
GCTGGAACTGCCCTGCAGTTCCTAACATCGATAGCAAAGATGTGGTACTTTCTGTTTG
CCCATAGACAGATATACGAATTGCTACGAAAGGCGCGCTGCCATGAATTACTCCAAAA
GTGTGAGCTCTTTGAAAGGATGTCAGATCTACCTGTTATCAAAGAGATTCGCCAGCAG
GTTGAGTCCACGATGAATCGGTACTGGGCCAGCACTCGTCGGCAAATTCTTATCTATT
TGTACAGCTGTATTTGTATTACTACAAACTACTTTATCAACTCCTTCGTAATCAACCT
CTATCGCTATTTCACTAAACCGAAAGGATCCTACGACATAATGTTACCTCTGCCATCT
CTGTATCCCGCCTGGGAGCACAAGGGATTAGAGTTTCCCTACTATCATATACAGATGT
ACCTGGAAACCTGTTCTCTGTATATCTGCGGCATGTGTGCCGTTAGCTTTGATGGAGT
CTTTATTGTCCTGTGCCTTCATAGCGTGGGACTTATGAGGTCACTTAACCAAATGGTG
GAACAAGCCACATCTGAGTTGGTTCCTCCAGATCGCAGGGTTGAATACTTGCGATGCT
GTATTTATCAGTACCAACGAGTGGCGAACTTTGCAACCGAGGTTAACAACTGCTTTCG
GCACATCACTTTCACGCAGTTCCTGCTTAGCCTTTTCAACTGGGGCCTGGCCTTGTTC
CAAATGAGCGTCGGATTGGGCAACAACAGCAGCATCACCATGATCCGGATGACCATGT
ACCTGGTGGCAGCCGGCTATCAGATAGTTGTGTACTGCTACAATGGCCAGCGATTTGC
GACTGCTAGCGAGGAGATTGCCAACGCCTTTTACCAGGTGCGATGGTACGGAGAGTCC
AGGGAGTTCCGCCACCTCATCCGCATGATGCTGATGCGCACGAACCGGGGATTCAGGC
TGGACGTGTCCTGGTTCATGCAAATGTCCTTGCCCACACTCATGGCGGTGAGTAGCGG
AGCAGAGCAGAGCAGGGGTCCTGCAGGTCCTGCAGGTCCTGCAGGTCCACCCCCAAGG
GTCCCCTCCTACAGCCAGTTCCACTTGATTGATTCGCAGATGGTCCGGACAAGTGGAC
AGTACTTCCTGCTGCTGCAGAACGTCAACCAGAAA

DOR30

MAVSTRVATKQEVPESRRAFRNLFNCFYALGMQAPDGSRPTTSSTWQRIYACFSVVMY
VWQLLLVPTFFVISYRYMGGMEITQVLTSAQVAIDAVILPAKIVALAWNLPLLRRAEH
HLAALDARCREQEEFQLILDAVRFCNYLVWFYQICYAIYSSSTFVCAFLLGQPPYALY
LPGLDWQRSQMQFCIQAWIEFLIMNWTCLHQASDDVYAVIYLYVVRIQVQLLARRVEK
LGTDDSGQVEIYPDERRQEEHCAELQRCIVDHQTMLQLLDCISPVISRTIFVQFLITA
AIMGTTMINIFIFANTNTKIASIIYLLAVTLQTAPCCYQATSLMLDNERLALAIFQCQ
WLGQSARFRKMLLYYLHRAQQPITLTAMKLFPINLATYFSIAKFSFSLYTLIKGMNLG
ERFNRTN

DOR30nt

ATGGCGGTGAGCACTCGTGTGGCCACAAAGCAGGAAGTGCCCGAATCCCGGCGAGCGT
TTAGGAATCTCTTCAATTGCTTCTATGCCCTTGGCATGCAGGCACCGGATGGCAGTCG
ACCGACCACGAGCAGCACATGGCAACGCATCTACGCCTGCTTCTCGGTGGTCATGTAC
GTGTGGCAACTGCTGCTGGTGCCCACATTCTTTGTGATCAGCTATCGGTACATGGGCG
GCATGGAGATTACCCAGGTGCTGACCTCCGCCCAGGTGGCCATCGATGCGGTCATTCT
GCCGGCCAAGATTGTGGCACTGGCGTGGAATTTGCCATTGCTGCGCAGAGCAGAGCAT
CATCTGGCCGCCTTGGATGCGCGGTGCAGGGAACAGGAGGAGTTCCAATTGATCCTCG
ATGCGGTGAGGTTTTGCAACTATCTGGTATGGTTCTACCAGATCTGCTATGCCATCTA
CTCCTCGTCGACATTTGTGTGCGCCTTCCTGCTGGGCCAACCGCCATATGCCCTCTAT
TTGCCTGGCCTCGATTGGCAGCGTTCCCAGATGCAGTTCTGCATCCAGGCCTGGATTG
AGTTCCTTATCATGAACTGGACGTGCCTGCACCAAGCTAGCGATGATGTGTACGCCGT
TATCTATCTGTATGTGGTCCGGATTCAAGTGCAATTGCTGGCCAGGCGGGTGGAGAAG
CTGGGCACGGATGATAGTGGCCAGGTGGAGATCTATCCCGATGAGCGGCGGCAGGAGG
AGCATTGCGCGGAACTGCAGCGCTGCATTGTAGATCACCAGACGATGCTGCAGCTGCT
CGACTGCATTAGTCCCGTCATCTCGCGTACCATATTCGTTCAGTTCCTGATCACCGCC
GCCATCATGGGCACCACCATGATCAACATTTTCATTTTCGCCAATACGAACACGAAGA
TCGCATCGATCATTTACCTGCTGGCGGTGACCCTGCAGACGGCTCCATGTTGCTATCA
GGCCACCTCGCTGATGTTGGACAACGAGAGGCTGGCCCTGGCCATCTTCCAGTGCCAG
TGGCTGGGCCAGAGTGCCCGGTTCCGTAAGATGCTGCTCTACTATCTTCATCGCGCCC
AGCAGCCCATCACGCTGACCGCCATGAAGCTGTTTCCCATCAATCTGGCCACGTACTT
CAGTATAGCCAAGTTCTCGTTTTCGCTCTACACGCTCATCAAGGGGATGAATCTCGGC
GAGCGATTCAACAGGACAAAT

DOR31

MIFKYIQEPVLGSLFRSRDSLIYLNRSIDQMGWRLPPRTKPYWWLYYIWTLVVIVLVF
IFIPYGLIMTGIKEFKNFTTTDLFTYVQVPVNTNASIMKGIIVLFMRRRFSRAQKMMD

AMDIRCTKMEEKVQVHRAAALCNRVVVIYHCIYFGYLSMALTGALVIGKTPFCLYNPL
VNPDDHFYLATAIESVTMAGIILANLILDVYPIIYVVVLRIHMELLSERIKTLRTDVE
KGDDQHYAELVECVKDHKLIVEYGNTLRPMISATMFIQLLSVGLLLGLAAVSMQFYNT
VMERVVSGVYTIAILSQTFPFCYVCEQLSSDCESLTNTLFHSKWIGAERRYRTTMLYF
IHNVQQSILFTAGGIFPICLNTNIKMAKFAFSVVTIVNEMDLAEKLRRE

DOR31nt

ATGATTTTTAAGTACATTCAAGAGCCAGTCCTTGGATCCTTATTTCGATCCCGGGATT
CGCTGATCTACTTAAACAGATCCATAGATCAAATGGGATGGAGACTGCCGCCACGAAC
TAAGCCGTACTGGTGGCTCTATTACATTTGGACATTGGTGGTCATAGTACTCGTCTTT
ATCTTTATACCCTATGGACTGATAATGACTGGAATAAAGGAGTTCAAGAACTTCACGA
CCACGGATCTGTTTACGTATGTCCAGGTGCCGGTTAACACCAATGCTTCGATCATGAA
GGGCATTATAGTGTTGTTTATGCGGCGGCGATTTTCAAGGGCTCAGAAGATGATGGAC
GCCATGGACATTCGATGCACCAAGATGGAGGAGAAAGTCCAGGTGCACCGAGCAGCAG
CCTTATGCAATCGTGTTGTTGTGATTTACCATTGCATATACTTCGGCTATCTATCCAT
GGCCTTAACCGGAGCTCTGGTGATTGGGAAGACTCCATTCTGTTTGTACAATCCACTG
GTTAACCCCGACGATCATTTCTATCTGGCCACTGCCATTGAATCGGTCACCATGGCTG
GCATTATTCTGGCCAATCTCATTTTGGACGTATATCCCATCATATATGTGGTCGTTCT
GCGGATCCACATGGAGCTCTTGAGTGAGCGAATCAAGACGCTGCGTACTGATGTGGAA
AAAGGCGACGATCAACATTATGCCGAGCTGGTGGAGTGTGTAAAGGATCACAAGCTAA
TTGTCGAATATGGAAACACTCTGCGTCCCATGATATCCGCCACGATGTTCATCCAACT
ACTATCCGTTGGCTTACTTTTGGGTCTGGCAGCGGTGTCCATGCAGTTCTATAACACC
GTAATGGAGCGTGTTGTCTCCGGGGTCTACACCATAGCCATTCTATCCCAGACCTTTC
CATTTTGCTATGTCTGTGAGCAGCTGAGCAGCGATTGCGAATCCCTGACCAACACACT
GTTCCATTCCAAGTGGATTGGAGCTGAGCGACGATACAGAACCACGATGTTGTACTTC
ATTCACAATGTTCAGCAGTCGATTTTGTTCACTGCGGGCGGAATTTTCCCCATATGTC
TAAACACCAATATAAAGATGGCCAAGTTCGCTTTCTCAGTGGTGACCATTGTAAATGA
GATGGACTTGGCCGAGAAATTGAGAAGGGAG

DOR32

MEPVQYSYEDFARLPTTVFWIMGYDMLGVPKTRSRRILYWIYRFLCLASHGVCVGVMV
FRMVEAKTIDNVSLIMRYATLVTYIINSDTKFATVLQRSAIQSLNSKLAELYPKTTLD
RIYHRVNDHYWTKSFVYLVIIYIGSSIMVVIGPIITSIIAYFTHNVFTYMHCYPYFLY
DPEKDPVWIYISIYALEWLHSTQMVISNIGADIWLLYFQVQINLHFRGIIRSLADHKP
SVKHDQEDRKFIAKIVDKQVHLVSLQNDLNGIFGKSLLLSLLTTAAVICTVAVYTLIQ
GPTLEGFTYVIFIGTSVMQVYLVCYYGQQVLDLSGEVAHAVYNHDFHDASIAYKRYLL
IIIIRAQQPVELNAMGYLSISLDTFKQLMSVSYRVITMLMQMIQ

DOR32nt

ATGGAACCTGTGCAGTACAGCTACGAGGATTTCGCTCGATTGCCCACGACGGTGTTCT
GGATCATGGGCTACGACATGCTGGGCGTTCCGAAGACCCGCTCTCGCAGGATACTATA
CTGGATATATCGTTTCCTCTGTCTCGCCAGCCATGGGGTCTGTGTAGGAGTCATGGTA
TTTCGTATGGTGGAGGCAAAGACCATTGACAATGTTTCGCTGATCATGCGGTATGCCA
CTCTGGTCACCTATATCATCAACTCGGATACGAAATTCGCAACTGTCTTACAAAGGAG
TGCAATTCAAAGTCTAAACTCAAAACTGGCCGAACTATATCCGAAGACCACGCTGGAC
AGGATCTATCACCGGGTGAATGATCACTATTGGACCAAGTCATTTGTATATTTGGTTA
TTATCTACATTGGTTCGTCGATTATGGTTGTTATTGGACCGATTATTACGTCGATTAT
AGCTTACTTCACGCACAACGTTTTCACCTACATGCACTGCTATCCGTACTTTTTGTAT
GATCCTGAGAAGGATCCGGTTTGGATCTACATCAGCATCTATGCTCTGGAATGGTTGC
ACAGCACACAGATGGTCATTTCGAACATTGGCGCGGATATCTGGCTGCTGTACTTTCA
GGTGCAGATAAATCTCCACTTCAGGGGCATTATACGATCACTGGCGGATCACAAGCCC
AGTGTGAAGCACGACCAGGAGGACAGGAAATTCATTGCGAAAATTGTCGACAAGCAGG
TGCACCTGGTCAGTTTGCAAAACGATCTGAATGGTATCTTTGGAAAATCGCTGCTTCT
AAGCCTGCTGACCACCGCAGCGGTTATCTGCACGGTGGCGGTGTACACTCTGATTCAG
GGTCCCACCTTGGAGGGCTTCACCTATGTGATCTTCATCGGGACTTCTGTGATGCAGG
TCTACCTGGTGTGCTATTACGGTCAGCAAGTTCTCGACTTGAGCGGCGAGGTGGCCCA
CGCCGTGTACAATCATGATTTTCACGATGCTTCTATAGCGTACAAGAGGTACCTGCTC
ATAATCATTATCAGGGCGCAGCAGCCCGTGGAACTTAATGCCATGGGCTACCTGTCCA
TTTCGCTGGACACCTTTAAACAGCTGATGAGCGTCTCCTACCGGGTTATAACCATGCT
CATGCAGATGATTCAG

DOR37

protein sequence is incomplete and is in progress
KVDSTRALVNHWRIFRIMGIHPPGKRTFWGRHYTAYSMVWNVTFHICIWVSFSVNLLQ
SNSLETFCESLCVTMPHTLYMLKLINVRRMRGQMISSHWLLRLLDKRLGCDDERQIIM
AGIERAEFIFRTIFRGLACTVVLGIIYISASSEPTLMYPTWIPWNWRDSTSAYLATAM
LHTTALMANATLVLNLSSYPGTYLILVSVHTKALALRVSKLGYGAPLPAVRMQAILVG
YIHDHQIILR*VSGNLISQCKNF*SISGVLTFIERRMYTHFGVPNIFIVIEDYYILFL
NYSLFKSLERSLSMTCFLQFFSTACAQCTICYFLLFGNVGIMRFMNMLFLLVILTTET
LLLCYTAELPCKEGESLLTAVYSCNWLSQSVNFRRLLLLMLARCQIPMILVSGVIVPI
SMKTF

DOR37nt

information on nucleotide sequence is in progress

DOR38

MRLIKISYSALNEVCVWLKLNGSWPLTESSRPWRSQSLLATAYIVWAWYVIASVGITI
SYQTAFLLNNLSDIIITTENCCTTFMGVLNFVRLIHLRLNQRKFRQLIENFSYEIWIP
NSSKNNVAAECRRRMVTFSIMTSLLACLIIMYCVLPLVEIFFGPAFDAQNKPFPYKMI
FPYDAQSSWIRYVMTYIFTSYAGICVVTTLFAEDTILGFFITYTCGQFHLLHQRIAGL
FAGSNAELAESIQLERLKRIVEKHNNIISANSV

DOR38nt

ATGCGTTTGATCAAAATTTCATATTCGGCACTTAATGAGGTGTGCGTTTGGCTGAAAC
TGAATGGTTCTTGGCCATTAACCGAATCATCGAGGCCATGGAGGAGCCAATCCTTATT
GGCCACCGCCTACATCGTGTGGGCGTGGTACGTCATTGCATCTGTGGGCATAACAATC
AGCTATCAGACGGCCTTTTTGCTGAACAACCTTTCGGACATTATTATCACCACGGAAA
ATTGTTGCACCACCTTTATGGGTGTCCTGAACTTTGTCCGACTCATCCATCTTCGCCT
CAATCAGAGGAAATTCCGCCAGCTTATTGAGAACTTTTCCTACGAAATTTGGATACCT
AATTCTTCCAAAAACAATGTTGCCGCCGAGTGTCGCAGACGCATGGTTACCTTCAGCA
TAATGACATCCTTGCTAGCGTGCCTGATCATAATGTATTGTGTCCTGCCGCTGGTGGA
GATCTTCTTTGGACCCGCCTTCGATGCACAGAACAAGCCGTTTCCCTACAAGATGATC
TTTCCGTACGATGCCCAGAGCAGTTGGATCCGATATGTGATGACCTACATCTTCACCT
CCTACGCGGGAATCTGTGTGGTCACCACCTTGTTTGCAGAGGACACCATTCTTGGCTT
CTTCATAACCTACACTTGTGGCCAATTTCATTTGCTACACCAACGAATCGCAGGTTTA
TTTGCGGGTTCCAATGCGGAATTGGCCGAGAGCATTCAGCTGGAGCGACTCAAACGTA
TTGTGGAAAAACACAACAATATTATCAGCGCAAATTCTGTA

DOR44

MKSTFKEERIKDDSKRRDLFVFVRQTMCIAAMYPFGYYVNGSGVLAVLVRFCDLTYEL
FNYFVSVHIAGLYICTIYINYGQGDLDFFVNCLIQTIIYLWTIAMKLYFRRFRPGLLN
TILSNINDEYETRSAVGFSFVTMAGSYRMSKLWIKTYVYCCYIGTIFWLALPIAYRDR
SLPLACWYPFDYTQPGVYEVVFLLQAMGQIQVAASFASSSGLHMVLCVLISGQYDVLF
CSLKNVLASSYVLMGANMTELNQLQAEQSAADVEPGQYAYSVEEETPLQELLKVGSSM
DFSSAFRLSFVRCIQHHRYIVAALKKIESFYSPIWFVKIGEVTFLMCLVAFVSTKSTA
ANSFMRMVSLGQYLLLVLYELFIICYFADIVFQNSQRCGEALWRSPWQRHLKDVRSDY
MFFMLNSRRQFQLTAGKISNLNVDRFRGVGILT

DOR44nt

ATGAAGAGCACATTCAAGGAAGAAAGGATTAAGGACGACTCCAAGCGTCGCGACCTGT
TTGTATTCGTGAGGCAAACCATGTGTATAGCGGCCATGTATCCCTTCGGTTACTACGT
GAATGGATCTGGAGTCCTGGCCGTTCTGGTGCGATTCTGTGACTTGACCTACGAGCTC

TTTAACTACTTCGTTTCGGTACACATAGCTGGCCTGTACATCTGCACCATCTACATCA
ACTATGGGCAAGGCGATTTGGACTTCTTCGTGAACTGTTTGATACAAACCATTATTTA
TCTGTGGACAATAGCGATGAAACTCTACTTTCGGAGGTTCAGACCTGGTTTGTTGAAT
ACCATTCTGTCCAACATCAATGATGAGTACGAGACACGTTCGGCTGTGGGATTCAGTT
TCGTCACAATGGCGGGATCCTATCGGATGTCCAAGCTATGGATCAAAACCTATGTGTA
TTGCTGCTACATAGGCACCATTTTCTGGCTGGCTCTTCCCATTGCCTACCGGGATAGG
AGTCTTCCTCTTGCCTGCTGGTATCCCTTTGACTATACACAACCCGGTGTCTATGAGG
TAGTGTTCCTTCTCCAGGCGATGGGACAGATCCAAGTGGCCGCATCCTTTGCCTCCTC
CAGTGGCCTGCATATGGTGCTTTGTGTGCTGATATCAGGGCAGTACGATGTCCTCTTT
TGCAGTCTCAAGAATGTATTAGCCAGCAGCTATGTCCTTATGGGAGCCAATATGACGG
AACTGAATCAATTGCAGGCTGAGCAATCTGCGGCCGATGTCGAGCCAGGTCAGTATGC
TTACTCCGTGGAGGAGGAGACACCTTTGCAAGAACTTCTAAAAGTTGGGAGCTCAATG
GACTTCTCCTCCGCATTCAGGCTGTCTTTTGTGCGGTGCATTCAGCACCATCGATACA
TAGTGGCGGCACTGAAGAAAATTGAGAGTTTCTACAGTCCCATATGGTTCGTGAAGAT
TGGCGAAGTCACCTTTCTTATGTGCCTGGTAGCCTTCGTCTCCACGAAGAGCACCGCG
GCCAACTCATTCATGCGAATGGTCTCCTTGGGCCAGTACCTGCTCTTAGTTCTCTACG
AGCTGTTCATCATCTGCTACTTCGCGGACATCGTTTTCAGAACAGCCAGCGGTGCGG
TGAAGCCCTCTGGCGAAGTCCTTGGCAGCGACATTTGAAGGATGTTCGCAGTGATTAC
ATGTTCTTTATGCTGAATTCCCGCAGGCAGTTCCAACTTACGGCCGGAAAAATAAGCA
ATCTAAACGTGGATCGTTTCAGAGGGGTGGGTATCCTTACT

DOR46

MAEVRVDSLEFFKSHWTAWRYLGVAHFRVENWKNLYVFYSIVSNLLVTLCYPVHLGIS
LFRNRTITEDILNLTTFATCTACSVKCLLYAYNIKDVLEMERLLRLLDERVVGPEQRS
IYGQVRVQLRNVLYVFIGIYMPCALFAELSFLFKEERGLMYPAWFPFDWLHSTRNYYI
ANAYQIVGISFQLLQNYVSDCFPAVVLCLISSHIKMLYNRFEEVGLDPARDAEKDLEA
CITDHKHILELFRRIEAFISLPMLIQFTVTALNVCIGLAALVFFVSEPMARMYFIFYS
LAMPLQIFPSCFFGTDNEYWFGRLHYAAFSCNWHTQNRSFKRKMMLFVEQSLKKSTAV
AGGMMRIHLDTFFSTLKGAYSLFTIIIRMRK

DOR46nt

ATGGCAGAGGTCAGAGTGGACAGTCTGGAGTTTTTCAAGAGCCATTGGACCGCCTGGC
GGTACTTGGGAGTGGCTCATTTTCGGGTCGAGAACTGGAAGAACCTTTACGTGTTTTA
CAGCATTGTGTCGAATCTTCTCGTGACCCTGTGCTACCCCGTTCACCTGGGAATATCC
CTCTTTCGCAACCGCACCATCACCGAGGACATCCTCAACCTGACCACCTTTGCGACCT
GCACAGCCTGTTCGGTGAAGTGCCTGCTCTACGCCTACAACATCAAGGATGTGCTGGA
GATGGAGCGGCTGTTGAGGCTTTTGGATGAACGCGTCGTGGGTCCGGAGCAACGCAGC
ATCTACGGACAAGTGAGGGTCCAGCTGCGAAATGTGCTATACGTGTTCATCGGCATCT

ACATGCCGTGTGCCCTGTTCGCCGAGCTATCCTTTCTGTTCAAGGAGGAGCGCGGTCT
GATGTATCCCGCCTGGTTTCCCTTCGACTGGCTGCACTCCACCAGGAACTATTACATA
GCGAACGCCTATCAGATAGTGGGCATCTCGTTTCAGCTGCTGCAAAACTATGTTAGCG
ACTGCTTTCCGGCGGTGGTGCTGTGCCTGATCTCATCCCACATCAAAATGTTGTACAA
CAGATTCGAGGAGGTGGGCCTGGATCCAGCCAGAGATGCGGAGAAGGACCTGGAGGCC
TGCATCACCGATCACAAGCATATTCTAGAGTGGGCAGGCGGCTCATTGGTTCGTGTTC
TATTCACTTTCCAACTTTTTTCCAGACTATTCCGACGCATCGAGGCCTTCATTTCCCT
GCCCATGCTAATTCAGTTCACAGTGACCGCCTTGAATGTGTGCATCGGTTTAGCAGCC
CTGGTGTTTTTCGTCAGCGAGCCCATGGCACGGATGTACTTCATCTTCTACTCCCTGG
CCATGCCGCTGCAGATCTTTCCGTCCTGCTTTTTCGGCACCGACAACGAGTACTGGTT
CGGACGCCTCCACTACGCGGCCTTCAGTTGCAATTGGCACACACAGAACAGGAGCTTT
AAGCGGAAAATGATGCTGTTCGTTGAGCAATCGTTGAAGAAGAGCACCGCTGTGGCTG
GCGGAATGATGCGTATCCACCTGGACACGTTCTTTTCCACCCTAAAGGGGGCCTACTC
CCTCTTTACCATCATTATTCGGATGAGAAAG

DOR48

MERHYFMVPKFALSLIGFYPEQKRTVLVKLWSFFNFFILTYGCYAEAYYGIHYIPINI
ATALDALCPVASSILSLVKMVAIWWYQDELRSLIERRFYTLATQLTFLLLCCGFCTST
SYSVRHLIDNILRRTHGKDWIYETPFKMMFPDLLLRLPLYPITYILVHWHGYITVVCF
VGADGFFLGFCLYFTVLLLCLQDDVCDLLEVENIEKSPSEAEEARIVREMEKLVDRHN
EVAELTERLSGVMVEITLAHFVTSSLIIGTSVVDILLFSGLGIIVYVVYTCAVGVEIF
LYCLGGSHIMEACSNLARSTFSSHWYGHSVRVQKMTLLMVARAQRVLTIKIPFFSPSL
ETLTSILRFTGSLIALAKSVI

DOR48nt

ATGGAGCGCCATTATTTCATGGTGCCAAAGTTTGCATTATCGCTGATTGGTTTTTATC
CCGAACAGAAGCGAACGGTTTTGGTGAAACTTTGGAGTTTCTTCAACTTTTTCATCCT
CACCTACGGCTGTTATGCAGAGGCTTACTATGGCATACACTATATACCGATTAACATA
GCCACTGCATTGGATGCCCTTTGTCCTGTGGCCTCCAGCATTTTGTCGCTGGTGAAAA
TGGTCGCCATTTGGTGGTATCAAGATGAATTAAGGAGTTTGATAGAGCGGGTAAGATT
TTTAACAGAGCAACAGAAGTCCAAGAGGAAACTGGGCTATAAGAAGAGGTTCTATACA
CTGGCAACGCAACTAACATTCCTGCTACTATGCTGTGGATTTTGCACCAGTACTTCCT
ATTCCGTCAGACATTTGATTGATAATATCCTGAGACGCACCCATGGCAAGGACTGGAT
CTACGAGACTCCGTTCAAGATGATGTAAGGAAAGGGAAGAATGGTTTATATATACTTT
TGGAACGAAATAATGATGTGATCTAAACAAGATGCACTTTTTTTTAGGTTCCCCGATC
TTCTCCTGCGTTTGCCACTCTATCCCATCACCTATATACTCGTGCATTGGCATGGCTA
CATTACTGTGGTTTGTTTTGTCGGCGCGGATGGTTTCTTCCTGGGGTTCTGTTTGTAC
TTCACTGTTTTGCTGCTCTGTCTGCAGGACGATGTTTGTGATTTACTAGAGGTTGAAA

ACATCGAGAAGAGTCCCTCCGAAGCGGAGGAAGCTCGCATAGTTCGGGAAATGGAAAA
ACTGGTGGACCGGCATAACGAGGTGGCCGAGCTGACAGAAAGATTGTCGGGTGTTATG
GTGGAAATAACACTGGCCCACTTTGTTACTTCGAGTTTGATAATCGGAACCAGCGTGG
TGGATATTTTATTAGTGGGTATTTACATTTGATTAGATCCTTTCGATATATGTTCTTA
AATTCTAGTTTTCCGGCCTGGGAATCATTGTGTATGTGGTCTACACTTGTGCCGTAGG
TGTGGAAATATTTCTATACTGTTTAGGAGGATCTCATATTATGGAAGCGGTATATTCA
TAAGAAACTACTATAAAGTTACTTTTAAATTCATTGCATTTCTTAGTGTTCCAATCTA
GCGCGCTCCACATTTTCCAGCCACTGGTATGGCCACAGTGTTCGGGTCCAAAAGATGA
CCCTTTTGATGGTAGCTCGTGCTCAACGAGTTCTCACAATTAAAATTCCTTTCTTTTC
CCCATCATTAGAGACTCTAACTTCGGTAAGCTTATGCGAAAATGTTATGGTACACACA
AGTCTACATTTCTATGAGGTCTTGTAGATTTTGCGCTTCACTGGATCTCTGATTGCCC
TGGCAAAGTCGGTTATA

DOR53

MLSKFFPHIKEKPLSERVKSRDAFIYLDRVMWSFGWTEPENKRWILPYKLWLAFVNIV
MLILLPISISIEYLHRFKTFSAGEFLSSLEIGVNMYGSSFKCAFTLIGFKKRQEAKVL
LDQLDKRCLSDKERSTVHRYVAMGNFFDILYHIFYSTFVVMNFPYFLLERRHAWRMYF
PYIDSDEQFYISSIAECFLMTEAIYMDLCTDVCPLISMLMARCHISLLKQRLRNLRSK
PGRTEDEYLEELTECIRDHRLLLDYVDALRPVFSGTIFVQFLLIGTVLGLSMINLMFF
STFWTGVATCLFMFDVSMETFPFCYLCNMIIDDCQEMSNCLFQSDWTSADRRYKSTLV
YFLHNLQQPITLTAGGVFPISMQTNLAMVKLAFSVVTVIKQFNLAERFQ

DOR53nt

TCAAACAAAGCCACGGACAAGATGTTAAGCAAGTTTTTCCCCACATAAAAGAAAAGC
CATTGAGCGAGCGGGTTAAGTCCCGAGATGCCTTCATTTACTTGGATCGGGTGATGTG
GTCCTTTGGCTGGACAGAGCCTGAAAACAAAAGGTGGATCCTTCCTTATAAACTGTGG
TTAGCGTTCGTGAACATAGTAATGCTCATCCTTCTGCCGATCTCGATAAGCATCGAGT
ACCTCCACCGATTTAAAACCTTCTCGGCGGGGGAGTTCCTTAGTTCCCTCGAGATTGG
AGTCAACATGTACGGAAGCTCTTTTAAGTGCGCCTTCACCTTGATTGGATTCAAGAAA
AGACAGGAAGCTAAGGTTTTACTGGATCAGCTGGACAAGAGATGCCTTAGCGATAAGG
AGAGGTCCACTGTTCATCGCTATGTCGCCATGGGAAACTTTTTCGATATTTTGTATCA
CATTTTTTACTCCACCTTCGTGGTAATGAACTTCCCGTATTTTCTGCTTGAGAGACGC
CATGCTTGGCGCATGTACTTTCCATATATCGATTCCGACGAACAGTTTTACATCTCCA
GCATCGCCGAGTGTTTTCTGATGACGGAGGCCATCTACATGGATCTCTGTACGGACGT
GTGTCCCTTGATCTCCATGCTTATGGCTCGATGCCACATCAGCCTCCTGAAACAGCGA
CTGAGAAATCTCCGATCGAAGCCAGGAAGGACCGAAGATGAGTACTTGGAGGAGCTCA
CCGAGTGCATTCGGGATCATCGATTGCTATTGGACTATGTTGACGCATTGCGACCCGT
CTTTTCGGGAACCATTTTTGTGCAGTTCCTCCTGATCGGTACTGTACTGGGTCTCTCA

ATGATAAATCTAATGTTCTTCTCGACATTTTGGACTGGTGTCGCCACTTGCCTTTTTA
TGTTCGACGTGTCCATGGAGACGTTCCCCTTTTGCTATTTGTGCAACATGATTATCGA
TGACTGCCAGGAAATGTCCAATTGCCTCTTTCAATCGGACTGGACCTCTGCCGATCGT
CGCTACAAATCCACTTTGGTATACTTTCTTCACAATCTTCAGCAACCCATTACTCTCA
CGGCTGGTGGAGTGTTTCCTATTTCCATGCAAACAAATTTGGCTATGGTGAAGCTGGC
ATTTTCTGTGGTTACGGTAATTAAGCAATTTAACTTGGCCGAAAGGTTTCAATAAGTT
GAGAGGGACGAGCTCTGCTACTATTATATTATATATTATATTATATTATATATATATT
ATTTTATATTATATATTGCTGTACCCTAATAAATATTTAGTAATAAAAAAAAAAAAAA
AAAA

DOR56

MDPVEMPIFGSTLKLMKFWSYLFVHNWRRYVAMTPYIIINCTQYVDIYLSTESLDFII
RNVYLAVLFTNTVVRGVLLCVQRFSYERFINILKSFYIELLVSTERLSQKCILHKWAV
LPYGMYLPTIDEYKYASPYYEIFFVIQAIMAPMGCCMYIPYTNMVVTFTLFAILMCRV
LQHKLRSLEKLKNEQVRGEIAQTIAQTVIVIAYMVMIFANSVVLYYVANELYFQSFDI
AIAAYESNWMDFDVDTQKTLKFLIMRSQKPLASLVGGTYPMNLKMLQSLLNAIYSFFT
LLRRVYG

DOR56nt

ATGGATCCGGTGGAGATGCCCATTTTTGGTAGCACTCTGAAGCTAATGAAGTTCTGGT
CATATCTGTTTGTTCACAACTGGCGCCGCTATGTCGCAATGACTCCGTACATCATTAT
CAACTGTACTCAGTATGTGGATATATATCTGAGCACCGAATCCTTGGACTTTATCATC
AGAAATGTATACCTGGCTGTATTGTTTACCAACACGGTGGTCAGAGGTGTATTGTTAT
GCGTACAGCGGTTTAGCTACGAGCGTTTCATTAATATTTTGAAAAGCTTTTACATTGA
GTTGTTGGTGAGTACCGAAAGATTATCTCAAAAATGCATATTGCATAAATGGGCAGTT
CTGCCATATGGCATGTATTTGCCCACTATTGATGAATACAAATACGCATCACCTTACT
ACGAGATTTTCTTTGTGATTCAAGCCATTATGGCTCCAATGGGGTGTTGCATGTACAT
ACCATACACAAACATGGTAGTGACATTTACCCTTTTCGCCATTCTCATGTGTCGAGTG
TTGCAACATAAGTTGAGAAGCCTAGAAAAGCTGAAAAATGAACAAGTACGTGGTGAAA
TCGCTCAAACAATTGCTCAGACCGTCATAGTCATCGCATACATGGTAATGATATTTGC
CAACAGTGTAGTCCTTTACTACGTGGCCAATGAGCTATACTTTCAAAGCTTTGATATT
GCCATTGCTGCCTATGAGAGCAATTGGATGGACTTTGATGTGGACACACAAAAGACTT
TGAAGTTCCTCATCATGCGCTCGCAAAAGCCCTTGGCGAGTCTGGTGGGTGGCACATA
TCCCATGAACTTGAAAATGCTTCAGTCACTACTAAATGCCATTTACTCCTTCTTCACC
CTTCTGCGTCGCGTTTACGGC

DOR58

MDASYFAVQRRALEIVGFDPSTPQLSLKHPIWAGILILSLISHNWPMVVYALQDLSDL
TRLTDNFAVFMQGSQSTFKFLVMMAKRRRIGSLIHRLHKLNQAASATPNHLEKIEREN
QLDRYVARSFRNAAYGVICASAIAPMLLGLWGYVETGVFTPTTPMEFNFWLDERKPHF
YWPIYVWGVLGVAAAAWLAIATDTLFSWLTHNVVIQFQLLELVLEEKDLNGGDSRLTG
FVSRHRIALDLAKELSSIFGEIVFVKYMLSYLQLCMLAFRFSRSGWSAQVPFRATFLV
AIIIQLSSYCYGGEYIKQQSLAIAQAVYGQINWPEMTPKKRRLWQMVIMRAQRPAKIF
GFMFVVDLPLLLWVIRTAGSFLAMLRTFER

DOR58nt

ATGGACGCCAGCTACTTTGCCGTCCAGAGAAGAGCTCTGGAAATAGTTGGATTCGATC
CCAGTACTCCGCAACTGAGTCTGAAACATCCCATCTGGGCCGGGATTCTCATCCTGTC
CTTGATCTCTCACAACTGGCCCATGGTAGTCTATGCCCTGCAGGATCTCTCCGACTTG
ACCCGTCTGACGGACAACTTTGCGGTGTTTATGCAAGGATCACAGAGCACCTTCAAGT
TCCTGGTCATGATGGCGAAACGAAGGCGCATTGGATCGTTGATTCACCGTTTGCATAA
GCTAAACCAGGCGGCCAGTGCCACGCCCAATCACCTGGAGAAGATCGAGAGGGAAAAC
CAACTGGATAGGTATGTCGCCAGGTCCTTTAGAAATGCCGCCTACGGAGTGATTTGTG
CCTCGGCCATAGCGCCCATGTTGCTTGGCCTGTGGGGATATGTGGAGACGGGTGTATT
TACCCCCACCACACCCATGGAGTTCAACTTCTGGCTGGACGAGCGAAAGCCTCACTTT
TATTGGCCCATCTACGTTTGGGGCGTACTGGGCGTGGCAGCTGCCGCCTGGTTGGCCA
TTGCAACGGACACCCTGTTCTCCTGGCTGACTCACAATGTGGTGATTCAGTTCCAACT
ACTGGAGCTTGTTCTCGAAGAGAAGGATCTGAATGGCGGAGACTCTCGCCTGACCGGG
TTTGTTAGTCGTCATCGTATAGCTCTGGATTTGGCCAAGGAACTAAGTTCGATTTTCG
GGGAGATCGTCTTTGTGAAATACATGCTCAGTTACCTGCAACTCTGCATGTTGGCCTT
TCGCTTCAGCCGCAGTGGCTGGAGTGCCCAGGTGCCATTTAGAGCCACCTTCCTAGTG
GCCATCATCATCCAACTGAGTTCGTATTGCTATGGAGGCGAGTATATAAAGCAGCAAA
GTTTGGCCATCGCACAAGCCGTTTATGGTCAAATCAATTGGCCAGAAATGACGCCAAA
GAAAAGAAGACTCTGGCAAATGGTGATCATGAGGGCGCAGCGACCGGCTAAGATTTTT
GGATTCATGTTCGTTGTGGACTTGCCACTGCTGCTTTGGGTCATCAGAACTGCGGGCT
CATTTCTGGCCATGCTTAGGACTTTCGAGCGT

DOR59

MHEADNREMELLVATQAYTRTITLLIWIPSVIAGLMAYSDCIYRSLFLPKSVFNVPAV
RRGEEHPILLFQLFPFGELCDNFVVGYLGPWYALGLGITAIPLWHTFITCLMKYVNLK
LQILNKRVEEMDITRLNSKLVIGRLTASELTFWQMQLFKEFVKEQLRIRKFVQELQYL
ICVPVMADFIIFSVLICFLFFALTVGHDELSLAYFSCGWYNFEMPLQKMLVFMMMHAQ
RPMKMRALLVDLNLRTFIDIGRGAYSYFNLLRSSHLY

DOR59nt

ATGCACGAAGCAGATAATCGGGAGATGGAACTTTTGGTCGCCACTCAGGCTTATACAC
GAACCATTACCCTGTTGATCTGGATACCATCGGTTATTGCTGGCCTAATGGCCTATTC
AGACTGCATCTACAGGAGTCTGTTTCTGCCGAAATCGGTTTTCAATGTGCCAGCTGTG
CGACGTGGTGAGGAGCATCCCATTCTGCTATTTCAGCTGTTTCCCTTCGGAGAACTTT
GCGATAACTTCGTTGTTGGATACTTGGGACCTTGGTATGCTCTGGGCCTGGGAATCAC
GGCTATCCCATTGTGGCACACCTTTATCACTTGCCTCATGAAGTACGTAAATCTCAAG
CTGCAAATACTCAACAAGCGAGTGGAGGAGATGGATATTACCCGACTTAATTCCAAAT
TGGTAATTGGTCGCCTAACTGCCAGTGAGTTAACCTTCTGGCAAATGCAACTCTTCAA
GGAATTTGTAAAGGAACAGCTGAGGATTCGAAAATTTGTCCAGGAACTACAGTATCTG
ATTTGCGTGCCTGTGATGGCAGATTTCATTATCTTCTCGGTTCTCATTTGCTTTCTCT
TTTTTGCCTTGACAGTTGGCCACGATGAACTGAGCCTTGCTTACTTTTCTTGCGGATG
GTACAACTTCGAAATGCCTTTGCAGAAAATGCTGGTTTTTATGATGATGCATGCCCAA
AGGCCGATGAAGATGCGCGCCCTGCTGGTCGATTTGAATCTGAGGACCTTCATAGACA
TTGGCCGTGGAGCCTACAGCTACTTCAATTTGCTGCGTAGCTCCCACTTGTAT

DOR61

MGHKDDMDSTDSTALSLKHISSLIFVISAQYPLISYVAYNRNDMEKVTACLSVVFTNM
LTVIKISTFLANRKDFWEMIHRFRKMHEQCKYREGLDYVAEANKLASFLGRAYCVSCG
LTGLYFMLGPIVKIGVCRWHGTTCDKELPMPMKFPFNDLESPGYEVCFLYTVLVTVVV
VAYASAVDGLFISFAINLRAHFQTLQRQIENWEFPSSEPDTQIRLKSIVEYHVLLLSL
SRKLRSIYTPTVMGQFVITSLQVGVIIYQLVTNMDSVMDLLLYASFFGSIMLQLFIYC
YGGEIIKAESLQVDTAVRLSNWHLASPKTRTSLSLIILQSQKEVLIRAGFFVASLANF
PYRLITLIKSIDSIC

DOR61nt

information on nucleotide sequence is in progress

DOR62

MEKQEDFKLNTHSAVYYHWRVWELTGLMRPPGVSSLLYVVYSITVNLVVTVLFPLSLL
ARLLFTTNMAGLCENLTITITDIVANLKFANVYMVRKQLHEIRSLLRLMDARARLVGD
PEEISALRKEVNIAQGTFRTFASIFVFGTTLSCVRVVVRPDRELLYPAWFGVDWMHST
RNYVLINIYQLFGLIVQAIQNCASDSYPPAFLCLLTGHMRALELRVRRIGCRTEKSNK
GQTYEAWREEVYQELIECIRDLARVHRLREIIQRVLSVPCMAQFVCSAAVQCTVAMHF
LYVADDHDHTAMIISIVFFSAVTLEVFVICYFGDRMRTQSEALCDAFYDCNWIEQLPK
FKRELLFTLARTQRPSLIYAGNYIALSLETFEQVMRFTYSVFTLLLRAK

DOR62nt

ATGGAGAAGCAAGAGGATTTCAAACTGAACACCCACAGTGCTGTGTACTACCACTGGC
GCGTTTGGGAGCTCACTGGCCTGATGCGTCCTCCGGGCGTTTCAAGCCTGCTTTACGT
GGTATACTCCATTACGGTCAACTTGGTGGTCACCGTGCTGTTTCCCTTGAGCTTGCTG
GCCAGGCTGCTGTTCACCACCAACATGGCCGGATTGTGCGAGAACCTGACCATAACTA
TTACCGATATTGTGGCCAATTTGAAGTTTGCGAATGTGTACATGGTGAGGAAGCAGCT
CCATGAGATTCGCTCTCTCCTAAGGCTCATGGACGCTAGAGCCCGGCTGGTGGGCGAT
CCCGAGGAGATTTCTGCCTTGAGGAAGGAAGTGAATATCGCACAGGGCACTTTCCGCA
CCTTTGCCAGTATTTTCGTATTTGGCACTACTTTGAGTTGCGTCCGCGTGGTCGTTCG
CCCGGATCGAGAGCTCCTGTATCCGGCCTGGTTCGGCGTTGACTGGATGCACTCCACC
AGAAACTATGTGCTCATCAATATCTACCAGCTCTTCGGCTTGATAGTGCAGGCTATAC
AGAACTGCGCTAGTGACTCCTATCCGCCTGCGTTTCTCTGCCTGCTCACGGGTCATAT
GCGTGCTTTGGAGCTGAGGGTGCGGCGGATTGGCTGCAGGACGGAAAAGTCCAATAAA
GGGCAGACATATGAAGCCTGGCGGGAGGAGGTGTACCAGGAACTCATCGAGTGCATCC
GCGATCTGGCGCGGGTCCATCGGCTGAGGGAGATCATTCAGCGGGTCCTTTCAGTGCC
CTGCATGGCCCAGTTCGTCTGCTCCGCCGCCGTCCAGTGTACCGTCGCCATGCACTTC
CTGTACGTAGCGGATGACCACGACCACACCGCCATGATCATCTCGATTGTATTTTTCT
CGGCCGTCACCTTGGAGGTGTTTGTAATCTGCTATTTTGGGGACAGGATGCGGACACA
GAGCGAGGCGCTGTGCGATGCCTTCTACGATTGCAACTGGATAGAACAGCTGCCCAAG
TTCAAGCGCGAACTGCTCTTCACCCTGGCCAGGACGCAGCGGCCTTCTCTTATTTACG
CAGGCAACTACATCGCACTCTCGCTGGAGACCTTCGAGCAGGTCATGAGGTTCACATA
CTCTGTTTTCACACTCTTGCTGAGGGCCAAGTAAGAACTTTATAATCTCTTTTTGGGG
AGAAAAATTTTAAAGCACAATAGCAGAAAAATATATCAGATAATATAACAAAAAAAAA
AAAAAAAAA

DOR64

MKLSETLKIDYFRVQLNAWRICGALDLSEGRYWSWSMLLCILVYLPTPMLLRGVYSFE
DPVENNFSLSLTVTSLSNLMKFCMYVAQLTKMVEVQSLIGQLDARVSGESQSERHRNM
TEHLLRMSKLFQITYAVVFIIAAVPFVFETELSLPMPMWFPFDWKNSMVAYIGALVFQ
EIGYVFQIMQCFAADSFPPLVLYLISEQCQLLILRISEIGYGYKTLEENEQDLVNCIR
DQNALYRLLDVTKSLVSYPMMVQFMVIGINIAITLFVLIFYVETLYDRIYYLCFLLGI
TVQTYPLCYYGTMVQESFAELHYAVFCSNWVDQSASYRGHMLILAERTKRMQLLLAGN
LVPIHLSTYVACWKGAYSFFTLMADRDGLGS

DOR64nt

GGCACGAGCCAAGAATTCAAAATGAAACTCAGCGAAACCCTAAAAATCGACTATTTTC
GAGTCCAGTTGAATGCCTGGCGAATTTGTGGTGCCTTGGATCTCAGCGAGGGTAGGTA

CTGGAGTTGGTCGATGCTATTGTGCATCTTGGTGTACCTGCCGACACCCATGCTACTG
AGAGGAGTATACAGTTTCGAGGATCCGGTGGAAAATAATTTCAGCTTGAGCCTGACGG
TCACATCGCTGTCCAATCTCATGAAGTTCTGCATGTACGTGGCCCAACTAACAAAGAT
GGTCGAGGTCCAGAGTCTTATTGGTCAGCTGGATGCCCGGGTTTCTGGCGAGAGCCAG
TCTGAGCGTCATAGAAATATGACCGAGCACCTGCTAAGGATGTCCAAGCTGTTCCAGA
TCACCTACGCTGTAGTCTTCATCATTGCTGCAGTTCCCTTCGTTTTCGAAACTGAGCT
AAGCTTACCCATGCCCATGTGGTTTCCCTTCGACTGGAAGAACTCGATGGTGGCCTAC
ATCGGAGCTCTGGTTTTCCAGGAGATTGGCTATGTCTTTCAAATTATGCAATGCTTTG
CAGCTGACTCGTTTCCCCCGCTCGTACTGTACCTGATCTCCGAGCAATGTCAATTGCT
GATCCTGAGAATCTCTGAAATCGGATATGGTTACAAGACTCTGGAGGAGAACGAACAG
GATCTGGTCAACTGCATCAGGGATCAAAACGCGCTGTATAGATTACTCGATGTGACCA
AGAGTCTCGTTTCGTATCCCATGATGGTGCAGTTTATGGTTATTGGCATCAACATCGC
CATCACCCTATTTGTCCTGATATTTTACGTGGAGACCTTGTACGATCGCATCTATTAT
CTTTGCTTTCTCTTGGGCATCACCGTGCAGACATATCCATTGTGCTACTATGGAACCA
TGGTGCAGGAGAGTTTTGCTGAGCTTCACTATGCGGTATTCTGCAGCAACTGGGTGGA
TCAAAGTGCCAGCTATCGTGGGCACATGCTCATCCTGGCGGAGCGCACTAAGCGGATG
CAGCTTCTCCTCGCCGGCAACCTGGTGCCCATCCACCTGAGCACCTACGTGGCCTGTT
GGAAGGGAGCCTACTCCTTCTTCACCCTGATGGCCGATCGAGATGGCCTGGGTTCTTA
GTAGCCCAGTCATTTCACTCACATTCTACATCAAGTAGTACTACCACTGAACACGAAC
ACGAATATTTCAAAAGTAAACACATAATATTCACAATAGTGTATCACTTTAATAAAAT
TTTTGGTTACCATGAAAAAAAAAAAAAAAAAA

<u>DOR67</u>

MLSQFFPHIKEKPLSERVKSRDAFVYLDRVMWSFGWTVPENKRWDLHYKLWSTFVTLV
IFILLPISVSVEYIQRFKTFSAGEFLSSIQIGVNMYGSSFKSYLTMMGYKKRQEAKMS
LDELDKRCVCDEERTIVHRHVALGNFCYIFYHIAYTSFLISNFLSFIMKRIHAWRMYF
PYVDPEKQFYISSIAEVILRGWAVFMDLCTDVCPLISMVIARCHITLLKQRLRNLRSE
PGRTEDEYLKELADCVRDHRLILDYVDALRSVFSGTIFVQFLLIGIVLGLSMINIMFF
STLSTGVAVVLFMSCVSMQTFPFCYLCNMIMDDCQEMADSLFQSDWTSADRRYKSTLV
YFLHNLQQPIILTAGGVFPISMQTNLNMVKLAFTVVTIVKQFNLAEKFQ

<u>DOR67nt</u>

GGCACGAGGAAATGTTAAGCCAGTTCTTTCCCCACATTAAAGAAAAGCCATTGAGCGA
GCGGGTTAAGTCCCGAGATGCCTTCGTTTACTTAGATCGGGTGATGTGGTCCTTTGGC
TGGACAGTGCCTGAAAACAAAAGGTGGGATCTACATTACAAACTGTGGTCAACTTTCG
TGACATTGGTGATATTTATCCTTCTGCCGATATCGGTAAGCGTTGAGTATATTCAGCG
GTTCAAGACCTTCTCGGCGGGTGAGTTTCTTAGCTCAATCCAGATTGGCGTTAACATG
TACGGAAGCAGCTTTAAAAGTTATTTGACCATGATGGGATATAAGAAGAGACAGGAGG

CTAAGATGTCACTGGATGAGCTGGACAAGAGATGCGTTTGTGATGAGGAGAGGACCAT
TGTACATCGACATGTCGCCCTGGGAAACTTTTGCTATATTTTCTATCACATTGCGTAC
ACTAGCTTTTTGATTTCAAACTTTTTGTCATTTATAATGAAGAGAATCCATGCCTGGC
GCATGTACTTTCCCTACGTCGACCCCGAAAAGCAATTTTACATCTCTAGCATCGCCGA
AGTCATTCTTAGGGGGTGGGCCGTCTTCATGGATCTCTGCACGGATGTGTGTCCTTTG
ATCTCCATGGTAATAGCACGATGCCACATCACCCTTCTGAAACAGCGCCTGCGAAATC
TACGATCGGAACCAGGAAGGACGGAAGATGAGTACTTGAAGGAGCTCGCCGACTGCGT
TCGAGATCACCGCTTGATATTGGACTATGTCGACGCATTGCGATCCGTCTTTTCGGGG
ACAATTTTTGTGCAGTTCCTCTTGATCGGTATTGTACTGGGTCTGTCAATGATAAATA
TAATGTTTTTCTCAACACTTTCGACTGGTGTCGCCGTTGTCCTTTTTATGTCCTGCGT
ATCTATGCAGACGTTCCCCTTTTGCTATTTGTGTAACATGATTATGGATGACTGCCAA
GAGATGGCCGACTCCCTTTTTCAATCGGACTGGACATCTGCCGATCGTCGCTACAAAT
CCACTTTGGTATACTTTCTTCACAATCTTCAGCAGCCCATTATTCTTACGGCTGGTGG
AGTCTTTCCTATTTCCATGCAAACAAATTTAAATATGGTGAAGCTGGCCTTTACTGTG
GTTACAATAGTGAAACAATTTAACTTGGCAGAAAAGTTTCAATAAGTTAAGATATGCA
AGCTCTGCTATTATAAACCTACACTCGAGAAAATATTTCTTCACATTAATAAACCTTC
AGTACTTACTGCTTGTGGCGCCCCCGGAAAAAAAAAAAAAAAAAA

DOR68

MSKLIEVFLGNLWTQRFTFARMGLDLQPDKKGNVLRSPLLYCIMCLTTSFELCTVCAF
MVQNRNQIVLCSEALMHGLQMVSSLLKMAIFLAKSHDLVDLIQQIQSPFTEEDLVGTE
WRSQNQRGQLMAAIYFMMCAGTSVSFLLMPVALTMLKYHSTGEFAPVSSFRVLLPYDV
TQPHVYAMDCCLMVFVLSFFCCSTTGVDTLYGWCALGVSLQYRRLGQQLKRIPSCFNP
SRSDFGLSGIFVEHARLLKIVQHFNYSFMEIAFVEVVIICGLYCSVICQYIMPHTNQN
FAFLGFFSLVVTTQLCIYLFGAEQVRLEAERFSRLLYEVIPWQNLPPKHRKLFLFPIE
RAQRETVLGAYFFELGRPLLVWVSIFLFIVLLF

DOR68nt

ATGTCAAAGCTAATCGAGGTGTTTCTGGGTAATCTGTGGACCCAGCGTTTTACCTTCG
CCCGAATGGGTTTGGATTTGCAGCCCGATAAAAAGGGCAATGTTTTGCGATCTCCGCT
TCTTTATTGTATTATGTGTCTGACAACAAGCTTTGAGCTCTGCACCGTGTGCGCCTTT
ATGGTCCAAAATCGCAACCAAATCGTGCTTTGTTCCGAGGCCCTGATGCACGGACTAC
AGATGGTCTCCTCGCTACTGAAGATGGCTATATTCTTGGCCAAATCTCACGACCTGGT
GGACCTAATTCAACAGATTCAGTCGCCTTTTACAGAGGAGGATCTTGTAGGTACAGAG
TGGAGATCCCAAAATCAAAGGGGACAACTAATGGCTGCCATTTACTTTATGATGTGTG
CCGGTACGAGTGTGTCATTTCTGTTGATGCCAGTGGCTTTGACCATGCTTAAGTACCA
TTCCACTGGGGAATTCGCGCCTGTCAGCTCGTTCCGGGTTCTGCTTCCATACGATGTG
ACACAACCGCATGTTTATGCCATGGACTGCTGCTTGATGGTATTTGTGTTAAGTTTTT

TTTGCTGCTCCACCACCGGAGTGGATACCTTATATGGATGGTGTGCTTTAGGCGTGAG
TTTACAATACCGTCGCCTCGGTCAACAACTTAAAAGGATACCCTCCTGTTTCAATCCA
TCTCGGTCTGACTTTGGATTAAGTGGGATTTTTGTGGAGCATGCTCGTCTGCTTAAAA
TAGTCCAACATTTTAATTATAGTTTTATGGAGATCGCATTTGTGGAGGTTGTTATAAT
CTGTGGACTCTATTGCTCAGTAATTTGTCAGTATATAATGCCACACACCAACCAAAAC
TTCGCCTTTCTGGGTTTCTTTTCATTGGTAGTTACCACACAGCTGTGCATCTATCTTT
TCGGTGCCGAACAGGTCCGTTTGGAGGCTGAGCGATTTTCCCGGCTGCTATACGAAGT
AATTCCTTGGCAAAACCTTCCTCCTAAACACCGGAAACTTTTCCTTTTTCCAATTGAG
CGCGCCCAACGAGAAACTGTTCTCGGTGCTTATTTCTTCGAACTAGGCAGACCTCTTC
TTGTTTGGGTAAGCATATTCCTTTTTATTGTATTATTATTT

DOR71

MVIIDSLSFYRPFWICMRLLVPTFFKDSSRPVQLYVVLLHILVTLWFPLHLLLHLLLL
PSTAEFFKNLTMSLTCVACSLKHVAHLYHLPQIVEIESLIEQLDTFIASEQEHRYYRD
HVHCHARRFTRCLYISFGMIYALFLFGVFVQVISGNWELLYPAYFPFDLESNRFLGAV
ALGYQVFSMLVEGFQGLGNDTYTPLTLCLLAGHVHLWSIRMGQLGYFDDETVVNHQRL
LDYIEQHKLLVRFHNLVSRTISEVQLVQLGGCGATLCIIVSYMLFFVGDTISLVYYLV
FFGVVCVQLFPSCYFASEVAEELERLPYAIFSSRWYDQSRDHRFDLLIFTQLTLGNRG
WIIKAGGLIELNLNAFFATLKMAYSLFAVVHRETGNPLQREH

DOR71nt

ATGGTCATTATCGACAGTCTTAGTTTTTATCGTCCATTCTGGATCTGCATGCGATTGC
TGGTACCGACTTTCTTCAAGGATTCCTCACGTCCTGTCCAGCTGTACGTGGTGTTGCT
GCACATCCTGGTCACCTTGTGGTTTCCACTGCATCTGCTGCTGCATCTTCTGCTACTT
CCATCTACCGCTGAGTTCTTTAAGAACCTGACCATGTCTCTGACTTGTGTGGCCTGCA
GTCTGAAGCATGTGGCCCACTTGTATCACTTGCCGCAGATTGTGGAAATCGAATCACT
GATCGAGCAATTAGACACATTTATTGCCAGCGAACAGGAGCATCGTTACTATCGGGAT
CACGTACATTGCCATGCTAGGCGCTTTACAAGATGTCTCTATATTAGCTTTGGCATGA
TCTATGCGCTTTTCCTGTTCGGCGTCTTCGTTCAGGTTATTAGCGGAAATTGGGAACT
TCTCTATCCAGCCTATTTCCCATTCGACTTGGAGAGCAATCGCTTTCTCGGCGCAGTA
GCCTTGGGCTATCAGGTATTCAGCATGTTAGTTGAAGGCTTCCAGGGGCTGGGCAACG
ATACCTATACCCCACTGACCCTATGCCTTCTGGCCGGACATGTCCATTTGTGGTCCAT
ACGAATGGGTCAACTGGGATACTTCGATGACGAGACGGTGGTGAATCATCAGCGTTTG
CTGGATTACATTGAGCAGCATAAACTCTTGGTGCGGTTCCACAACCTGGTGAGCCGGA
CCATCAGCGAAGTGCAACTGGTGCAGCTGGGCGGATGTGGAGCCACTCTGTGCATCAT
TGTCTCCTACATGCTCTTCTTTGTGGGCGACACAATCTCGCTGGTCTACTACTTGGTG
TTCTTTGGAGTGGTCTGCGTGCAGCTCTTTCCCAGCTGCTATTTTGCCAGCGAAGTAG
CCGAGGAGTTGGAACGGCTGCCATATGCGATCTTCTCCAGCAGATGGTACGATCAATC

GCGGGATCATCGATTCGATTTGCTCATCTTTACACAATTAACACTGGGAAACCGGGGG
TGGATCATCAAGGCAGGAGGTCTTATCGAGCTGAATTTGAATGCCTTTTTCGCCACCC
TGAAGATGGCCTATTCCCTTTTTGCAGTTGTGGTGCGGGCAAAGGGTATA

DOR72

MDLKPRVIRSEDIYRTYWLYWHLLGLESNFFLNRLLDLVITIFVTIWYPIHLILGLFM
ERSLGDVCKGLPITAACFFASFKFICFRFKLSEIKEIEILFKELDQRALSREECEFFN
QNTRREANFIWKSFIVAYGLSNISAIASVLFGGGHKLLYPAWFPYDVQATELIFWLSV
TYQIAGVSLAILQNLANDSYPPMTFCVVAGHVRLLAMRLSRIGQGPEETIYLTGKQLI
ESIEDHRKLMKIVELLRSTMNISQLGQFISSGVNISITLVNILFFADNNFAITYYGVY
FLSMVLELFPCCYYGTLISVEMNQLTYAIYSSNWMSMNRSYSRILLIFMQLTLAEVQI
KAGGMIGIGMNAFFATVRLAYSFFTLAMSLR

DOR72nt

ATGGACTTAAAACCGCGAGTCATTCGAAGTGAAGATATCTACAGAACCTATTGGTTAT
ATTGGCATCTTTTGGGCCTGGAAAGCAATTTCTTTCTGAATCGCTTGTTGGATTTGGT
GATTACAATTTTCGTAACCATTTGGTATCCAATTCACCTGATTCTGGGACTGTTTATG
GAAAGATCTTTGGGGGATGTCTGCAAGGGTCTACCAATTACGGCAGCATGCTTTTTCG
CCAGCTTTAAATTTATTTGTTTTCGCTTCAAGCTATCTGAAATTAAAGAAATCGAAAT
ATTATTTAAAGAGCTGGATCAGCGAGCTTTAAGTCGAGAGGAATGCGAGTTTTTCAAT
CAAAATACGAGACGTGAGGCGAATTTCATTTGGAAAAGTTTCATTGTGGCCTATGGAC
TGTCGAATATCTCGGCTATTGCATCAGTTCTTTTCGGCGGTGGACATAAGCTATTATA
TCCCGCCTGGTTTCCATACGATGTGCAGGCCACGGAACTAATATTTTGGCTAAGTGTA
ACATACCAAATTGCCGGAGTAAGTTTGGCCATACTTCAGAATTTGGCCAATGATTCCT
ATCCACCGATGACATTTTGCGTGGTTGCCGGTCATGTAAGACTTTTGGCGATGCGCTT
GAGTAGAATTGGCCAAGGTCCAGAGGAAACAATATACTTAACCGGAAAGCAATTAATC
GAAAGCATCGAGGATCACCGAAAACTAATGAAGATAGTGGAATTACTGCGCAGCACCA
TGAATATTTCGCAGCTCGGCCAGTTTATTTCAAGTGGTGTTAATATTTCCATAACACT
AGTCAACATTCTCTTCTTTGCGGATAATAATTTCGCTATAACCTACTACGGAGTGTAC
TTCCTATCGATGGTGTTGGAATTATTCCCGTGCTGCTATTACGGCACCCTGATATCCG
TGGAGATGAACCAGCTGACCTATGCGATTTACTCAAGTAACTGGATGAGTATGAATCG
GAGCTACAGCCGCATCCTACTGATCTTCATGCAACTCACCCTGGCGGAAGTGCAGATC
AAGGCCGGTGGGATGATTGGCATCGGAATGAACGCCTTCTTTGCCACCGTGCGATTGG
CCTACTCCTTCTTCACTTTGGCCATGTCGCTGCGT

DOR73

MDSRRKVRSENLYKTYWLYWRLLGVEGDYPFRRLVDFTITSFITILFPVHLILGMYKK
PQIQVFRSLHFTSECLFCSYKFFCFRWKLKEIKTIEGLLQDLDSRVESEEERNYFNQN
PSRVARMLSKSYLVAAISAIITATVAGLFSTGRNLMYLGWFPYDFQATAAIYWISFSY
QAIGSSLLILENLANDSYPPITFCVVSGHVRLLIMRLSRIGHDVKLSSSENTRKLIEG
IQDHRKLMKIIRLLRSTLHLSQLGQFLSSGINISITLINILFFAENNFAMLYYAVFFA
AMLIELFPSCYYGILMTMEFDKLPYAIFSSNWLKMDKRYNRSLIILMQLTLVPVNIKA
GGIVGIDMSAFFATVRMAYSFYTLALSFRV

DOR73nt

ATGGATTCAAGAAGGAAAGTCCGAAGTGAAAATCTTTACAAAACCTATTGGCTTTACT
GGCGACTTCTGGGAGTCGAGGGCGATTATCCTTTTCGACGGCTAGTGGATTTTACAAT
CACGTCTTTCATTACGATTTTATTTCCCGTGCATCTTATACTGGGAATGTATAAAAAG
CCCCAGATTCAAGTCTTCAGGAGTCTGCATTTCACATCGGAATGCCTTTTCTGCAGCT
ATAAGTTTTTCTGTTTTCGTTGGAAACTTAAAGAAATAAAGACCATCGAAGGATTGCT
CCAGGATCTCGATAGTCGAGTTGAAAGTGAAGAAGAACGCAACTACTTTAATCAAAAT
CCAAGTCGTGTGGCTCGAATGCTTTCGAAAAGTTACTTGGTAGCTGCTATATCGGCCA
TAATCACTGCAACTGTAGCTGGTTTATTTAGTACTGGTCGAAATTTAATGTATCTGGG
TTGGTTTCCCTACGATTTTCAAGCAACCGCCGCAATCTATTGGATTAGTTTTTCCTAT
CAGGCGATTGGCTCTAGTCTGTTGATTCTGGAAAATCTGGCCAACGATTCATATCCGC
CGATTACATTTTGTGTGGTCTCTGGACATGTGAGACTATTGATAATGCGTTTAAGTCG
AATTGGTCACGATGTAAAATTATCAAGTTCGGAAAATACCAGAAAACTCATCGAAGGT
ATCCAGGATCACAGGAAACTAATGAAGATAATACGCCTACTTCGCAGCACTTTACATC
TTAGCCAACTGGGCCAGTTCCTTTCTAGTGGAATCAACATTTCCATAACACTCATCAA
CATCCTGTTCTTTGCGGAAAACAACTTTGCAATGCTTTATTATGCGGTGTTCTTTGCT
GCAATGTTAATAGAACTATTTCCAAGTTGTTACTATGGAATTCTGATGACAATGGAGT
TTGATAAGCTACCATATGCCATCTTCTCCAGCAACTGGCTTAAAATGGATAAAAGATA
CAATCGATCCTTGATAATTCTGATGCAACTAACACTGGTTCCAGTGAATATAAAAGCA
GGTGGTATTGTTGGCATCGATATGAGTGCATTTTTTGCCACAGTTCGGATGGCATATT
CCTTTTACACTTTAGCCTTGTCATTTCGAGTA

DOR77

MELMRVPVQFYRTIGEDIYAHRSTNPLKSLLFKIYLYAGFINFNLLVIGELVFFYNSI
QDFETIRLAIAVAPCIGFSLVADFKQAAMIRGKKTLIMLLDDLENMHPKTLAKQMEYK
LPDFEKTMKRVINIFTFLCLAYTTTFSFYPAIKASVKFNFLGYDTFDRNFGFLIWFPF
DATRNNLIYWIMYWDIAHGAYLAAFQVTESTVEVIIIYCIFLMTSMVQVFMVCYYGDT
LIAASLKVGDAAYNQKWFQCSKSYCTMLKLLIMRSQKPASIRPPTFPPISLVTYMKNP

FNNLPKHSSSLQINANRYI

DOR77nt

ATGGAATTGATGCGAGTGCCAGTACAGTTTTACAGAACGATTGGAGAGGATATCTACG
CCCATCGATCCACGAATCCCCTAAAATCGCTTCTCTTCAAGATCTATCTATATGCGGG
ATTCATAAATTTTAATCTGTTGGTAATCGGTGAACTGGTGTTCTTCTACAACTCAATT
CAGGACTTTGAAACCATTCGATTGGCCATCGCGGTGGCTCCATGTATCGGATTTTCTC
TGGTTGCTGATTTTAAACAAGCTGCCATGATTAGAGGCAAGAAAACACTAATTATGCT
ACTCGATGATTTGGAGAACATGCATCCGAAAACCCTGGCAAAGCAAATGGAATACAAA
TTGCCGGACTTTGAAAAGACCATGAAACGTGTGATCAATATATTCACCTTTCTCTGCT
TGGCCTATACGACTACGTTCTCCTTTTATCCGGCCATCAAGGCATCCGTGAAATTTAA
TTTCTTGGGCTACGACACCTTTGATCGAAATTTTGGTTTCCTCATCTGGTTTCCCTTC
GATGCAACAAGGAATAATTTGATATACTGGATCATGTACTGGGACATAGCCCATGGGG
CCTATCTAGCGGCCTTTCAGGTCACCGAATCAACAGTGGAAGTGATTATTATTTACTG
CATTTTTTTGATGACCTCGATGGTTCAGGTATTTATGGTGTGCTACTATGGGGATACT
TTAATTGCCGCGAGCTTGAAAGTGGGCGATGCCGCTTACAACCAAAAGTGGTTTCAGT
GCAGCAAATCCTATTGCACCATGTTGAAGTTGCTAATCATGAGGAGTCAGAAACCAGC
TTCAATAAGACCGCCGACTTTTCCCCCCATATCCTTGGTTACCTATATGAAGAATCCC
TTCAACAATCTACCCAAACACAGCTCTTCCCTGCAAATCAACGCCAATCGCTATATC

DOR78

MKFMKYAVFFYTSVGIEPYTIDSRSKKASLWSHLLFWANVINLSVIVFGEILYLGVAY
SDGKFIDAVTVLSYIGFVIVGMSKMFFIWWKKTDLSDLVKELEHIYPNGKAEEEMYRL
DRYLRSCSRISITYALLYSVLIWTFNLFSIMQFLVYEKLLKIRVVGQTLPYLMYFPWN
WHENWTYYVLLFCQNFAGHTSASGQISTDLLLCAVATQVVMHFDYLARVVEKQVLDRD
WSENSRFLAKTVQYHQRILRLMDVLNDIFGIPLLLNFMVSTFVICFVGFQMTVGVPPD
IMIKLFLFLFSSLSQVYLICHYGQLIADAVRDFRSSSLSISAYKQNWQNADIRYRRAL
VFFIARPQRTTYLKATIFMNITRATMTDVRYNLKCH

DOR78nt

ATGAAGTTCATGAAGTACGCAGTTTTCTTTTACACATCGGTGGGCATTGAGCCGTATA
CGATTGACTCGCGGTCCAAAAAAGCGAGCCTATGGTCACATCTTCTCTTCTGGGCCAA
TGTGATCAATTTAAGTGTCATTGTTTTCGGAGAGATCCTCTATCTGGGAGTGGCCTAT
TCCGATGGAAAGTTCATTGATGCCGTCACTGTACTGTCATATATCGGATTCGTAATCG
TGGGCATGAGCAAGATGTTCTTCATATGGTGGAAGAAGACCGATCTAAGCGATTTGGT
TAAGGAATTGGAGCACATCTATCCAAATGGCAAAGCTGAGGAGGAGATGTATCGGTTG
GATAGGTATCTGCGATCTTGTTCACGAATTAGCATTACCTATGCACTACTCTACTCCG

TACTCATCTGGACCTTCAATCTGTTCAGTATCATGCAATTCCTTGTCTATGAAAAGTT
GCTTAAAATCCGAGTGGTCGGCCAAACGCTGCCATATTTGATGTACTTTCCCTGGAAC
TGGCATGAAAACTGGACGTATTATGTGCTGCTGTTCTGTCAAAACTTCGCAGGACATA
CTTCGGCATCGGGACAGATCTCTACGGATCTTTTGCTTTGTGCTGTTGCTACCCAGGT
GGTAATGCACTTCGATTACTTGGCCAGAGTGGTGGAAAAACAAGTGTTAGATCGCGAT
TGGAGCGAAAACTCCAGATTTTTGGCAAAAACTGTACAATATCATCAGCGCATTCTTC
GGCTAATGGACGTTCTCAACGATATATTCGGGATACCGCTACTGCTTAACTTTATGGT
CTCCACATTTGTCATCTGCTTTGTGGGATTCCAAATGACCGTGGGTGTCCCGCCGGAC
ATCATGATTAAGCTCTTCTTGTTCCTGTTCTCGTCCTTGTCGCAAGTGTACTTGATAT
GCCACTACGGCCAGCTGATTGCCGATGCGGTAAGAGACTTTCGAAGCTCTAGCTTATC
GATTTCTGCATATAAGCAGAATTGGCAAAATGCTGACATTCGCTATCGTCGGGCTCTG
GTATTCTTTATAGCTCGACCTCAGAGGACAACTTATCTAAAAGCTACAATTTTCATGA
ATATAACAAGGGCCACCATGACGGACGTAAGATACAATTTGAAATGTCAT

DOR81

MMETLRNSGLNLKNDFGIGRKIWRVFSFTYNMVILPVSFPINYVIHLAEFPPELLLQS
LQLCLNTWCFALKFFTLIVYTHRLELANKHFDELDKYCVKPAEKRKVRDMVATITRLY
LTFVVVYVLYATSTLLDGLLHHRVPYNTYYPFINWRVDRTQMYIQSFLEYFTVGYAIY
VATATDSYPVIYVAALRTHILLLKDRIIYLGDPSNEGSSDPSYMFKSLVDCIKAHRTM
LNFCDAIQPIISGTIFAQFIICGSILGIIMINMVLFADQSTRFGIVIYVMAVLLQTFP
LCFYCNAIVDDCKELAHALFHSAWWVQDKRYQRTVIQFLQKLQQPMTFTAMNIFNINL
ATNINVSPLLSVRTGKEAKSELQSLQVAKFAFTVYAIASGMNLDQKLSIKE

DOR81nt

ATGATGGAGACGCTGCGAAATTCGGGCTTGAATTTGAAGAACGATTTCGGTATAGGCC
GCAAGATTTGGAGGGTGTTTTCGTTCACCTACAATATGGTGATACTTCCCGTAAGTTT
CCCAATCAACTATGTGATACATCTGGCGGAGTTCCCGCCGGAGCTGCTGCTGCAATCC
CTGCAACTGTGCCTCAACACTTGGTGCTTCGCTCTGAAGTTCTTCACTCTGATCGTCT
ATACGCACCGCTTGGAGCTGGCCAACAAGCACTTTGACGAATTGGATAAGTACTGCGT
GAAGCCGGCGGAGAAGCGCAAGGTTCGCGACATGGTGGCCACTATTACAAGACTGTAC
CTGACCTTCGTCGTGGTCTACGTCCTCTACGCCACCTCCACGCTACTGGACGGACTAC
TGCACCACCGTGTTCCCTACAATACGTACTATCCGTTCATAAACTGGCGAGTCGATCG
GACCCAGATGTACATCCAGAGTTTTCTGGAGTACTTCACCGTGGGTTATGCCATATAT
GTGGCCACCGCCACCGATTCCTACCCTGTGATTTACGTGGCAGCCCTGCGAACTCATA
TTCTCTTGCTCAAGGACCGTATCATTTACTTGGGCGATCCCAGCAACGAGGGTAGCAG
CGACCCGAGCTACATGTTTAAATCGTTGGTGGATTGTATCAAGGCACACAGAACCATG
CTAAAGTGCAGTTTTTGTGATGCCATTCAACCAATCATCTCTGGCACGATATTTGCCC
AATTCATCATATGCGGATCGATCCTGGGCATAATTATGATCAACATGGTATTGTTCGC

TGATCAATCGACCCGATTCGGCATAGTCATCTACGTTATGGCCGTCCTTCTGCAGACT
TTTCCGCTTTGCTTCTACTGCAACGCCATCGTGGACGACTGCAAAGAACTGGCCCACG
CACTTTTCCATTCCGCCTGGTGGGTGCAGGACAAGCGATACCAGCGGACTGTCATCCA
GTTCCTGCAGAAACTGCAGCAGCCCATGACCTTCACCGCCATGAACATATTTAACATT
AATTTGGCCACTAACATCAATGTAAGTCCACTGCTCTCGGTTAGAACGGGGAAGGAAG
CAAAGTCCGAACTTCAATCCTTGCAGGTAGCCAAGTTCGCCTTCACCGTGTACGCCAT
CGCGAGCGGTATGAACCTGGACCAAAAGTTAAGCATTAAGGAA

DOR82

MACIPRYQWKGRPTERQFYASEQRIVFLLGTICQIFQITGVLIYWYCNGRLATETGTF
VAQLSEMCSSFCLTFVGFCNVYAISTNRNQIETLLEELHQIYPRYRKNHYRCQHYFDM
AMTIMRIEFLFYMILYVYYNSAPLWVLLWEHLHEEYDLSFKTQTNTWFPWKVHGSALG
FGMAVLSITVGSFVGVGFSIVTQNLICLLTFQLKLHYDGISSQLVSLDCRRPGAHKEL
SILIAHHSRILQLGDQVNDIMNFVFGSSLVGATIAICMSSVSIMLLDLASAFKYASGL
VAFVLYNFVICYMGTEVTLAVKIGSYMDGRRWIPKDSLLRSQRLQVLVAVGFFNICVL
SNRRPKIEILLRYYYHIMFYSFKLYFSLRKGSLWKILSSFTLLRI

DOR82nt

ATGGCATGCATACCAAGATATCAATGGAAAGGACGCCCTACTGAAAGACAGTTCTACG
CTTCGGAGCAAAGGATAGTGTTCCTTCTTGGAACCATTTGCCAGATATTCCAGATTAC
TGGAGTGCTTATCTATTGGTATTGCAATGGCCGTCTTGCCACGGAAACGGGCACCTTT
GTGGCACAATTATCTGAAATGTGCAGTTCTTTTTGTCTAACATTTGTGGGATTCTGTA
ACGTTTATGCGATCTCTACAAACCGCAATCAAATTGAAACATTACTCGAGGAGCTTCA
TCAGATATATCCGAGATACAGGAAAAATCACTATCGCTGCCAGCATTATTTTGACATG
GCCATGACAATAATGAGAATTGAGTTTCTTTTCTATATGATCTTGTACGTGTACTACA
ATAGTGCACCATTATGGGTGCTTCTTTGGGAACACTTGCACGAGGAATATGATCTTAG
CTTCAAGACGCAGACCAACACTTGGTTTCCATGGAAAGTCCATGGGTCGGCACTTGGA
TTTGGTATGGCTGTACTAAGCATAACCGTGGGATCCTTTGTGGGCGTAGGTTTCAGTA
TTGTCACCCAGAATCTTATCTGTTTGTTAACCTTCCAACTAAAGTTGCACTACGATGG
AATATCCAGTCAGTTAGTATCTCTCGATTGCCGTCGTCCTGGAGCTCATAAGGAGTTG
AGCATCCTCATCGCCCACCACAGCCGAATCCTTCAGCTGGGCGACCAAGTCAATGACA
TAATGAACTTTGTATTCGGCTCTAGCCTAGTAGGTGCCACTATTGCCATTTGTATGTC
AAGTGTTTCTATAATGCTACTGGACTTAGCATCTGCCTTCAAATATGCCAGTGGTCTA
GTGGCATTCGTCCTCTACAACTTTGTCATCTGCTACATGGGAACCGAGGTCACTTTAG
CTGTGAAGATTGGTTCATATATGGACGGAAGGCGGTGGATACCCAAAGATTCGTTGCT
GAGATCTCAGAGGCTACAGGTGCTCGTCGCAGTTGGATTTTTTAATATATGTGTCCTC
TCGAATCGTCGTCCTAAAATTGAAATTTTGCTTAGATATTATTACCATATTATGTTTT
ATTCATTTAAATTATATTTTTCTTTAAGGAAAGGTAGCCTTTGGAAAATCTTGTCTTC

TTTCACCTTATTGAGGATC

DOR83

MQLEDFMRYPDLVCQAAQLPRYTWNGRRSLEVKRNLAKRIIFWLGAVNLVYHNIGCVM
YGYFGDGRTKDPIAYLAELASVASMLGFTIVGTLNLWKMLSLKTHFENLLNEFEELFQ
LIKHRAYRIHHYQEKYTRHIRNTFIFHTSAVVYYNSLPILLMIREHFSNSQQLGYRIQ
SNTWYPWQVQGSIPGFFAAVACQIFSCQTNMCVNMFIQFLINFFGIQLEIHFDGLARQ
LETIDARNPHAKDQLKYLIVYHTKLLNLADRVNRSFNFTFLISLSVSMISNCFLAFSM
TMFDFGTSLKHLLGLLLFITYNFSMCRSGTHLILTSGKVLPAAFYNNWYEGDLVYRRM
LLILMMRATKPYMWKTYKLAPVSITTYMAECKTKEAHEQRHFRRHERQKPRVARI

DOR83nt

ATGCAGTTGGAGGACTTTATGCGGTACCCGGACCTCGTGTGTCAAGCGGCCCAACTTC
CCAGATACACGTGGAATGGCAGACGATCCTTGGAAGTTAAACGCAACTTGGCAAAACG
CATTATCTTCTGGCTTGGAGCAGTAAATTTGGTTTATCACAATATTGGCTGCGTCATG
TATGGCTATTTCGGTGATGGAAGAACAAAGGATCCAATTGCGTATTTAGCTGAATTGG
CATCTGTGGCCAGCATGCTTGGTTTCACCATTGTGGGCACCCTCAACTTGTGGAAGAT
GCTGAGCCTTAAGACCCATTTTGAGAACCTACTAAATGAATTCGAGGAATTATTTCAA
CTAATCAAGCACAGGGCGTATCGCATACACCACTATCAAGAAAAGTATACGCGTCATA
TACGAAATACATTTATTTTCCATACCTCTGCCGTTGTCTACTACAACTCACTACCAAT
TCTTCTAATGATTCGGGAACATTTCTCGAACTCACAGCAGTTGGGCTATAGAATTCAG
AGTAATACCTGGTATCCCTGGCAGGTTCAGGGATCAATTCCTGGATTTTTTGCTGCAG
TCGCCTGTCAAATCTTTTCGTGCCAAACCAATATGTGCGTCAATATGTTTATCCAGTT
TCTGATCAACTTTTTTGGTATCCAGCTAGAAATACACTTCGATGGTTTGGCCAGGCAG
CTGGAGACCATCGATGCCCGCAATCCCCATGCCAAGGATCAATTGAAGTATCTGATTG
TATATCACACAAAATTGCTTAATCTAGCCGACAGAGTTAATCGATCGTTTAACTTTAC
GTTTCTCATAAGTCTGTCGGTATCCATGATATCCAACTGTTTTCTGGCATTTTCCATG
ACCATGTTCGACTTTGGCACCTCTCTAAAACATTTACTCGGACTTTTGCTATTCATCA
CATATAATTTTTCAATGTGCCGCAGTGGTACGCACTTGATTTTAACGAGTGGCAAAGT
ATTGCCAGCGGCCTTTTATAACAATTGGTATGAAGGCGATCTTGTTTATCGAAGGATG
CTCCTCATCCTGATGATGCGTGCTACGAAACCTTATATGTGGAAAACCTACAAGCTGG
CACCTGTATCCATAACTACATATATGGCAGAATGCAAAACAAAAGAAGCCCATGAACA
ACGCCATTTTAGACGCCATGAAAGACAAAAACCTCGGGTTGCACGAATA

DOR84

MVFSFYAEVATLVDRLRDNENFLESCILLSYVSFVVMGLSKIGAVMKKKPKMTALVRQ
LETCFPSPSAKVQEEYAVKSWLKRCHIYTKGFGGLFMIMYFAHALIPLFIYFIQRVLL

HYPDAKQIMPFYQLEPWEFRDSWLFYPSYFHQSSAGYTATCGSIAGDLMIFAVVLQVI
MHYERLAKVLREFKIQAHNAPNGAKEDIRKLQSLVANHIDILRLTDLMNEVFGIPLLL
NFIASALLVCLVGVQLTIALSPEYFCKQMLFLISVLLEVYLLCSFSQRLIDAVC

DOR84nt

ATGGTGTTTAGTTTTTATGCCGAGGTAGCGACTCTGGTGGACAGGTTACGCGATAATG
AAAATTTTCTCGAGAGCTGCATCTTACTGAGCTACGTGTCCTTTGTGGTCATGGGCCT
CTCCAAGATAGGTGCTGTAATGAAAAAAAAGCCAAAAATGACAGCTTTGGTCAGGCAA
TTGGAGACCTGCTTTCCGTCGCCAAGTGCAAAGGTTCAAGAGGAATATGCTGTGAAGT
CCTGGCTGAAACGCTGCCATATATACACAAAGGGATTTGGTGGTCTCTTCATGATCAT
GTATTTCGCTCACGCTCTGATTCCCTTATTCATATACTTCATTCAAAGAGTGCTGCTC
CACTATCCGGATGCCAAGCAGATTATGCCGTTTTACCAACTCGAACCTTGGGAATTTC
GCGACTCCTGGTTGTTTTATCCAAGCTATTTTCACCAGTCGTCGGCCGGATATACGGC
TACATGTGGATCCATTGCCGGTGACCTAATGATCTTCGCTGTGGTCCTGCAGGTCATC
ATGCACTACGAAAGACTGGCCAAGGTTCTTAGGGAGTTTAAGATTCAAGCCCATAACG
CACCCAATGGAGCTAAGGAGGATATAAGGAAGTTGCAGTCCCTAGTCGCCAATCACAT
TGATATACTTCGACTCACTGATCTGATGAACGAGGTCTTTGGAATTCCCTTGTTGCTA
AACTTTATTGCATCTGCGCTGCTGGTCTGCCTGGTGGGAGTTCAATTAACCATCGCTT
TAAGTCCAGAGTATTTTTGCAAGCAGATGCTATTTCTGATTTCCGTACTGCTTGAGGT
CTATCTCCTTTGCTCCTTCAGCCAGAGGTTAATAGATGCTGTATGT

DOR87

MTIEDIGLVGINVRMWRHLAVLYPTPGSSWRKFAFVLPVTAMNLMQFVYLLRMWGDLP
AFILNMFFFSAIFNALMRTWLVIIKRRQFEEFLGQLATLFHSILDSTDEWGRGILRRA
EREARNLAILNLSASFLDIVGALVSPLFREERAHPFGVALPGVSMTSSPVYEVIYLAQ
LPTPLLLSMMYMPFVSLFAGLAIFGKAMLQILVHRLGQIGGEEQSEEERFQRLASCIA
YHTQVMRYVWQLNKLVANIVAVEAIIFGSIICSLLFCLNIITSPTQVISIVMYILTML
YVLFTYYNRANEICLENNRVAEAVYNVPWYEAGTRFRKTLLIFLMQTQHPMEIRVGNV
YPMTLAMFQSLLNASYSYFTMLRGVTGK

DOR87nt

GGCACGAGGCTTATAGAAAGTGCCGAGCAATGACAATCGAGGATATCGGCCTGGTGGG
CATCAACGTGCGGATGTGGCGACACTTGGCCGTGCTGTACCCCACTCCGGGCTCCAGC
TGGCGCAAGTTCGCCTTCGTGCTGCCGGTGACTGCGATGAATCTGATGCAGTTCGTCT
ACCTGCTGCGGATGTGGGGCGACCTGCCCGCCTTCATTCTGAACATGTTCTTCTTCTC
GGCCATTTTCAACGCCCTGATGCGCACGTGGCTGGTCATAATCAAGCGGCGCCAGTTC
GAGGAGTTTCTCGGCCAACTGGCCACTCTGTTCCATTCGATTCTCGACTCCACCGACG

AGTGGGGGCGTGGCATCCTGCGGAGGGCGGAACGGGAGGCTCGGAACCTGGCCATCCT
TAATTTGAGTGCCTCCTTCCTGGACATTGTCGGTGCTCTGGTATCGCCGCTTTTCAGG
GAGGAGAGAGCTCATCCCTTCGGCGTAGCTCTACCAGGAGTGAGCATGACCAGTTCAC
CCGTCTACGAGGTTATCTACTTGGCCCAACTGCCTACGCCCCTGCTGCTGTCCATGAT
GTACATGCCTTTCGTCAGCCTTTTTGCCGGCCTGGCCATCTTTGGGAAGGCCATGCTG
CAGATCCTGGTACACAGGCTGGGCCAGATTGGCGGAGAAGAGCAGTCGGAGGAGGAGC
GCTTCCAAAGGCTGGCCTCCTGCATTGCGTACCACACGCAGGTGATGCGCTATGTGTG
GCAGCTCAACAAACTGGTGGCCAACATTGTGGCGGTGGAAGCAATTATTTTTGGCTCG
ATAATCTGCTCACTGCTCTTCTGTCTGAATATTATAACCTCACCCACCCAGGTGATCT
CGATAGTGATGTACATTCTGACCATGCTGTACGTTCTCTTCACCTACTACAATCGGGC
CAATGAAATATGCCTCGAGAACAACCGGGTGGCGGAGGCTGTTTACAATGTGCCCTGG
TACGAGGCAGGAACTCGGTTTCGCAAAACCCTCCTGATCTTCTTGATGCAAACACAAC
ACCCGATGGAGATAAGAGTCGGCAACGTTTACCCCATGACATTGGCCATGTTCCAGAG
TCTGTTGAATGCGTCCTACTCCTACTTTACCATGCTGCGTGGCGTCACCGGCAAATGA
GCTGAAAGACCGAAAAAACCGGAGTATCCCCTTCCATATTCCCCCTGCTCCTTTATTT
TCCTTTCCTTTTCCCTTTCCGTTTTCCCATTCGCTTTTCCAGCAATCCGGGTAATGCA
AAAAGTTGTTGCTGGCTGTGGTCCTGGCTGCTTGTTTGGCATTTGCATATGCTTGTCG
TTTGAAAGGATTTAATCGGACTGCTGGCACGGAGTCGGCATCCTGGCTCCTGGATCCT
GGCATGCAAATAGTTGGCTTCTTAGATTGTTACACAAAATAGATTGTAGATTGCAGCT
GAATGTTGTGCTTGGAATAAAGTCAAAAGGATGTGGAGTCGGCCCAAGGCTCTGCCCA
TTCTGTTTGCTCGGGATGCCCGAAAGTATGAAAAAAAAAAAAAAAAAA

DOR91

MVRYVPRFADGQKVKLAWPLAVFRLNHIFWPLDPSTGKWGRYLDKVLAVAMSLVFMQH
NDAELRYLRFEASNRNLDAFLTGMPTYLILVEAQFRSLHILLHFEKLQKFLEIFYANI
YIDPRKEPEMFRKVDGKMIINRLVSAMYGAVISLYLIAPVFSIINQSKDFLYSMIFPF
DSDPLYIFVPLLLTNVWVGIVIDTMMFGETNLLCELIVHLNGSYMLLKRDLQLAIEKI
LVARDRPHMAKQLKVLITKTLRKNVALNQFGQQLEAQYTVRVFIMFAFAAGLLCALSF
KAYTTDSLSTMYYLTHWEQILQYSTNPSENLRLLKLINLAIEMNSKPFYVTGLKYFRV
SLQAGLKRQKFLRSASSSTLSTADVLAFAFAFTRWLL

DOR91nt

ATGGTTCGTTACGTGCCCCGGTTCGCTGATGGTCAGAAAGTAAAGTTGGCTTGGCCCT
TGGCGGTTTTTCGGTTAAATCACATATTCTGGCCATTGGATCCGAGCACAGGGAAATG
GGGCCGATATCTGGACAAGGTTCTAGCTGTTGCGATGTCCTTGGTTTTTATGCAACAC
AACGATGCAGAGCTGAGGTACTTGCGCTTCGAGGCAAGTAATCGGAATTTGGATGCCT
TTCTCACAGGAATGCCAACGTATTTAATCCTCGTGGAGGCTCAATTTAGAAGTCTTCA
CATTCTACTGCACTTCGAGAAGCTTCAGAAGTTTTTAGAAATATTCTACGCAAATATT

TATATTGATCCCCGTAAGGAACCCGAAATGTTTCGAAAAGTGGATGGAAAGATGATAA
TTAACAGATTAGTTTCGGCCATGTACGGTGCAGTTATCTCTCTGTATCTAATCGCACC
CGTTTTTTCCATCATTAACCAAAGCAAAGATTTTCTATACTCTATGATCTTTCCGTTC
GATTCGGATCCCTTGTACATATTTGTGCCACTGCTTTTGACAAACGTATGGGTTGGCA
TTGTAATAGATACCATGATGTTCGGGGAGACGAATTTGTTGTGTGAACTAATTGTCCA
CCTAAATGGTAGTTATATGTTGCTCAAGAGGGACTTGCAGTTGGCCATTGAAAAGATA
TTAGTTGCAAGGGACCGTCCGCATATGGCCAAACAGCTAAAGGTTTTAATTACAAAAA
CTCTCCGAAAGAATGTGGCTCTAAATCAGTTTGGCCAGCAGCTGGAGGCTCAGTATAC
TGTGCGGGTTTTTATTATGTTTGCATTCGCTGCGGGCCTTTTATGTGCTCTTTCTTTT
AAGGCTTATACGACGGATTCCCTCAGCACAATGTACTACCTTACCCATTGGGAGCAAA
TCCTGCAGTACTCTACAAATCCCAGCGAAAATCTGCGATTACTAAAGCTCATTAACTT
GGCCATTGAGATGAACAGCAAGCCCTTCTATGTGACAGGGCTAAAATATTTTCGCGTT
AGTCTGCAGGCTGGCTTAAAACGTCAAAAGTTTCTGCGGTCTGCCAGCTCATCCACCC
TTAGCACCGCTGATGTGTTGGCATTTGCTTTTGCTTTTACTCGCTGGCTGCTT

DOR92

MSEWLRFLKRDQQLDVYFFAVPRLSLDIMGYWPGKTGDTWPWRSLIHFAILAIGVATE
LHAGMCFLDRQQITLALETLCPAGTSAVTLLKMFLMLRFRQDLSIMWNRLRGLLFDPN
WERPEQRDIRLKHSAMAARINFWPLSAGFFTCTTYNLKPILIAMILYLQNRYEDFVWF
TPFNMTMPKVLLNYPFFPLTYIFIAYTGYVTIFMFGGCDGFYFEFCAHLSALFEVLQA
EIESMFRPYTDHLELSPVQLYILEQKMRSVIIRHNAIIDLTRFFRDRYTIITLAHFVS
AAMVIGFSMVNLLTLGNNGLGAMLYVAYTVAALSQLLVYCYGGTLVAESSTGLCRAMF
SCPWQLFKPKQRRLVQLLILRSQRPVSMAVPFFSPSLATFAAILQTSGSIIALVKSFQ

DOR92nt

ATGTCCGAGTGGTTACGCTTTCTGAAACGCGATCAACAGCTGGATGTGTACTTTTTTG
CAGTGCCCCGCTTGAGTTTAGACATAATGGGCTATTGGCCGGGCAAAACTGGTGATAC
ATGGCCCTGGAGATCCCTGATTCACTTCGCAATCCTGGCCATTGGCGTGGCCACCGAA
CTGCATGCTGGCATGTGTTTTCTAGACCGACAGCAGATTACCTTGGCACTGGAGACCC
TCTGTCCAGCTGGCACATCGGCGGTCACGCTGCTCAAGATGTTCCTAATGCTGCGCTT
TCGTCAGGATCTCTCCATTATGTGGAACCGCCTGAGGGGCCTGCTCTTCGATCCCAAC
TGGGAGCGACCCGAGCAGCGGGACATCCGGCTAAAGCACTCGGCCATGGCGGCTCGCA
TCAATTTCTGGCCCCTGTCAGCCGGATTCTTCACATGCACCACCTACAACCTAAAGCC
GATACTGATCGCAATGATATTGTATCTCCAGAATCGTTACGAGGACTTCGTTTGGTTT
ACACCCTTCAATATGACTATGCCCAAAGTTCTGCTAAACTATCCATTTTTTCCCCTGA
CCTACATATTTATTGCCTATACGGGCTATGTGACCATCTTTATGTTCGGCGGCTGTGA
TGGTTTTTATTTCGAGTTCTGTGCCCACCTATCAGCTCTTTTCGAAGTGCTCCAGGCG
GAGATAGAATCAATGTTTAGACCCTACACTGATCACTTGGAACTGTCGCCAGTGCAGC

TTTACATTTTAGAGCAAAAGATGCGATCAGTAATCATTAGGCACAATGCCATCATCGA
TTTGACCAGATTTTTTCGTGATCGCTATACCATTATTACCCTGGCCCATTTTGTGTCC
GCCGCCATGGTGATTGGATTCAGCATGGTTAATCTCCTGACATTGGGCAATAATGGTC
TGGGCGCAATGCTCTATGTGGCCTACACGGTTGCCGCTTTGAGCCAACTGCTGGTTTA
TTGCTATGGCGGAACTCTGGTGGCCGAAAGTAGCACTGGTCTGTGCCGAGCCATGTTC
TCCTGTCCGTGGCAGCTTTTTAAGCCTAAACAACGTCGACTCGTTCAGCTTTTGATTC
TCAGATCGCAGCGTCCTGTTTCCATGGCAGTGCCATTCTTTTCGCCATCGTTGGCTAC
CTTTGCTGCGATTCTTCAAACTTCGGGTTCCATAATTGCGCTGGTTAAGTCCTTTCAG

DOR95

MSDKVKGKKQEEKDQSLRVQILVYRCMGIDLWSPTMANDRPWLTFVTMGPLFLFMVPM
FLAAHEYITQVSLLSDTLGSTFASMLTLVKFLLFCYHRKEFVGLIYHIRAILAKEIEV
WPDAREIIEVENQSDQMLSLTYTRCFGLAGIFAALKPFVGIILSSIRGDEIHLELPHN
GVYPYDLQVVMFYVPTYLWNVMASYSAVTMALCVDSLLFFFTYNVCAIFKIAKHRMIH
LPAVGGKEELEGLVQVLLLHQKGLQIADHIADKYRPLIFLQFFLSALQICFIGFQVAD
LFPNPQSLYFIAFVGSLLIALFIYSKCGENIKSASLDFGNGLYETNWTDFSPPTKRAL
LIAAMRAQRPCQMKGYFFEASMATFSTIVRSAVSYIMMLRSFNA

DOR95nt

ATGAGCGACAAGGTGAAGGGAAAAAAGCAGGAGGAAAAGGATCAATCCTTGCGGGTG CAAATTC
CCAGCTATAGTGCTGTAACCATGGCACTCTGCGTGGACTCGCTGCTCTTCTTTTTCAC
CTACAACGTGTGCGCCATTTTCAAGATCGCCAAGCACCGGATGATCCATCTGCCGGCG
GTGGGCGGAAAGGAGGAGCTGGAGGGGCTCGTCCAGGTGCTGCTGCTGCACCAGAAGG
GCCTCCAGATCGCCGATCACATTGCGGACAAGTACCGGCCGCTGATCTTTTTGCAGTT
CTTTCTGTCCGCCTTGCAGATCTGCTTCATTGGATTCCAGGTGGCTGATCTGTTTCCC
AATCCGCAGAGTCTCTACTTTATCGCCTTTGTGGGCTCGCTGCTCATCGCACTGTTCA
TCTACTCGAAGTGCGGCGAAAATATCAAGAGTGCCAGCCTGGATTTCGGAAACGGGCT
GTACGAGACCAACTGGACCGACTTCTCGCCACCCACTAAAAGAGCCCTCCTCATTGCC
GCCATGCGCGCCCAGCGACCTTGCCAGATGAAGGGCTACTTTTTCGAGGCCAGCATGG
CCACCTTCTCGACGATTGTTCGCTCTGCCGTGTCGTACATCATGATGTTGCGCTCCTT
TAATGCC

DOR99

MEEFLRPQMFQEVAQMVHFQWRRNPVDNSMVNASMVPFCLSAFLNVLFFGCNGWDIIG
HFWLGHPANQNPPVLSITIYFSIRGLMLYLKRKEIVEFVNDLDRECPRDLVSQLDMQM
DETYRNFWQRYRFIRIYSHLGGPMFCVVPLALFLLTHEGKDTPVAQHEQLLGGWLPCG
VRKDPNFYLLVWSFDLMCTTCGVSFFVTFDNLFNVMQGHLVMHLGHLARQFSAIDPRQ

SLTDEKRFFVDLRLLVQRQQLLNGLCRKYNDIFKVAFLVSNFVGAGSLCFYLFMLSET
SDVLIIAQYILPTLVLVGFTFEICLRGTQLEKASEGLESSLRSQEWYLGSRRYRKFYL
LWTQYCQRTQQLGAFGLIQVNMVHFTEIMQLAYRLFTFLKSH

DOR99nt

ATGGAGGAGTTTCTGCGTCCGCAGATGTTCCAGGAGGTGGCTCAGATGGTGCATTTCC
AGTGGCGGAGAAATCCGGTGGACAACAGCATGGTGAACGCATCCATGGTCCCCTTCTG
CTTGTCGGCGTTTCTTAATGTCCTGTTTTTCGGCTGCAATGGTTGGGACATCATAGGA
CATTTTTGGCTGGGACATCCTGCCAACCAGAATCCGCCCGTGCTTAGCATCACCATTT
ACTTCTCGATCAGGGGATTGATGCTATACCTGAAACGAAAGGAAATCGTTGAGTTTGT
TAACGACTTGGATCGGGAGTGTCCGCGGGACTTGGTCAGCCAGTTGGACATGCAAATG
GATGAGACGTACCGAAACTTTTGGCAGCGCTATCGCTTCATCCGTATCTACTCCCATT
TGGGTGGTCCGATGTTCTGCGTTGTGCCATTAGCTCTATTCCTCCTGACCCACGAGGG
TAAAGATACTCCTGTTGCCCAGCACGAGCAGCTCCTTGGAGGATGGCTGCCATGCGGT
GTGCGAAAGGACCCAAATTTCTACCTTTTAGTCTGGTCCTTCGACCTGATGTGCACCA
CTTGCGGCGTCTCCTTTTTCGTTACCTTCGACAACCTATTCAATGTGATGCAGGGACA
TTTGGTCATGCATTTGGGCCATCTTGCTCGCCAGTTTTCGGCCATCGATCCTCGACAG
AGTTTGACCGATGAGAAGCGATTCTTTGTGGATCTTAGGTTATTAGTTCAGAGGCAGC
AGCTTCTTAATGGATTGTGCAGAAAATACAACGACATCTTTAAAGTGGCCTTCCTGGT
GAGCAATTTTGTAGGCGCCGGTTCCCTCTGCTTCTACCTCTTTATGCTCTCGGAGACA
TCAGATGTCCTTATCATCGCCCAGTATATATTACCCACTTTGGTCCTGGTGGGCTTCA
CATTTGAGATTTGTCTACGGGGAACCCAACTGGAAAAGGCGTCGGAGGGACTGGAATC
GTCGTTGCGAAGCCAGGAATGGTATTTGGGAAGTAGGCGGTACCGGAAGTTCTATTTG
CTCTGGACGCAATATTGCCAGCGAACACAGCAACTGGGCGCCTTTGGGCTAATCCAAG
TCAATATGGTGCACTTCACTGAAATAATGCAGCTGGCCTATAGACTCTTCACTTTTCT
CAAATCTCAT

DORA45

MTTSMQPSKYTGLVADLMPNIRAMKYSGLFMHNFTGGSAFMKKVYSSVHLVFLLMQFT
FILVNMALNAEEVNELSGNTITTLFFTHCITKFIYLAVNQKNFYRTLNIWNQVNTHPL
FAESDARYHSIALAKMRKLFFLVMLTTVASATAWTTITFFGDSVKMVVDHETNSSIPV
EIPRLPIKSFYPWNASHGMFYMISFAFQIYYVLFSMIHSNLCDVMFCSWLIFACEQLQ
HLKGIMKPLMELSASLDTYRPNSAALFRSLSANSKSELIHNEEKDPGTDMDMSGIYSS
KADWGAQFRAPSTLQSFGGNGGGGNGLVNGANPNGLTKKQEMMVRSAIKYWVERHKHV
VRLVAAIGDTYGAALLLHMLTSTIKLTLLAYQATKINGVNVYAFTVVGYLGYALAQVF
HFCIFGNRLIEESSSVMEAAYSCHWYDGSEEAKTFVQIVCQQCQKAMSISGAKFFTVS
LDLFASVLGAVVTYFMVLVQLK

<u>DORA45nt</u>

GGCACGAGCTGGTTCCGGAAAGCCTCATATCTCGTATCTTAAAGTATCCCGGTTAAGC
CTTAAAGAGTGAAATGATTGCCTAGACGATTGCTGCATTACTGGCACTCAATTAACCC
AAGTGTACCAGACAACAATTACATTTGTATTTTTAAAGTTCAATAGCAAGGATGACAA
CCTCGATGCAGCCGAGCAAGTACACGGGCCTGGTCGCCGACCTGATGCCCAACATCCG
GGCGATGAAGTACTCCGGCCTGTTCATGCACAACTTCACGGGCGGCAGTGCCTTCATG
AAGAAGGTGTACTCCTCCGTGCACCTGGTGTTCCTCCTCATGCAGTTCACCTTCATCC
TGGTCAACATGGCCCTGAACGCCGAGGAGGTCAACGAGCTGTCGGGCAACACGATCAC
GACCCTCTTCTTCACCCACTGCATCACGAAGTTTATCTACCTGGCTGTTAACCAGAAG
AATTTCTACAGAACATTGAATATATGGAACCAGGTGAACACGCATCCCTTGTTCGCCG
AGTCGGATGCTCGTTACCATTCGATCGCACTGGCGAAGATGAGGAAGCTGTTCTTTCT
GGTGATGCTGACCACAGTCGCCTCGGCCACCGCCTGGACCACGATCACCTTCTTTGGC
GACAGCGTAAAAATGGTGGTGGACCATGAGACGAACTCCAGCATCCCGGTGGAGATAC
CCCGGCTGCCGATTAAGTCCTTCTACCCGTGGAACGCCAGCCACGGCATGTTCTACAT
GATCAGCTTTGCCTTTCAGATCTACTACGTGCTCTTCTCGATGATCCACTCCAATCTA
TGCGACGTGATGTTCTGCTCTTGGCTGATATTCGCCTGCGAGCAGCTGCAGCACTTGA
AGGGCATCATGAAGCCGCTGATGGAGCTGTCCGCCTCGCTGGACACCTACAGGCCCAA
CTCGGCGGCCCTCTTCAGGTCCCTGTCGGCCAACTCCAAGTCGGAGCTAATTCATAAT
GAAGAAAAGGATCCCGGCACCGACATGGACATGTCGGGCATCTACAGCTCGAAAGCGG
ATTGGGGCGCTCAGTTTCGAGCACCCTCGACACTGCAGTCCTTTGGCGGGAACGGGGG
CGGAGGCAACGGGTTGGTGAACGGCGCTAATCCCAACGGGCTGACCAAAAAGCAGGAG
ATGATGGTGCGCAGTGCCATCAAGTACTGGGTCGAGCGGCACAAGCACGTGGTGCGAC
TGGTGGCTGCCATCGGCGATACTTACGGAGCCGCCCTCCTCCTCCACATGCTGACCTC
GACCATCAAGCTGACCCTGCTGGCATACCAGGCCACCAAAATCAACGGAGTGAATGTC
TACGCCTTCACAGTCGTCGGATACCTAGGATACGCGCTGGCCCAGGTGTTCCACTTTT
GCATCTTTGGCAATCGTCTGATTGAAGAGAGTTCATCCGTCATGGAGGCCGCCTACTC
GTGCCACTGGTACGATGGCTCCGAGGAGGCCAAGACCTTCGTCCAGATCGTGTGCCAG
CAGTGCCAGAAGGCGATGAGCATATCGGGAGCGAAATTCTTCACCGTCTCCCTGGATT
TGTTTGCTTCGGTTCTGGGTGCCGTCGTCACCTACTTTATGGTGCTGGTGCAGCTCAA
GTAAGTTGCTGCGAAGCTGATGGATTTTTGTACCAGAAAAGCGAATGCCAAGAAGCCA
CCTACCGCCCCTTGCCCCCTCCGCACTGTGCAACCAGCAATATCACAGAGCAATTATA
ACGCAAATTATATATTTTATACCTGCGACGAGCGAGCCTCGTGGGGCATAATGGAGAC
ATTCTGGGGCACATAGAAGCCTGCAAATACTTATCGATTTTGTACACGCGTAGAGCTT
TTAATGTAAACTCAAGATGCAAACTAAATAAATGTGTAGTGAAAAAAAAAAAAAAAAA
AAA

GENBANK ACCESSION NUMBERS

The accession numbers for the sequences reported in this paper are AF127921-AF127926.

REFERENCES

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

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389-3402.

Amrein, H., Gorman, M., and Nothiger, R. (1988). The sex-determining gene *tra-2* of Drosophila encodes a putative RNA binding protein [published erratum appears in Cell 1989 Jul 28;58(2):following 419]. Cell *55*, 1025-1035.

Ayer, R.K., Jr., and Carlson, J. (1992) Olfactory physiology in the Drosophila antenna and maxillary palp: *acj6* distinguishes two classes of odorant pathways. J. Neurobiol. *23*, 965-982.

Bargmann, C. I., and Kaplan, J. M. (1998). Signal transduction in the Caenorhabditis elegans nervous system. Annu Rev Neurosci *21*, 279-308.

Bargmann, C. I., Hartwieg, E., and Horvitz, H. R. (1993). Odorant-selective genes and neurons mediate olfaction in C. elegans. Cell *74*, 515-527.

Bargmann, C. I., and Horvitz, H. R. (1991). Chemosensory neurons with overlapping functions direct chemotaxis to multiple chemicals in C. elegans. Neuron 7, 729-742.

Ben-Arie, N., Lancet, D., Taylor, C., Khen, M., Walker, N., Ledbetter, D. H., Carrozzo, R., Patel, K., Sheer, D., Lehrach, H., and et al. (1994). Olfactory receptor gene cluster on human chromosome 17: possible duplication of an ancestral receptor repertoire. Hum. Mol. Genet. *3*, 229-235.

Buck, L. B. (1996). Information coding in the vertebrate olfactory system. Annu Rev Neurosci *19*, 517-44.

Buck, L., and Axel, R. (1991). A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. Cell *65*, 175-187.

Burge, C., and Karlin, S. (1997). Prediction of complete gene structures in human genomic DNA. J. Mol. Biol. *268*, 78-94.

Carlson, J. R. (1996). Olfaction in Drosophila: from odor to behavior. Trends Genet. *12*, 175-180.

Chess, A., Simon, I., Cedar, H., and Axel, R. (1994). Allelic inactivation regulates olfactory receptor gene expression. Cell *78*, 823-834.

Colbert, H.A., and Bargmann, C.I. (1995) Odorant-specific adaptation pathways generate olfactory plasticity in C. elegans. Neuron *14*, 803-812.

Cserzo, M., Wallin, E., Simon, I., von Heijne, G., and Elofsson, A. (1997). Prediction of transmembrane -helices in prokaryotic membrane proteins: the dense alignment surface method. Protein Eng. *10*, 673-676.

Doe, C. Q., and Skeath, J. B. (1996). Neurogenesis in the insect central nervous system. Curr. Opin. Neurobiol. *6*, 18-24.

Dulac, C., and Axel, R. (1995). A novel family of genes encoding putative pheromone receptors in mammals. Cell *83*, 195-206.

Faber, T., Joerges, J., and Menzel, R. (1998). Associative learning modifies neural representations of odors in the insect brain. Nature Neurosci. *2*, 74-78.

Friedrich, R. W., and Korsching, S. I. (1997). Combinatorial and chemotopic odorant coding in the zebrafish olfactory bulb visualized by optical imaging. Neuron *18*, 737-752.

Grillenzoni, N., van Helden, J., Dambly-Chaudiere, C., and Ghysen, A. (1998). The *iroquois* complex controls the somatotopy of Drosophila notum mechanosensory projections. Development *125*, 3563-9.

Hartl, D. L., Nurminsky, D. I., Jones, R. W., and Lozovskaya, E. R. (1994). Genome structure and evolution in Drosophila: applications of the framework P1 map. Proc. Natl. Acad. Sci. U S A *91*, 6824-6829.

Herrada, G., and Dulac, C. (1997). A novel family of putative pheromone receptors in mammals with a topographically organized and sexually dimorphic distribution. Cell *90*, 763-773.

Imamura, K., Mataga, N., and Mori, K. (1992). Coding of odor molecules by mitral/tufted cells in rabbit olfactory bulb. I. Aliphatic compounds. J. Neurophysiol. *68*, 1986-2002.

Joerges, J., Kuttner, A., Galizia, C. G., and Menzel, R. (1997). Presentations of odours and odour mixtures visualized in the honeybee brain. Nature *387*, 285-288.

Katoh, K., Koshimoto, H., Tani, A., and Mori, K. (1993). Coding of odor molecules by mitral/tufted cells in rabbit olfactory bulb. II. Aromatic compounds. J. Neurophysiol. *70*, 2161-2175.

Kauer, J. S., Senseman, D. M., and Cohen, L. B. (1987). Odor-elicited activity monitored simultaneously from 124 regions of the salamander olfactory bulb using a voltage-sensitive dye. Brain Res. *418*, 255-261.

Kim, M. S., Repp, A., and Smith, D. P. (1998). LUSH odorant-binding protein mediates chemosensory responses to alcohols in Drosophila melanogaster. Genetics *150*, 711-21.

Kimmerly, W., Stultz, K., Lewis, S., Lewis, K., Lustre, V., Romero, R., Benke, J., Sun, D., Shirley, G., Martin, C., and Palazzolo, M. (1996). A P1-based physical map of the Drosophila euchromatic genome. Genome Res. *6*, 414-430.

Kyte, J., and Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. J Mol. Biol. *157*, 105-132.

Laissue, P. P., Reiter, C., Hiesinger, P. R., Halter, S., Fischbach, K. F., and Stocker, R. F. (1999) 3D reconstruction of the antennal lobe in Drosophila melanogaster. J. Comp. Neurol., in press.

Lancet, D., Greer, C. A., Kauer, J. S., and Shepherd, G. M. (1982). Mapping of odor-related neuronal activity in the olfactory bulb by high-resolution 2-deoxyglucose autoradiography. Proc. Natl. Acad. Sci. USA *79*, 670-674.

Levy, N. S., Bakalyar, H. A., and Reed, R. R. (1991). Signal transduction in olfactory neurons. J. Steroid Biochem. Mol. Biol. *39*, 633-637.

Lin, D. M., and Goodman, C. S. (1994). Ectopic and increased expression of Fasciclin II alters motoneuron growth cone guidance. Neuron *13*, 507-523.

Matsunami, H., and Buck, L. B. (1997). A multigene family encoding a diverse array of putative pheromone receptors in mammals. Cell *90*, 775-84.

McKenna, M. P., Hekmat-Scafe, D. S., Gaines, P., and Carlson, J. R. (1994). Putative Drosophila pheromone-binding proteins expressed in a subregion of the olfactory system. J Biol. Chem. *269*, 16340-16347.

McLatchie, L. M., Fraser, N. J., Main, M. J., Wise, A., Brown, J., Thompson, N., Solari, R., Lee, M. G., and Foord, S. M. (1998). RAMPs regulate the transport and ligand specificity of the calcitonin- receptor-like receptor. Nature *393*, 333-9.

Merritt, D. J., and Whitington, P. M. (1995). Central projections of sensory neurons in the Drosophila embryo correlate with sensory modality, soma position, and proneural gene function. J Neurosci *15*, 1755-67.

Mitchell, R., McCulloch, D., Lutz, E., Johnson, M., MacKenzie, C., Fennell, M., Fink, G., Zhou, W., and Sealfon, S. C. (1998). Rhodopsin-family receptors associate with small G proteins to activate phospholipase D. Nature *392*, 411-4.

Mombaerts, P., Wang, F., Dulac, C., Chao, S. K., Nemes, A., Mendelsohn, M., Edmondson, J., and Axel, R. (1996). Visualizing an olfactory sensory map. Cell *87*, 675-686.

Mori, K., Mataga, N., and Imamura, K. (1992). Differential specificities of single mitral cells in rabbit olfactory bulb for a homologous series of fatty acid odor molecules. J. Neurophysiol. *67*, 786-789.

Ngai, J., Chess, A., Dowling, M. M., Necles, N., Macagno, E. R., and Axel, R. (1993). Coding of olfactory information:

topography of odorant receptor expression in the catfish olfactory epithelium. Cell *72*, 667-680.

Parmentier, M., Libert, F., Schurmans, S., Schiffmann, S., Lefort, A., Eggericks, D., Ledent, C., Molleareau, C., Gerard, D., and et al. (1992). Expression of members of the putative olfactory receptor gene family in mammalian germ cells. Nature *355*, 453-455.

Pelosi, P. (1994). Odorant-binding proteins. Crit. Rev. Biochem. Mol. Biol. *29*, 199-228.

Persson, B., and Argos, P. (1994). Prediction of transmembrane segments in proteins utilizing multiple sequence alignments. J. Mol. Biol. *237*, 182-192.

Pikielny, C. W., Hasan, G., Rouyer, F., and Rosbash, M. (1994). Members of a family of Drosophila putative odorant-binding proteins are expressed in different subsets of olfactory hairs. Neuron *12*, 35-49.

Ray, K., and Rodrigues, V. (1995). Cellular events during development of the olfactory sense organs in Drosophila melanogaster. Dev. Biol. *167*, 426-38.

Reddy, G. V., Gupta, B., Ray, K., and Rodrigues, V. (1997). Development of the Drosophila olfactory sense organs utilizes cell-cell interactions as well as lineage. Development *124*, 703-712.

Ressler, K. J., Sullivan, S. L., and Buck, L. B. (1993). A zonal organization of odorant receptor gene expression in the olfactory epithelium. Cell *73*, 597-609.

Ressler, K. J., Sullivan, S. L., and Buck, L. B. (1994). Information coding in the olfactory system: evidence for a stereotyped and highly organized epitope map in the olfactory bulb. Cell *79*, 1245-1255.

Robertson, H. M. (1998). Two large families of chemoreceptor genes in the nematodes Caenorhabditis elegans and Caenorhabditis briggsae reveal extensive gene duplication, diversification, movement, and intron loss. Genome Res. *8*, 449-463.

Robinow, S. and White, K. (1988) The locus *elav* of Drosophila melanogaster is expressed in neurons at all developmental stages. Dev. Biol. *126*, 294-303.

Rodrigues, V. (1988). Spatial coding of olfactory information in the antennal lobe of Drosophila melanogaster. Brain Res. *453*, 299-307.

Ryba, N. J., and Tirindelli, R. (1997). A new multigene family of putative pheromone receptors. Neuron *19*, 371-379.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual, Second Edition (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).

Schaeren-Wiemers, N., and Gerfin-Moser, A. (1993). A single protocol to detect transcripts of various types and expression levels in neural tissue and cultured cells: *in situ* hybridization using digoxigenin-labelled cRNA probes. Histochemistry *100*, 431-440.

Sengupta, P., Chou, J. H., and Bargmann, C. I. (1996). *odr-10* encodes a seven transmembrane domain olfactory receptor required for responses to the odorant diacetyl. Cell *84*, 899-909.

Siddiqi, O. (1987). Neurogenetics of olfaction in Drosophila melanogaster. Trends Genet. *3*, 137-142.

Siden-Kiamos, I., Saunders, R. D., Spanos, L., Majerus, T., Treanear, J., Savakis, C., Louis, C., Glover, D. M., Ashburner, M., and Kafatos, F. C. (1990). Towards a physical map of the Drosophila melanogaster genome: mapping of cosmid clones within defined genomic divisions. Nucleic Acids Res. *18*, 6261-6270.

Singh, R. N., and Nayak, S. (1985). Fine structure and primary sensory projections of sensilla on the maxillary palp of Drosophila melanogaster Meigen (Diptera: Drosophilidae). Int. J. Insect Morphol. Embryol. *14*, 291-306.

Stewart, W. B., Kauer, J. S., and Shepherd, G. M. (1979). Functional organization of rat olfactory bulb analyzed by the 2-deoxyglucose method. J. Comp. Neurol. *185*, 715-734.

Stocker, R. F. (1994). The organization of the chemosensory system in Drosophila melanogaster: a review. Cell Tissue Res. *275*, 3-26.

Stocker, R. F., Lienhard, M. C., Borst, A., and Fischbach, K. F. (1990). Neuronal architecture of the antennal lobe in Drosophila melanogaster. Cell Tissue Res. *262*, 9-34.

Stocker, R. F., Singh, R. N., Schorderet, M., and Siddiqi, O. (1983). Projection patterns of different types of antennal sensilla in the antennal glomeruli of Drosophila melanogaster. Cell Tissue Res. *232*, 237-248.

Thummel, C. S., Boulet, A. M., and Lipshitz, H. D. (1988). Vectors for Drosophila melanogaster P-element-mediated transformation and tissue culture transfection. Gene *74*, 771-784.

Troemel, E. R., Chou, J. H., Dwyer, N. D., Colbert, H. A., and Bargmann, C. I. (1995). Divergent seven transmembrane receptors are candidate chemosensory receptors in C. elegans. Cell *83*, 207-218.

Troemel, E. R., Kimmel, B. E., and Bargmann, C. I. (1997). Reprogramming chemotaxis responses: sensory neurons define olfactory preferences in C. elegans. Cell *91*, 161-169.

Vassar, R., Chao, S. K., Sitcheran, R., Nuñez, J. M., Vosshall, L. B., and Axel, R. (1994). Topographic organization of sensory projections to the olfactory bulb. Cell *79*, 981-991.

Vassar, R., Ngai, J., and Axel, R. (1993). Spatial segregation of odorant receptor expression in the mammalian olfactory epithelium. Cell *74*, 309-318.

Venkatesh, S., and Singh, R. N. (1984). Sensilla on the third antennal segment of Drosophila melanogaster Meigen (Diptera: Drosophilidae). Int. J. Insect Morphol. Embryol. *13*, 51-63.

Wang, F., Nemes, A., Mendelsohn, M., and Axel, R. (1998). Odorant receptors govern the formation of a precise topographic map. Cell *93*, 47-60.

What is claimed is:

1. An isolated nucleic acid molecule encoding an insect odorant receptor.

2. An isolated DNA, cDNA, genomic DNA, synthetic DNA or RNA of claim 1.

3. An isolated nucleic acid molecule of claim 1, wherein the nucleic acid molecule encodes a Drosophila odorant receptor.

4. The isolated nucleic acid molecule of claim 3 which comprise:
   (a) one of the nucleic acid sequences as set forth in Figure 8,
   (b) a sequence being degenerated to a sequence of (a) as a result of the genetic code; or
   (c) a sequence encoding one of the amino acid sequences as set forth in Figure 8.

5. A nucleic acid molecule of at least 12 nucleotides capable of specifically hybridizing with the sequence of a nucleic acid molecule of claim 1.

6. A DNA, cDNA, genomic DNA, synthetic DNA or RNA of claim 5.

7. A vector which comprises the isolated nucleic acid molecule of claim 1, or 5.

8. An isolated nucleic acid molecule of claim 7 operatively linked to a regulatory element.

9. A plasmid of claim 8.

10. A host vector system for the production of a polypeptide having the biological activity of an insect

odorant receptor which comprises the vector of claim 7 and a suitable host.

11. A host vector system of claim 10, wherein the suitable host is a bacterial cell, yeast cell, insect cell, or animal cell.

12. A method of producing a polypeptide having the biological activity of an insect odorant receptor which comprising growing the host vector system of claim 11 under conditions permitting production of the polypeptide and recovering the polypeptide so produced.

13. A purified, insect odorant receptor.

14. A polypeptide encoded by the isolated nucleic acid molecule of claim 1.

15. An antibody capable of specifically binding to an insect odorant receptor.

16. An antibody capable of competitively inhibiting the binding of the antibody of claim 15.

17. A monoclonal antibody of claim 15 or 16.

18. A method for identifying cDNA inserts encoding an insect odorant receptors comprising:
   (a) generating a cDNA library which contains clones carrying cDNA inserts from antennal or maxillary palp sensory neurons;
   (b) hybridizing nucleic acid molecules of the clones from the cDNA libraries generated in step (a) with probes prepared from the antenna or maxillary palp neurons and probes from heads lacking antenna or maxillary palp neurons or from virgin female body tissue;

(c) selecting clones which hybridized with probes from the antenna or maxillary palp neurons but not from head lacking antenna or maxillary palp neurons or virgin female body tissue; and

(d) isolating clones which carry the hybridized inserts, thereby identifying the inserts encoding odorant receptors.

19. A method of claim 18, after step (c), further comprising:

(a) amplifying the inserts from the selected clones by polymerase chain reaction;

(b) hybridizing the amplified inserts with probes from the antennal or maxillary palp neurons; and

(c) isolating the clones which carry the hybridized inserts, thereby identifying the inserts encoding the odorant receptors.

20. A method of claim 19, wherein the probes are cDNA probes.

21. The cDNA inserts identified by the method of claim 18 or 19.

22. A method for identifying DNA inserts encoding an insect odorant receptors comprising:

(a) generating DNA libraries which contain clones carrying inserts from a sample which contains at least one antennal or maxillary palp neuron;

(b) contacting clones from the cDNA libraries generated in step (a) with nucleic acid molecule of claim 5 in appropriate conditions permitting the hybridization of the nucleic acid molecules of the clones and the nucleic acid molecule;

(c) selecting clones which hybridized with the nucleic acid molecule; and

(d) isolating the clones which carry the hybridized inserts, thereby identifying the inserts encoding the odorant receptors.

23. A method to identify DNA inserts encoding an insect odorant receptors comprising:
(a) generating DNA libraries which contain clones with inserts from a sample which contains at least one antenna or maxillary palp sensory neuron;
(b) contacting the clones from the DNA libraries generated in step (a) with appropriate polymerase chain reaction primers capable of specifically binding to nucleic acid molecules encoding odorant receptors in appropriate conditions permitting the amplification of the hybridized inserts by polymerase chain reaction;
(c) selecting the amplified inserts; and
(d) isolating the amplified inserts, thereby identifying the inserts encoding the odorant receptors.

24. A method to isolate DNA molecules encoding insect odorant receptors comprising:
(a) contacting a biological sample known to contain nucleic acids with appropriate polymerase chain reaction primers capable of specifically binding to nucleic acid molecules encoding insect odorant receptors in appropriate conditions permitting the amplification of the hybridized molecules by polymerase chain reaction;
(b) isolating the amplified molecules, thereby identifying the DNA molecules encoding the insect odorant receptors.

25. A method of transforming cells which comprises transfecting a host cell with a suitable vector of claim 7.

26. Transformed cells produced by the method of claim 25.

27. The transformed cells of claim 26, wherein the host cells are not usually expressing odorant receptors.

28. The transformed cells of claim 26, wherein the host cells are expressing odorant receptors.

29. A method of identifying a compound capable of specifically bind to an insect odorant receptor which comprises contacting a transfected cells or membrane fractions of the transfected cells of claim 26 with an appropriate amount of the compound under conditions permitting binding of the compound to such receptor, detecting the presence of any such compound specifically bound to the receptor, and thereby determining whether the compound specifically binds to the receptor.

30. A method of identifying a compound capable of specifically bind to an insect odorant receptor which comprises contacting an appropriate amount of the purified odorant receptor of claim 13 with an appropriate amount of the compound under conditions permitting binding of the compound to such purified receptor, detecting the presence of any such compound specifically bound to the receptor, and thereby determining whether the compound specifically binds to the receptor.

31. A method of claim 30, wherein the purified receptor is embedded in a lipid bilayer.

32. A method of identifying a compound capable of activating the activity of an insect odorant receptor which comprises contacting the transfected cells or membrane fractions of the transfected cells of claim 26 with the compound under conditions permitting the

activation of a functional odorant receptor response, the activation of the receptor indicating that the compound is capable of activating the activity of a odorant receptor.

33. A method of identifying a compound capable of activating the activity of an odorant receptor which comprises contacting a purified odorant receptor of claim 13 with the compound under conditions permitting the activation of a functional odorant receptor response, the activation of the receptor indicating that the compound is capable of activating the activity of a odorant receptor.

34. A method of claim 33, wherein the purified receptor is embedded in a lipid bilayer.

35. A method of identifying a compound capable of inhibiting the activity of a odorant receptor which comprises contacting the transfected cells or membrane fractions of the transfected cells of claims 26 with an appropriate amount of the compound under conditions permitting the inhibition of a functional odorant receptor response, the inhibition of the receptor response indicating that the compound is capable of inhibiting the activity of a odorant receptor.

36. A method of identifying a compound capable of inhibiting the activity of a odorant receptor which comprises contacting an appropriate amount of the purified odorant receptor of claim 13 with an appropriated amount of the compound under conditions permitting the inhibition of a functional odorant receptor response, the inhibition of the receptor response indicating that the compound is capable of activating the activity of a odorant receptor.

37. A method of claim 30, wherein the purified receptor is embedded in a lipid bilayer.

38. A method of claim 29, 30, 31, 32, 33, 34, 35, 36, or 37 wherein the compound is not previously known.

39. The compound identified by the method of claim 38.

40. A method of controlling pest populations which comprises identifying odorant ligands by the method of claim 29, 30, 31, 32, 33, 34, 35, 36, or 37 which are alarm odorant ligands and spraying the desired area with the identified odorant ligands.

41. A method of controlling a pest population which comprises identifying odorant ligands by the method of claim 29, 30, 31, 32, 33, 34, 35, 36, or 37 which interfere with the interaction between the odorant ligands and the odorant receptors which are associated with fertility.

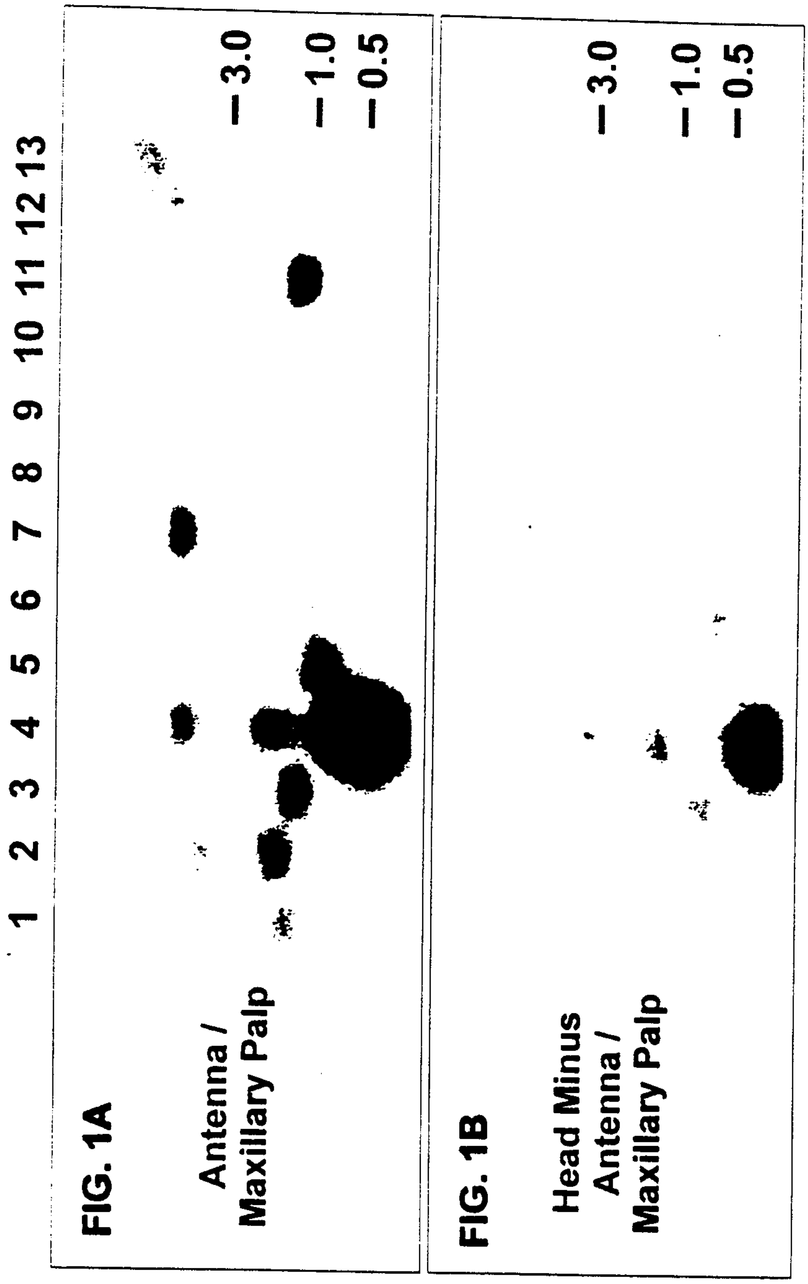

1 2 3 4 5 6 7 8 9 10 11 12 13
Antenna / Maxillary Palp
Head Minus Antenna / Maxillary Palp
—3.0
—1.0
—0.5

FIG. 1A

FIG. 1B

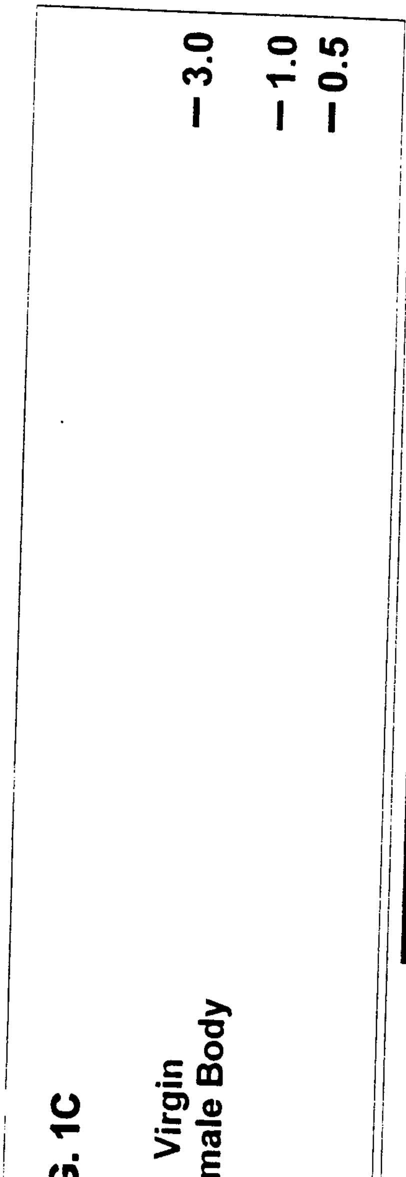


FIG. 1C

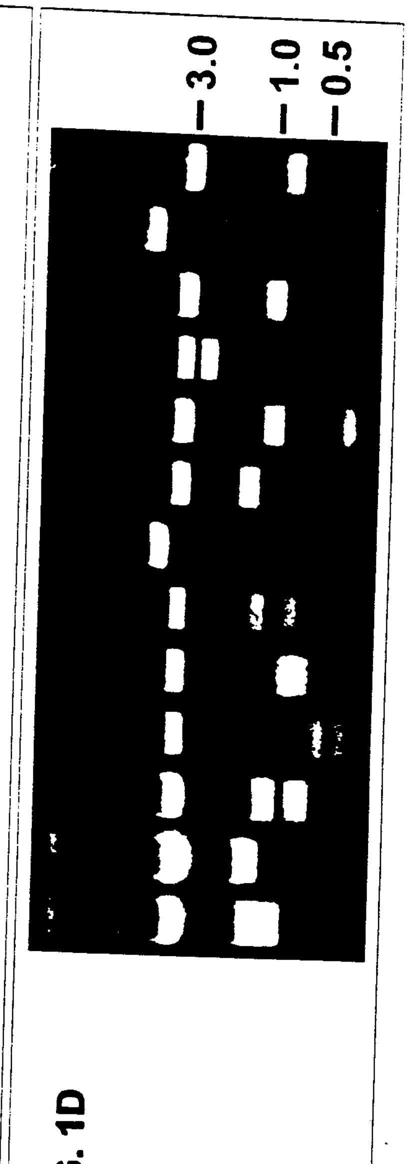


FIG. 1D

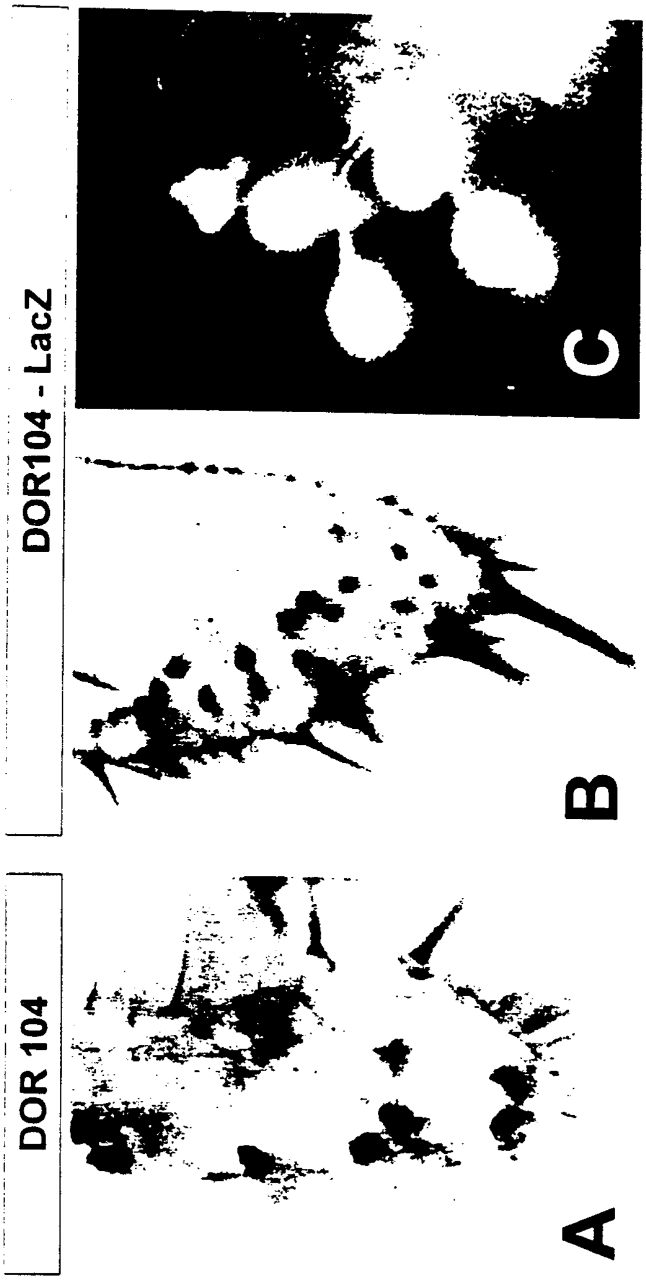


FIG. 2A

FIG. 2B

FIG. 2C

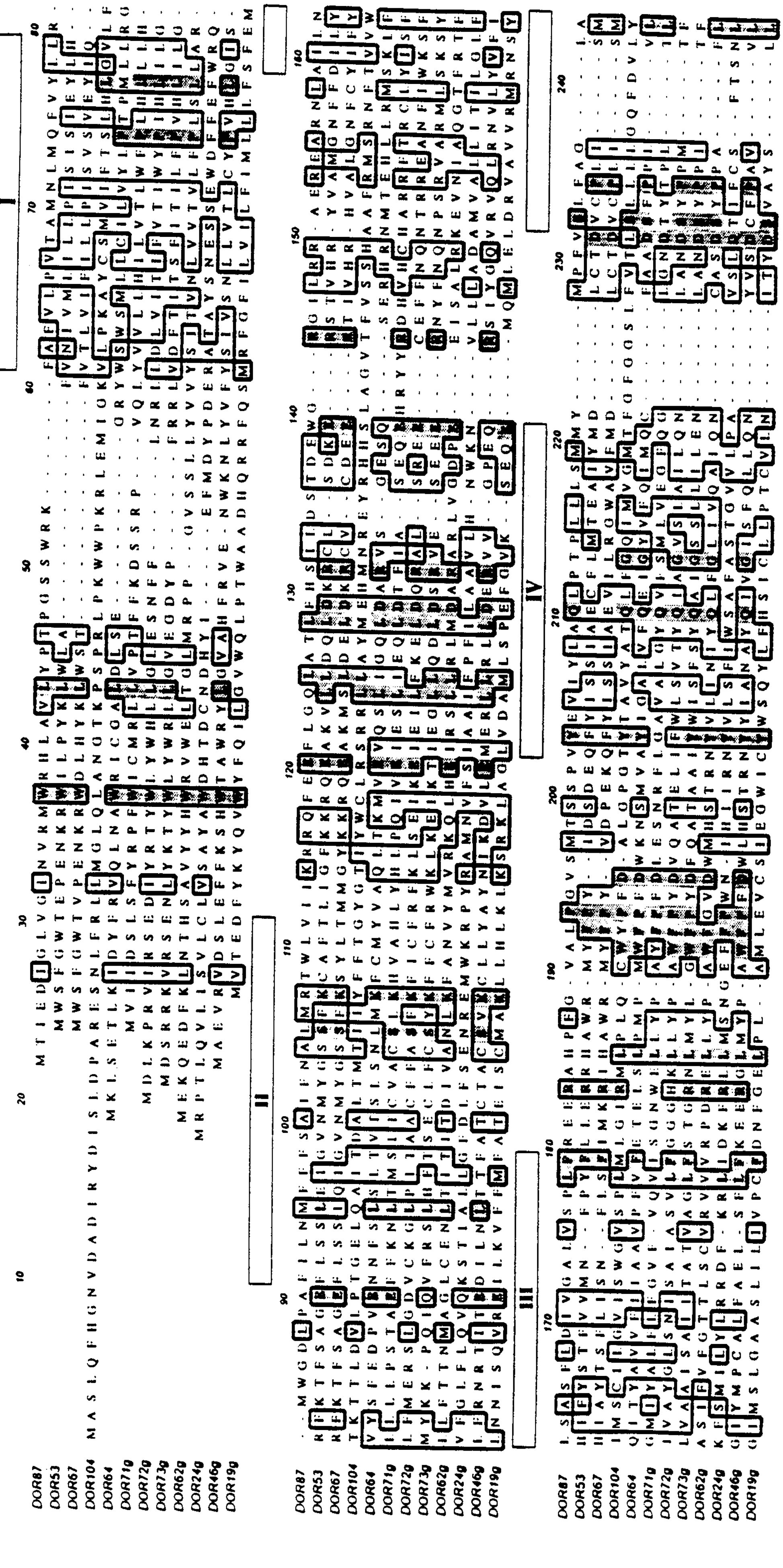
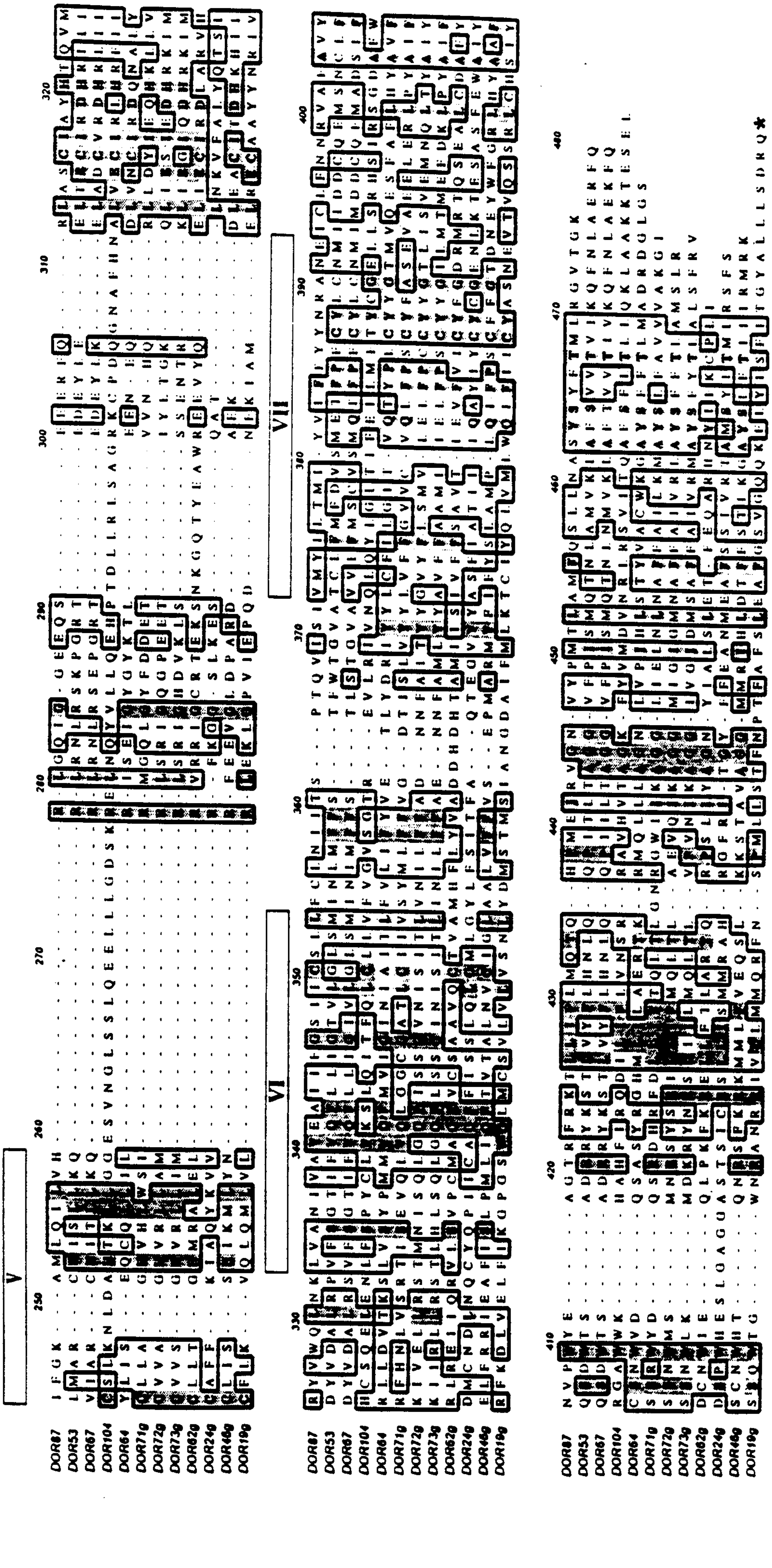


FIG. 3A

FIG. 3B


V
VI
VII
DOR87
DOR53
DOR67
DOR104
DOR64
DOR71g
DOR72g
DOR73g
DOR62g
DOR24g
DOR46g
DOR19g

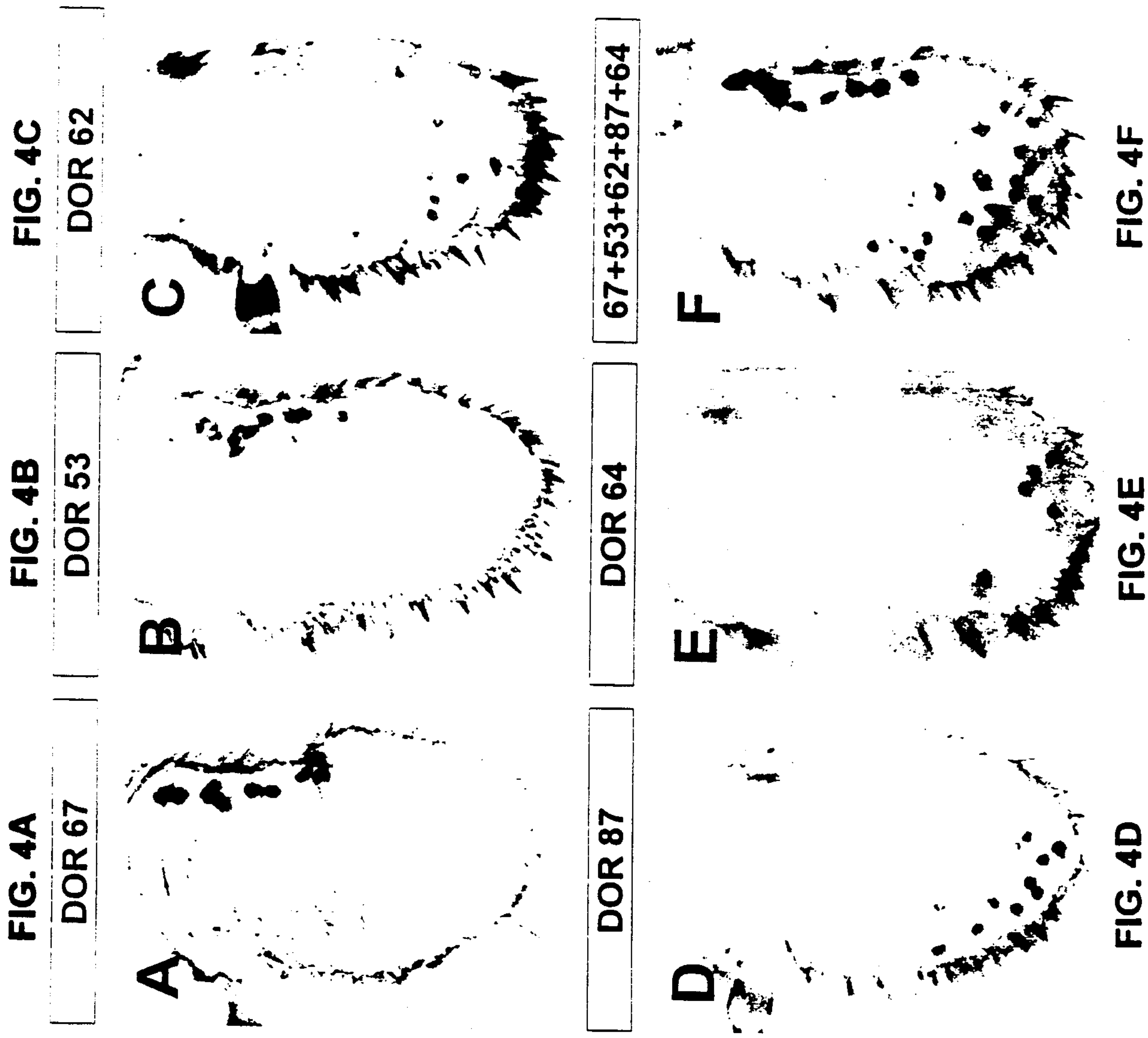


FIG. 4A FIG. 4B FIG. 4C

FIG. 4D FIG. 4E FIG. 4F

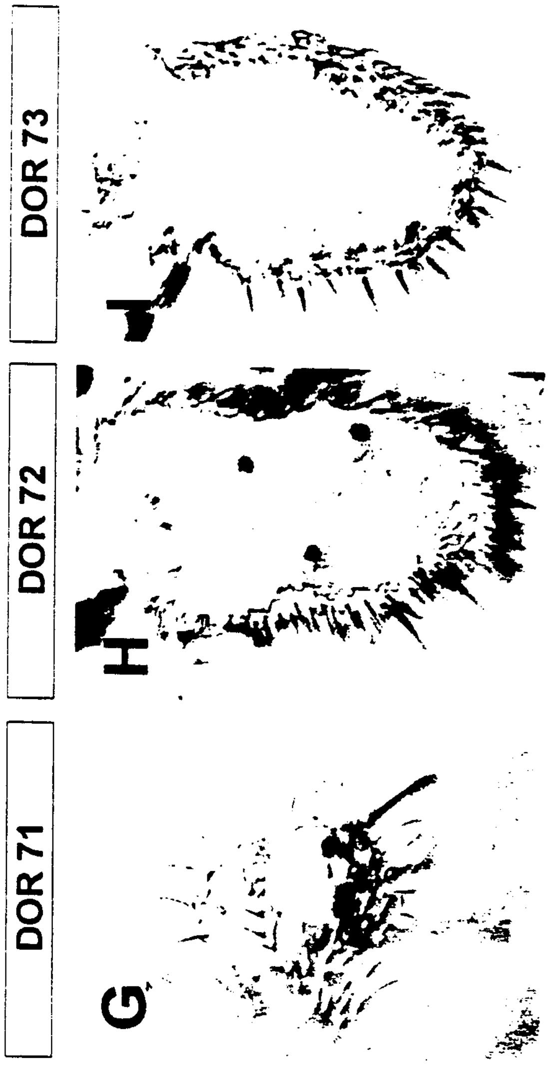


FIG. 4G FIG. 4H FIG. 4I

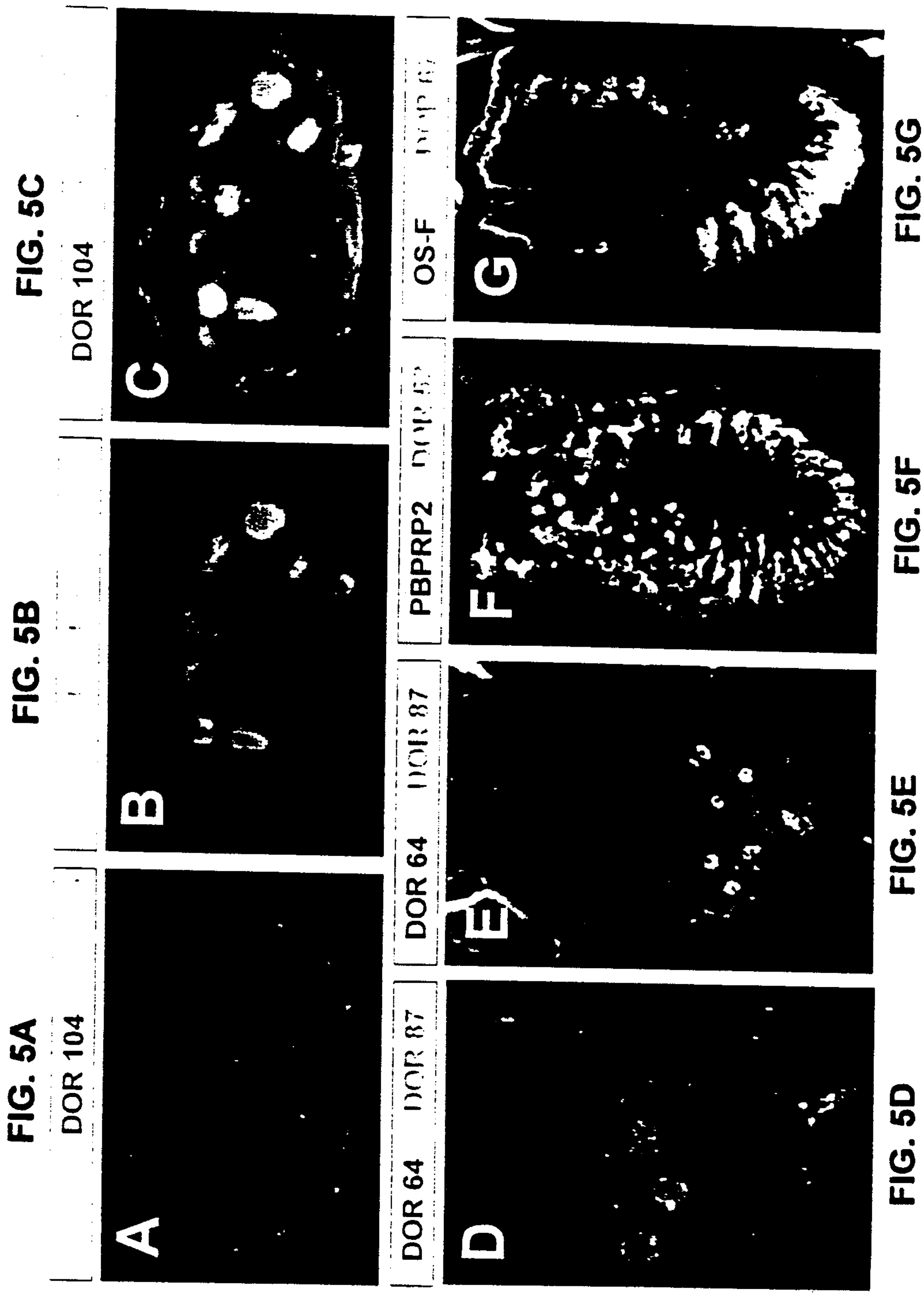


FIG. 5A FIG. 5B FIG. 5C

FIG. 5D FIG. 5E FIG. 5F FIG. 5G

FIG. 6A FIG. 6B FIG. 6C

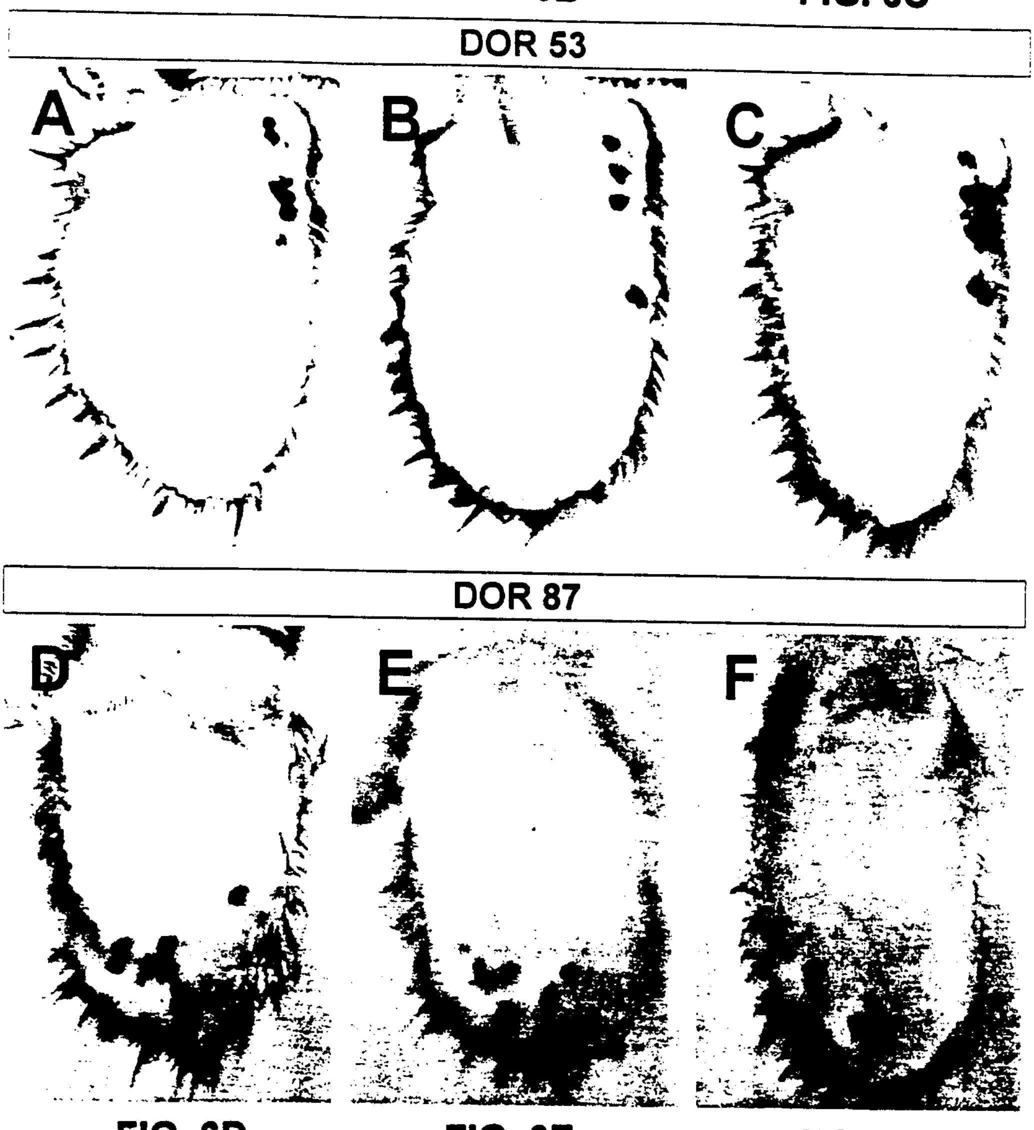


FIG. 6D FIG. 6E FIG. 6F

FIG. 7A

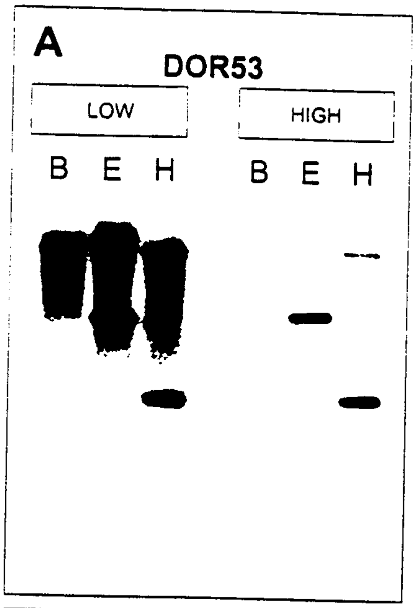
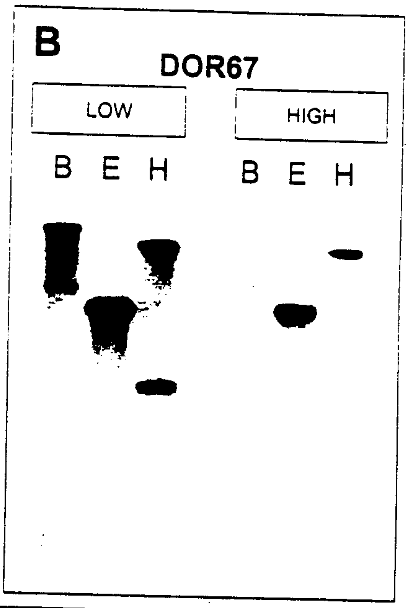


FIG. 7B



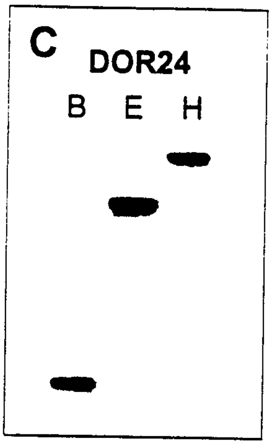


FIG. 7C

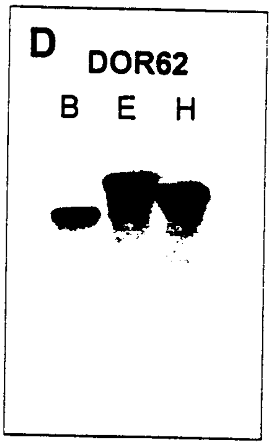


FIG. 7D

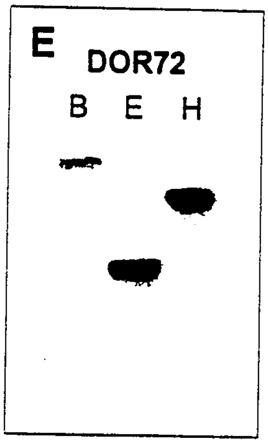


FIG. 7E

FIG. 8A

DOR62 [AF127921]

```
LOCUS       XXXXX          553 bp    mRNA            INV       02-FEB-1999
DEFINITION  Drosophila melanogaster putative odorant receptor DOR62 mRNA,
            partial cds.
ACCESSION   AF127921
KEYWORDS    .
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;
            Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 553)
  AUTHORS   Leslie B. Vosshall,Hubert Amrein, Pavel S. Morozov, Andrey
            Rzhetsky, and Richard Axel.
  TITLE     A Spatial Map of Olfactory Receptor Expression in the Drosophila
            Antenna
  JOURNAL   Cell (1999) In press
REFERENCE   2  (bases 1 to 553)
  AUTHORS   Vosshall,L.B., Amrein,H., Morozov,P.S., Rzhetsky,A. and
            Axel,Richard.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-FEB-1999) Department of Biochemistry and Molecular
            Biophysics, HHMI-Columbia University, 701 West 168th Street, New
            York, NY 10032, USA
FEATURES             Location/Qualifiers
     Comment:       /product='odorant receptor DOR62'
                    /tissue_type='adult antenna'
                    /note='putative seven transmembrane G protein coupled
                    receptor; genomic DNA sequenced by the European Drosophila
                    Genome Project and given accession number 62D9.q'
     source          1..553
                     /organism="Drosophila melanogaster"
                     /strain="Oregon R"
                     /db_xref="taxon:7227"
                     /chromosome="X"
                     /map="2F"
     gene            <1..553
                     /note="Drosophila melanogaster odorant receptor DOR62"
                     /gene="DOR62"
     CDS             <1..459
                     /gene="DOR62"
                     /note="odorant receptor"
                     /codon_start=1
                     /product="DOR62"

/translation="QELIECIRDLARVHRLREIIQRVLSVPCMAQFVCSAAVQCTVAM
HFLYVADDHDHTAMIISIVFFSAVTLEVFVICYFGDRMRTQSEALCDAFYDCNWIEQL
PKFKRELLFTLARTQRPSLIYAGNYIALSLETFEQVMRFTYSVFTLLLRAK"
BASE COUNT      141 a    145 c    132 g    135 t
ORIGIN
        1 caggaactca tcgagtgcat ccgcgatctg gcgcgggtcc atcggctgag ggagatcatt
       61 cagcgggtcc tttcagtgcc ctgcatggcc cagttcgtct gctccgccgc cgtccagtgt
      121 accgtcgcca tgcacttcct gtacgtagcg gatgaccacg accacaccgc catgatcatc
      181 tcgattgtat tttttctcggc cgtcaccttg gaggtgtttg taatctgcta ttttggggac
      241 aggatgcgga cacagagcga ggcgctgtgc gatgccttct acgattgcaa ctggatagaa
      301 cagctgccca agttcaagcg cgaactgctc ttcaccctgg ccaggacgca gcggcctttct
      361 cttatttacg caggcaacta catcgcactc tcgctggaga cctttcgagca ggtcatgagg
      421 ttcacatact ctgttttcac actcttgctg agggccaagt aagaacttta taatctcttt
      481 ttggggagaa aaattttaaa gcacaatagc agaaaaatat atcagataat ataacaaaaa
      541 aaaaaaaaaa aaa
```

FIG. 8B-1

DOR104[AF127922]

```
LOCUS       XXXXX        1493 bp    mRNA              INV       02-FEB-1999
DEFINITION  Drosophila melanogaster odorant receptor DOR104 mRNA, complete
cds.
ACCESSION   AF127922
KEYWORDS    .
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;
            Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1493)
  AUTHORS   Leslie B. Vosshall, Hubert Amrein, Pavel S. Morozov, Andrey
            Rzhetsky and Richard Axel.
  TITLE     A Spatial Map of Olfactory Receptor Expression in the Drosophila
            Antenna
  JOURNAL   Cell (1999) In press
REFERENCE   2  (bases 1 to 1493)
  AUTHORS   Vosshall,L.B., Amrein,H., Morozov,P.S., Rzhetsky,A. and Axel,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-FEB-1999) Department of Biochemistry and Molecular
            Biophysics, HHMI-Columbia University, 701 W. 168th Street, New
            York, NY 10025, USA
FEATURES             Location/Qualifiers
     Comment:       /tissue_type='adult maxillary palp'
                    /note='putative seven transmembrane G protein coupled
                    receptor'
     source          1..1493
                     /organism="Drosophila melanogaster"
                     /strain="Oregon R"
                     /db_xref="taxon:7227"
                     /chromosome="III"
                     /map="85B"
     gene            1..1493
                     /note="odorant receptor"
                     /gene="DOR104"
     CDS             21..1421
                     /gene="DOR104"
                     /note="odorant receptor"
                     /codon_start=1
                     /product="DOR104"

/translation="MASLQFHGNVDADIRYDISLDPARESNLFRLLMGLQLANGTKPS
PRLPKWWPKRLEMIGKVLPKAYCSMVIFTSLHLGVLFTKTTLDVLPTGELQAITDALT
MTIIYFFTGYGTIYWCLRSRRLLAYMEHMNREYRHHSLAGVTFVSSHAAFRMSRNFTV
VWIMSCLLGVISWGVSPLMLGIRMLPLQCWYPFDALGPGTYTAVYATQLFGQIMVGMT
FGFGGSLFVTLSLLLLGQFDVLYCSLKNLDAHTKLLGGESVNGLSSLQEELLLGDSKR
ELNQYVLLQEHPTDLLRLSAGRKCPDQGNAFHNALVECIRLHRFILHCSQELENLFSP
YCLVKSLQITFQLCLLVFVGVSGTREVLRIVNQLQYLGLTIFELLMFTYCGELLSRHS
IRSGDAFWRGAWWKHAHFIRQDILIFLVNSRRAVHVTAGKFYVMDVNRLRSVITQAFS
FLTLLQKLAAKKTESEL"
BASE COUNT      332 a    373 c    407 g    381 t
```

FIG. 8B-2

ORIGIN

```
   1 ggcacgagca gtcgatggcc agtcttcagt tccacggcaa cgtcgatgcg gacatcaggt
  61 atgatattag cctggatccg gctagggaat cgaatctctt ccgtctgcta atgggactcc
 121 agttggcgaa tggcacgaag ccatcgccgc ggttacccaa atggtggcca aagcggctgg
 181 aaatgattgg taaagtgctg cccaaagcct attgttccat ggtgattttc acctcccctgc
 241 atttgggtgt cctgttcacg aaaaccacac tggatgtcct gccgacgggg gagctgcagg
 301 ccataacgga tgccctcacc atgaccataa tatacttttt cacgggctac ggcaccatct
 361 actggtgcct gcgctcccgg cgcctcttgg cctacatgga gcacatgaac cgggagtatc
 421 gccatcattc gctggccggg gtgacctttg tgagtagcca tgcggccttt aggatgtcca
 481 gaaacttcac ggtggtgtgg ataatgtcct gcctgctggg cgtgatttcc tggggcgttt
 541 cgccactgat gctgggcatc cggatgctgc cgctccaatg ttggtatccc ttcgacgccc
 601 tgggtcccgg cacatatacg gcggtctatg ctacacaact tttcggtcag atcatggtgg
 661 gcatgacctt tggattcggg ggatcactgt ttgtcaccct gagcctgcta ctcctgggac
 721 aattcgatgt gctctactgc agcctgaaga acctggatgc ccataccaag ttgctgggcg
 781 gggagtctgt aaatggcctg agttcgctgc aagaggagtt gctgctgggg gactcgaaga
 841 gggaattaaa tcagtacgtt ttgctccagg agcatccgac ggatctgctg agattgtcgg
 901 caggacgaaa atgtcctgac caaggaaatg cgtttcacaa cgccttggtg gaatgcattc
 961 gcttgcatcg cttcattctg cactgctcac aggagttgga gaatctattc agtccatatt
1021 gtctggtcaa gtcactgcag atcacctttc agctttgcct gctggtcttt gtgggcgttt
1081 cgggtactcg agaggtcctg cggattgtca accagctaca gtacttggga ctgaccatct
1141 tcgagctcct aatgttcacc tattgtggcg aactcctcag tcggcatagt attcgatctg
1201 gcgacgcctt ttggaggggt gcgtggtgga agcacgccca tttcatccgc caggacatcc
1261 tcatctttct ggtcaatagt agacgtgcag ttcacgtgac tgccggcaag ttttatgtga
1321 tggatgtgaa tcgtctaaga tcggttataa cgcaggcgtt cagcttcttg actttgctgc
1381 aaaagttggc tgccaagaag acggaatcgg agctctaaac tggtaccacg catcgatatt
1441 tatttagcgc attaaaaaaa agtcgagtaa aagcaaaaaa aaaaaaaaaa aaa
```

FIG. 8C-1

DOR87 [AF127926]

```
LOCUS       XXXXX        1556 bp    mRNA            INV       02-FEB-1999
DEFINITION  Drosophila melanogaster odorant receptor DOR87 mRNA, complete
cds.
ACCESSION   AF127926
KEYWORDS    .
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;
            Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1556)
  AUTHORS   Leslie B. Vosshall, Hubert Amrein, Pavel S. Morozov, Andrey
            Rzhetsky and Richard Axel.
  TITLE     A Spatial Map of Olfactory Receptor Expression in the Drosophila
            Antenna
  JOURNAL   Cell (1999) In press
REFERENCE   2  (bases 1 to 1556)
  AUTHORS   Vosshall,L.B., Amrein,H., Morozov,P.S., Rzhetsky,A. and Axel,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-FEB-1999) Department of Biochemistry and Molecular
            Biophysics, HHMI-Columbia University, 701 West 168th Street, New
            York, NY 10032, USA
FEATURES             Location/Qualifiers
     Comment:      /tissue='adult antenna'
                   /note='putative seven transmembrane G protein coupled
                   receptor'
     source          1..1556
                     /organism="Drosophila melanogaster"


                     /strain="Oregon R"
                     /db_xref="taxon:7227"
                     /chromosome="II"
                     /map="43B1-2"
     gene            1..1556
                     /note="odorant receptor"
                     /gene="DOR87"
     CDS             30..1157
                     /gene="DOR87"
                     /note="odorant receptor"
                     /codon_start=1
                     /product="DOR87"

/translation="MTIEDIGLVGINVRMWRHLAVLYPTPGSSWRKFAFVLPVTAMNL
MQFVYLLRMWGDLPAFILNMFFFSAIFNALMRTWLVIIKRRQFEEFLGQLATLFHSIL
DSTDEWGRGILRRAEREARNLAILNLSASFLDIVGALVSPLFREERAHPFGVALPGVS
MTSSPVYEVIYLAQLPTPLLLSMMYMPFVSLFAGLAIFGKAMLQILVHRLGQIGGEEQ
SEEERFQRLASCIAYHTQVMRYVWQLNKLVANIVAVEAIIFGSIICSLLFCLNIITSP
TQVISIVMYILTMLYVLFTYYNRANEICLENNRVAEAVYNVPWYEAGTRFRKTLLIFL
MQTQHPMEIRVGNVYPMTLAMFQSLLNASYSYFTMLRGVTGK"
BASE COUNT      329 a     413 c     408 g     406 t
```

FIG. 8C-2

```
ORIGIN
        1 ggcacgaggc ttatagaaag tgccgagcaa tgacaatcga ggatatcggc ctggtgggca
       61 tcaacgtgcg gatgtggcga cacttggccg tgctgtaccc cactccgggc tccagctggc
      121 gcaagttcgc cttcgtgctg ccggtgactg cgatgaatct gatgcagttc gtctacctgc
      181 tgcggatgtg ggcgacctg cccgccttca ttctgaacat gttcttcttc tcggccattt
      241 tcaacgccct gatgcgcacg tggctggtca taatcaagcg gcgccagttc gaggagtttc
      301 tcggccaact ggccactctg ttccattcga ttctcgactc caccgacgag tgggggcgtg
      361 gcatcctgcg gagggcggaa cgggaggctc ggaacctggc catccttaat ttgagtgcct
      421 ccttcctgga cattgtcggt gctctggtat cgccgctttt cagggaggag agagctcatc
      481 ccttcggcgt agctctacca ggagtgagca tgaccagttc acccgtctac gaggttatct
      541 acttggccca actgcctacg cccctgctgc tgtccatgat gtacatgcct ttcgtcagcc
      601 tttttgccgg cctggccatc tttgggaagg ccatgctgca gatcctggta cacaggctgg
      661 gccagattgg cggagaagag cagtcggagg aggagcgctt ccaaaggctg gcctcctgca
      721 ttgcgtacca cacgcaggtg atgcgctatg tgtggcagct caacaaactg gtggccaaca
      781 ttgtggcggt ggaagcaatt atttttggct cgataatctg ctcactgctc ttctgtctga
      841 atattataac ctcacccacc caggtgatct cgatagtgat gtacattctg accatgctgt
      901 acgttctctt cacctactac aatcgggcca atgaaatatg cctcgagaac aaccgggtgg
      961 cggaggctgt ttacaatgtg ccctggtacg aggcaggaac tcggtttcgc aaaaccctcc
     1021 tgatcttctt gatgcaaaca caacacccga tggagataag agtcggcaac gtttacccca
     1081 tgacattggc catgttccag agtctgttga atgcgtccta ctcctacttt accatgctgc
     1141 gtggcgtcac cggcaaatga gctgaaagac cgaaaaaacc ggagtatccc cttccatatt
     1201 ccccctgctc ctttattttc ctttcctttt ccctttccgt tttcccattc gcttttccag
     1261 caatccgggt aatgcaaaaa gttgttgctg gctgtggtcc tggctgcttg tttggcattt
     1321 gcatatgctt gtcgtttgaa aggatttaat cggactgctg gcacggagtc ggcatcctgg
     1381 ctcctggatc ctggcatgca aatagttggc ttcttagatt gttacacaaa atagattgta
     1441 gattgcagct gaatgttgtg cttggaataa agtcaaaagg atgtggagtc ggcccaaggc
     1501 tctgcccatt ctgtttgctc gggatgcccg aaagtatgaa aaaaaaaaaa aaaaaa
```

FIG. 8D-1

DOR53 [AF127923]

```
LOCUS       XXXXX        1305 bp    mRNA              INV       06-FEB-1999
DEFINITION  Drosophila melanogaster putative odorant receptor DOR53 mRNA,
            complete cds.
ACCESSION   AF127923
KEYWORDS    .
SOURCE      fruit fly.

  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;
            Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1305)
  AUTHORS   Leslie B. Vosshall, Hubert Amrein, Pavel S. Morozov, Andrey
            Rzhetsky and Richard Axel.
  TITLE     A Spatial Map of Olfactory Receptor Expression in the Drosophila
            Antenna
  JOURNAL   Cell (1999) In press
REFERENCE   2  (bases 1 to 1305)
  AUTHORS   Vosshall,L.B., Amrein,H., Morozov,P.S., Rzhetsky,A. and Axel,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-FEB-1999) Department of Biochemistry and Molecular
            Biophysics, HHMI-Columbia University, 701 West 168th Street, New
            York, NY 10032, USA
FEATURES             Location/Qualifiers
     source          1..1305
                     /organism="Drosophila melanogaster"
                     /strain="Oregon R"
                     /db_xref="taxon:7227"
                  /note="_tissue_type='adult antenna'  _note='putative seven
                     transmembrane G protein coupled receptor';"
                     /chromosome="II"
                     /map="22A2-3"
     gene            1..1305
                     /note="odorant receptor"
                     /gene="DOR53"
     CDS             79..1179
                     /gene="DOR53"
                     /note="odorant receptor"
                     /codon_start=1
                     /product="DOR53"

/translation="MWSFGWTEPENKRWILPYKLWLAFVNIVMLILLPISISIEYLHR
FKTFSAGEFLSSLEIGVNMYGSSFKCAFTLIGFKKRQEAKVLLDQLDKRCLSDKERST
VHRYVAMGNFFDILYHIFYSTFVVMNFPYFLLERRHAWRMYFPYIDSDEQFYISSIAE
CFLMTEAIYMDLCTDVCPLISMLMARCHISLLKQRLRNLRSKPGRTEDEYLEELTECI
RDHRLLLDYVDALRPVFSGTIFVQFLLIGTVLGLSMINLMFFSTFWTGVATCLFMFDV
SMETFPFCYLCNMIIDDCQEMSNCLFQSDWTSADRRYKSTLVYFLHNLQQPITLTAGG
VFPISMQTNLAMVKLAFSVVTVIKQFNLAERFQ"
BASE COUNT       331 a     286 c     279 g     409 t
```

FIG. 8D-2

```
ORIGIN
        1 tttttcccc acataaaaga aaagccattg agcgagcggg ttaagtcccg agatgccttc
       61 atttacttgg atcgggtgat gtggtccttt ggctggacag agcctgaaaa caaaaggtgg
      121 atccttcctt ataaactgtg gttagcgttc gtgaacatag taatgctcat ccttctgccg
      181 atctcgataa gcatcgagta cctccaccga tttaaaacct tctcggcggg ggagttcctt
      241 agttccctcg agattggagt caacatgtac ggaagctctt ttaagtgcgc cttcaccttg
      301 attggattca agaaaagaca ggaagctaag gttttactgg atcagctgga caagagatgc
      361 cttagcgata aggagaggtc cactgttcat cgctatgtcg ccatgggaaa ctttttcgat
      421 attttgtatc acattttta ctccaccttc gtggtaatga acttcccgta ttttctgctt
      481 gagagacgcc atgcttggcg catgtacttt ccatatatcg attccgacga acagttttac
      541 atctccagca tcgccgagtg ttttctgatg acggaggcca tctacatgga tctctgtacg
      601 gacgtgtgtc ccttgatctc catgcttatg gctcgatgcc acatcagcct cctgaaacag
      661 cgactgagaa atctccgatc gaagccagga aggaccgaag atgagtactt ggaggagctc
      721 accgagtgca ttcgggatca tcgattgcta ttggactatg ttgacgcatt gcgacccgtc
      781 ttttcgggaa ccatttttgt gcagttcctc ctgatcggta ctgtactggg tctctcaatg
      841 ataaatctaa tgttcttctc gacattttgg actggtgtcg ccacttgcct tttatgttc
      901 gacgtgtcca tggagacgtt ccccttttgc tatttgtgca acatgattat cgatgactgc
      961 caggaaatgt ccaattgcct ctttcaatcg gactggacct ctgccgatcg tcgctacaaa
     1021 tccactttgg tatactttct tcacaatctt cagcaaccca ttactctcac ggctggtgga
     1081 gtgtttccta tttccatgca aacaaatttg gctatggtga agctggcatt ttctgtggtt
     1141 acggtaatta agcaatttaa cttggccgaa aggtttcaat aagttgagag ggacgagctc
     1201 tgctactatt atattatata ttatattata ttatatatat attattttat attatatatt
     1261 gctgtaccct aataaatatt tagtaataaa aaaaaaaaaa aaaaa
```

FIG. 8E-1

DOR67 [AF127924]

```
LOCUS       XXXXX          1321 bp    mRNA              INV       06-FEB-1999
DEFINITION  Drosophila melanogaster putative odorant receptor DOR67 mRNA,
            complete cds.
ACCESSION   AF127924
KEYWORDS    .
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;
            Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1321)
  AUTHORS   Leslie B. Vosshall, Hubert Amrein, Pavel S. Morozov, Andrey
            Rzhetsky and Richard Axel.
  TITLE     A Spatial Map of Olfactory Receptor Expression in the Drosophila
            Antenna
  JOURNAL   Cell (1999) In press
REFERENCE   2  (bases 1 to 1321)
  AUTHORS   Vosshall,L.B., Amrein,H., Morozov,P.S., Rzhetsky,A. and Axel,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-FEB-1999) Department of Biochemistry and Molecular
            Biophysics, HHMI-Columbia University, 701 West 168th Street, New
            York, NY 10032, USA
FEATURES             Location/Qualifiers
     Comment:      /tissue_type='adult antenna'
                   /note='putative seven transmembrane G protein coupled
                   receptor'
     source          1..1321
                     /organism="Drosophila melanogaster"
                     /strain="Oregon R"
                     /db_xref="taxon:7227"
                     /chromosome="II"
                     /map="22A3"
     gene            1..1321
                     /note="odorant receptor"
                     /gene="DOR67"
     CDS             102..1202
                     /gene="DOR67"
                     /note="odorant receptor"
                     /codon_start=1
                     /product="DOR67"

/translation="MWSFGWTVPENKRWDLHYKLWSTFVTLVIFILLPISVSVEYIQR
FKTFSAGEFLSSIQIGVNMYGSSFKSYLTMMGYKKRQEAKMSLDELDKRCVCDEERTI
VHRHVALGNFCYIFYHIAYTSFLISNFLSFIMKRIHAWRMYFPYVDPEKQFYISSIAE
VILRGWAVFMDLCTDVCPLISMVIARCHITLLKQRLRNLRSEPGRTEDEYLKELADCV
RDHRLILDYVDALRSVFSGTIFVQFLLIGIVLGLSMINIMFFSTLSTGVAVVLFMSCV
SMQTFPFCYLCNMIMDDCQEMADSLFQSDWTSADRRYKSTLVYFLHNLQQPIILTAGG
VFPISMQTNLNMVKLAFTVVTIVKQFNLAEKFQ"
BASE COUNT      348 a     281 c     294 g     398 t
```

FIG. 8E-2

```
ORIGIN
        1 ggcacgagga aatgttaagc cagttctttc cccacattaa agaaaagcca ttgagcgagc
       61 gggttaagtc ccgagatgcc ttcgtttact tagatcgggt gatgtggtcc tttggctgga
      121 cagtgcctga aaacaaaagg tgggatctac attacaaact gtggtcaact ttcgtgacat
      181 tggtgatatt tatccttctg ccgatatcgg taagcgttga gtatattcag cggttcaaga
      241 ccttctcggc gggtgagttt cttagctcaa tccagattgg cgttaacatg tacggaagca
      301 gctttaaaag ttatttgacc atgatgggat ataagaagag acaggaggct aagatgtcac
      361 tggatgagct ggacaagaga tgcgtttgtg atgaggagag gaccattgta catcgacatg
      421 tcgccctggg aaacttttgc tatattttct atcacattgc gtacactagc tttttgattt
      481 caaacttttt gtcatttata atgaagagaa tccatgcctg gcgcatgtac tttccctacg
      541 tcgaccccga aaagcaattt tacatctcta gcatcgccga agtcattctt aggggggtggg
      601 ccgtcttcat ggatctctgc acggatgtgt gtcctttgat ctccatggta atagcacgat
      661 gccacatcac ccttctgaaa cagcgcctgc gaaatctacg atcggaacca ggaaggacgg
      721 aagatgagta cttgaaggag ctcgccgact gcgttcgaga tcaccgcttg atattggact
      781 atgtcgacgc attgcgatcc gtcttttcgg ggacaatttt tgtgcagttc ctcttgatcg
      841 gtattgtact gggtctgtca atgataaata taatgttttt ctcaacactt tcgactggtg
      901 tcgccgttgt cctttttatg tcctgcgtat ctatgcagac gttccccttt tgctatttgt
      961 gtaacatgat tatggatgac tgccaagaga tggccgactc cctttttcaa tcggactgga
     1021 catctgccga tcgtcgctac aaatccactt tggtatactt tcttcacaat cttcagcagc
     1081 ccattattct tacggctggt ggagtctttc ctatttccat gcaaacaaat ttaaatatgg
     1141 tgaagctggc ctttactgtg gttacaatag tgaaacaatt taacttggca gaaaagtttc
     1201 aataagttaa gatatgcaag ctctgctatt ataaacctac actcgagaaa atatttcttc
     1261 acattaataa accttcagta cttactgctt gtggcgcccc cggaaaaaaa aaaaaaaaaa
     1321 a
```

FIG. 8F-1

DOR64 [AF127925]

```
LOCUS       XXXXX         1308 bp    mRNA              INV       06-FEB-1999
DEFINITION  Drosophila melanogaster putative odorant receptor DOR64 mRNA,
            complete cds.
ACCESSION   AF127925
KEYWORDS    .
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;
            Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1308)
  AUTHORS   Leslie B. Vosshall, Hubert Amrein, Pavel S. Morozov, Andrey
            Rzhetsky and Richard Axel.
  TITLE     A Spatial Map of Olfactory Receptor Expression in the Drosophila
            Antenna
  JOURNAL   Cell (1999) In press
REFERENCE   2  (bases 1 to 1308)
  AUTHORS   Vosshall,L.B., Amrein,H., Morozov,P.S., Rzhetsky,A. and Axel,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-FEB-1999) Department of Biochemistry and Molecular
            Biophysics, HHMI-Columbia University, 701 West 168th Street, New
            York, NY 10032, USA
FEATURES             Location/Qualifiers
     Comment:       /tissue_type='adult antenna'
                    /note='putative seven transmembrane G protein coupled
                    receptor'
     source            1..1308
                       /organism="Drosophila melanogaster"
                       /strain="Oregon R"
                       /db_xref="taxon:7227"
                       /chromosome="II"
                       /map="23A1-2"
     gene              1..1308
                       /note="odorant receptor"
                       /gene="DOR64"
     CDS               22..1158
                       /gene="DOR64"
                       /note="odorant receptor"
                       /codon_start=1
                       /product="DOR64"

/translation="MKLSETLKIDYFRVQLNAWRICGALDLSEGRYWSWSMLLCILVY
LPTPMLLRGVYSFEDPVENNFSLSLTVTSLSNLMKFCMYVAQLTKMVEVQSLIGQLDA
RVSGESQSERHRNMTEHLLRMSKLFQITYAVVFIIAAVPFVFETELSLPMPMWFPFDW
KNSMVAYIGALVFQEIGYVFQIMQCFAADSFPPLVLYLISEQCQLLILRISEIGYGYK
TLEENEQDLVNCIRDQNALYRLLDVTKSLVSYPMMVQFMVIGINIAITLFVLIFYVET
LYDRIYYLCFLLGITVQTYPLCYYGTMVQESFAELHYAVFCSNWVDQSASYRGHMLIL
AERTKRMQLLLAGNLVPIHLSTYVACWKGAYSFFTLMADRDGLGS"
BASE COUNT       331 a     314 c     305 g     358 t
```

FIG. 8F-2

```
ORIGIN
        1 ggcacgagcc aagaattcaa aatgaaactc agcgaaaccc taaaaatcga ctattttcga
       61 gtccagttga atgcctggcg aatttgtggt gccttggatc tcagcgaggg taggtactgg
      121 agttggtcga tgctattgtg catcttggtg tacctgccga cacccatgct actgagagga
      181 gtatacagtt tcgaggatcc ggtggaaaat aatttcagct tgagcctgac ggtcacatcg
      241 ctgtccaatc tcatgaagtt ctgcatgtac gtggcccaac taacaaagat ggtcgaggtc
      301 cagagtctta ttggtcagct ggatgcccgg gtttctggcg agagccagtc tgagcgtcat
      361 agaaatatga ccgagcacct gctaaggatg tccaagctgt tccagatcac ctacgctgta
      421 gtcttcatca ttgctgcagt tcccttcgtt ttcgaaactg agctaagctt acccatgccc
      481 atgtggtttc ccttcgactg gaagaactcg atggtggcct acatcggagc tctggttttc
      541 caggagattg gctatgtctt tcaaattatg caatgctttg cagctgactc gtttcccccg
      601 ctcgtactgt acctgatctc cgagcaatgt caattgctga tcctgagaat ctctgaaatc
      661 ggatatggtt acaagactct ggaggagaac gaacaggatc tggtcaactg catcagggat
      721 caaaacgcgc tgtatagatt actcgatgtg accaagagtc tcgtttcgta tcccatgatg
      781 gtgcagttta tggttattgg catcaacatc gccatcaccc tatttgtcct gatattttac
      841 gtggagacct tgtacgatcg catctattat ctttgctttc tcttgggcat caccgtgcag
      901 acatatccat tgtgctacta tggaaccatg gtgcaggaga gttttgctga gcttcactat
      961 gcggtattct gcagcaactg ggtggatcaa agtgccagct atcgtgggca catgctcatc
     1021 ctggcggagc gcactaagcg gatgcagctt ctcctcgccg gcaacctggt gcccatccac
     1081 ctgagcacct acgtggcctg ttggaaggga gcctactcct tcttcaccct gatggccgat
     1141 cgagatggcc tgggttctta gtagcccagt catttcactc acattctaca tcaagtagta
     1201 ctaccactga acacgaacac gaatatttca aaagtaaaca cataatattc acaatagtgt
     1261 atcactttaa taaaattttt ggttaccatg aaaaaaaaaa aaaaaaa
```

FIG. 8G-1

DOR71g

```
LOCUS       XXXXX        1252 bp    DNA             INV       07-FEB-1999
DEFINITION  Drosophila melanogaster odorant receptor DOR71g, GENSCAN
predicted complete cds.
ACCESSION   ;
KEYWORDS    .
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;
            Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1252)
  AUTHORS   Leslie B. Vosshall, Hubert Amrein, Pavel S. Morozov, Andrey
            Rzhetsky and Richard Axel.
  TITLE     A Spatial Map of Olfactory Receptor Expression in the Drosophila
            Antenna
  JOURNAL   Cell (1999) In press
REFERENCE   2  (bases 1 to 1252)
  AUTHORS   Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
            Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
            Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
            Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
            Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
            Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
            Pfeiffer,B., Poon,L., Punch,E., Sequeira,A., Sethi,H., Snir,E.,
            Svirskas,R.R., Twomey,B., Wan,K.H., Weinburg,T., Zhang,R.,
            Zieran,L.L. and Rubin,G.M.
  TITLE     Sequencing of Drosophila melanogaster
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 1252)
  AUTHORS   Vosshall,L.B., Amrein,H., Morozov,P.S., Rzhetsky,A. and Axel,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-FEB-1999) Department of Biochemistry and Molecular
            Biophysics, HHMI-Columbia University, 701 West 168th Street, New
            York, NY 10032, USA
FEATURES             Location/Qualifiers
     Comment:       /tissue_type='adult maxillary palp'
                    /note='putative seven transmembrane G protein coupled
                    receptor'
                    /note='sequence annotation for BDGP P1 clone DS07071;
                    predicted using GENSCAN; join (26150-24899)'
     source            1..1252
                       /organism="Drosophila melanogaster"
                       /db_xref="taxon:7227"
                       /chromosome="II"
                       /map="33B1-2"
     exon              1..731
                       /gene="DOR71g"
                       /note="exon"
     mRNA              join(1..731,811..1234)
                       /gene="DOR71g"
     CDS               join(1..731,811..1234)
                       /gene="DOR71g"
                       /note="odorant receptor"
                       /codon_start=1
                       /product="DOR71g"
```

FIG. 8G-2

```
/translation="MVIIDSLSFYRPFWICMRLLVPTFFKDSSRPVQLYVVLLHILVT
LWFPLHLLLHLLLLPSTAEFFKNLTMSLTCVACSLKHVAHLYHLPQIVEIESLIEQLD
TFIASEQEHRYYRDHVHCHARRFTRCLYISFGMIYALFLFGVFVQVISGNWELLYPAY
FPFDLESNRFLGAVALGYQVFSMLVEGFQGLGNDTYTPLTLCLLAGHVHLWSIRMGQL
GYFDDETVVNHQRLLDYIEQHKLLVRFHNLVSRTISEVQLVQLGGCGATLCIIVSYML
FFVGDTISLVYYLVFFGVVCVQLFPSCYFASEVAEELERLPYAIFSSRWYDQSRDHRF
DLLIFTQLTLGNRGWIIKAGGLIELNLNAFFATLKMAYSLFAVVVRAKGI"
     gene            1..1234
                     /note="odorant receptor"
                     /gene="DOR71g"
     intron          732..810
                     /gene="DOR71g"

                     /note="intron"
     exon            811..1231
                     /gene="DOR71g"
                     /note="exon"
     polyA_signal    1246..1252
                     /note="polyA_signal"
BASE COUNT      269 a    299 c    305 g    379 t
ORIGIN
        1 atggtcatta tcgacagtct tagtttttat cgtccattct ggatctgcat gcgattgctg
       61 gtaccgactt tcttcaagga ttcctcacgt cctgtccagc tgtacgtggt gttgctgcac
      121 atcctggtca ccttgtggtt tccactgcat ctgctgctgc atcttctgct acttccatct
      181 accgctgagt tctttaagaa cctgaccatg tctctgactt gtgtggcctg cagtctgaag
      241 catgtggccc acttgtatca cttgccgcag attgtggaaa tcgaatcact gatcgagcaa
      301 ttagacacat ttattgccag cgaacaggag catcgttact atcgggatca cgtacattgc
      361 catgctaggc gctttacaag atgtctctat attagctttg gcatgatcta tgcgcttttc
      421 ctgttcggcg tcttcgttca ggttattagc ggaaattggg aacttctcta tccagcctat
      481 ttccccattcg acttggagag caatcgcttt ctcggcgcag tagcccttggg ctatcaggta
      541 ttcagcatgt tagttgaagg cttccagggg ctgggcaacg atacctatac ccccactgacc
      601 ctatgccttc tggccggaca tgtccatttg tggtccatac gaatgggtca actgggatac
      661 ttcgatgacg agacggtggt gaatcatcag cgtttgctgg attacattga gcagcataaa
      721 ctcttggtgc ggtaagcttt gattaactaa cttttgacaa gaagtttatt cactttaact
      781 ggttccaaaa acgatgcact caatgtgcag attccacaac ctggtgagcc ggaccatcag
      841 cgaagtgcaa ctggtgcagc tgggcggatg tggagccact ctgtgcatca ttgtctccta
      901 catgctcttc tttgtgggcg acacaatctc gctggtctac tacttggtgt tctttggagt
      961 ggtctgcgtg cagctctttc ccagctgcta ttttgccagc gaagtagccg aggagttgga
     1021 acggctgcca tatgcgatct tctccagcag atggtacgat caatcgcggg atcatcgatt
     1081 cgatttgctc atctttacac aattaacact gggaaaccgg gggtggatca tcaaggcagg
     1141 aggtcttatc gagctgaatt tgaatgcctt tttcgccacc ctgaagatgg cctattccct
     1201 ttttgcagtt gtggtgcggg caaagggtat atagagagtc tgtttaatta aa
```

FIG. 8H-1

DOR72g

```
LOCUS       XXXXX         1321 bp    DNA              INV       08-FEB-1999
DEFINITION  Drosophila melanogaster odorant receptor DOR72g, GENSCAN
predicted complete cds.
ACCESSION   ;
KEYWORDS    .
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;
            Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1321)
  AUTHORS   Leslie B. Vosshall, Hubert Amrein, Pavel S. Morozov, Andrey
            Rzhetsky and Richard Axel.
  TITLE     A Spatial Map of Olfactory Receptor Expression in the Drosophila
            Antenna
  JOURNAL   Cell (1999) In press
REFERENCE   2  (bases 1 to 1321)
  AUTHORS   Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
            Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
            Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
            Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
            Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
            Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
            Pfeiffer,B., Poon,L., Punch,E., Sequeira,A., Sethi,H., Snir,E.,
            Svirskas,R.R., Twomey,B., Wan,K.H., Weinburg,T., Zhang,R.,
            Zieran,L.L. and Rubin,G.M.
  TITLE     Sequencing of Drosophila melanogaster
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 1321)
  AUTHORS   Vosshall,L.B., Amrein,H., Morozov,P.S., Rzhetsky,A. and Axel,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-FEB-1999) Department of Biochemistry and Molecular
            Biophysics, HHMI-Columbia University, 701 West 168th Street, New
            York, NY 10032, USA
FEATURES             Location/Qualifiers
     Comment:      /tissue_type='adult antenna'
                   /note='putative seven transmembrane G protein coupled
                   receptor'
                   /note='sequence annotation for BDGP P1 clone DS07071;
                   predicted using GENSCAN; join (27938-26618)'
     source           1..1321
                      /organism="Drosophila melanogaster"
                      /db_xref="taxon:7227"
                      /chromosome="II"
                      /map="33B1-2"
     mRNA             join(1..728,910..1321)
                      /gene="DOR72g"
     gene             1..1321
                      /note="odorant receptor"
                      /gene="DOR72g"
     exon             1..728
                      /gene="DOR72g"
                      /note="exon"
     CDS              join(1..728,910..1321)
                      /gene="DOR72g"
                      /note="odorant receptor"
                      /codon_start=1
                      /product="DOR72g"
```

FIG. 8H-2

```
/translation="MDLKPRVIRSEDIYRTYWLYWHLLGLESNFFLNRLLDLVITIFV
TIWYPIHLILGLFMERSLGDVCKGLPITAACFFASFKFICFRFKLSEIKEIEILFKEL
DQRALSREECEFFNQNTRREANFIWKSFIVAYGLSNISAIASVLFGGGHKLLYPAWFP
YDVQATELIFWLSVTYQIAGVSLAILQNLANDSYPPMTFCVVAGHVRLLAMRLSRIGQ
GPEETIYLTGKQLIESIEDHRKLMKIVELLRSTMNISQLGQFISSGVNISITLVNILF
FADNNFAITYYGVYFLSMVLELFPCCYYGTLISVEMNQLTYAIYSSNWMSMNRSYSRI
LLIFMQLTLAEVQIKAGGMIGIGMNAFFATVRLAYSFFTLAMSLR"
     intron          729..909
                     /gene="DOR72g"
                     /note="intron"
     exon            910..1321
                     /gene="DOR72g"
                     /note="exon"
BASE COUNT      361 a    255 c    283 g    422 t
ORIGIN
        1 atggacttaa aaccgcgagt cattcgaagt gaagatatct acagaaccta ttggttatat
       61 tggcatcttt tgggcctgga aagcaatttc tttctgaatc gcttgttgga tttggtgatt
      121 acaattttcg taaccatttg gtatccaatt cacctgattc tgggactgtt tatggaaaga
      181 tctttggggg atgtctgcaa gggtctacca attacggcag catgcttttt cgccagcttt
      241 aaatttattt gttttcgctt caagctatct gaaattaaag aaatcgaaat attatttaaa
      301 gagctggatc agcgagcttt aagtcgagag gaatgcgagt ttttcaatca aaatacgaga
      361 cgtgaggcga atttcatttg gaaaagtttc attgtggcct atggactgtc gaatatctcg
      421 gctattgcat cagttctttt cggcggtgga cataagctat tatatcccgc ctggtttcca
      481 tacgatgtgc aggccacgga actaatattt tggctaagtg taacatacca aattgccgga
      541 gtaagtttgg ccatacttca gaatttggcc aatgattcct atccaccgat gacattttgc
      601 gtggttgccg gtcatgtaag acttttggcg atgcgcttga gtagaattgg ccaaggtcca
      661 gaggaaacaa tatacttaac cggaaagcaa ttaatcgaaa gcatcgagga tcaccgaaaa
      721 ctaatgaagt aatgtacata tatagaatgg tttttagtta ttatcattaa atgaacgtgt
      781 tgtaggaaaa ccattctgtt tgtcgggtgt cacggaaatc gattttcctt aatttacata
      841 tgatattaaa tacttccttg caaacaatta tcatattagt aatttagaat ctttattatt
      901 tatttccaga atagtggaat tactgcgcag caccatgaat atttcgcagc tcggccagtt
      961 tatttcaagt ggtgttaata tttccataac actagtcaac attctcttct ttgcggataa
     1021 taatttcgct ataacctact acggagtgta cttcctatcg atggtgttgg aattattccc
     1081 gtgctgctat tacggcaccc tgatatccgt ggagatgaac cagctgacct atgcgattta
     1141 ctcaagtaac tggatgagta tgaatcggag ctacagccgc atcctactga tcttcatgca
     1201 actcaccctg gcggaagtgc agatcaaggc cggtgggatg attggcatcg gaatgaacgc
     1261 cttctttgcc accgtgcgat tggcctactc cttcttcact ttggccatgt cgctgcgtta
     1321 a
```

FIG. 8I-1

DOR73g

```
LOCUS       XXXXX        1212 bp    DNA             INV       08-FEB-1999
DEFINITION  Drosophila melanogaster odorant receptor DOR73g, GENSCAN
predicted complete cds.
ACCESSION   ;
KEYWORDS    .
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;
            Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1212)
  AUTHORS   Leslie B. Vosshall, Hubert Amrein, Pavel S. Morozov, Andrey
            Rzhetsky and Richard Axel.
  TITLE     A Spatial Map of Olfactory Receptor Expression in the Drosophila
            Antenna
  JOURNAL   Cell (1999) In press
REFERENCE   2  (bases 1 to 1212)
  AUTHORS   Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
            Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
            Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
            Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
            Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
            Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
            Pfeiffer,B., Poon,L., Punch,E., Sequeira,A., Sethi,H., Snir,E.,
            Svirskas,R.R., Twomey,B., Wan,K.H., Weinburg,T., Zhang,R.,
            Zieran,L.L. and Rubin,G.M.
  TITLE     Sequencing of Drosophila melanogaster
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 1212)
  AUTHORS   Vosshall,L.B., Amrein,H., Morozov,P.S., Rzhetsky,A. and Axel,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-FEB-1999) Department of Biochemistry and Molecular
            Biophysics, HHMI-Columbia University, 701 West 168th Street, New
            York, NY 10032, USA
FEATURES             Location/Qualifiers
     Comment:        /tissue_type='adult antenna'
                     /note='putative seven transmembrane G protein coupled
                     receptor'
                     /note='sequence annotation for BDGP P1 clone DS07071;
                     predicted using GENSCAN; join (29614-28403)'

  source             1..1212
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /chromosome="II"
                     /map="33B1-2"
  mRNA               join(1..722,798..1212)
                     /gene="DOR73g"
  gene               1..1212
                     /note="odorant receptor"
                     /gene="DOR73g"
  exon               1..722
                     /gene="DOR73g"
                     /note="exon"
  CDS                join(1..722,798..1212)
                     /gene="DOR73g"
                     /note="odorant receptor"
                     /codon_start=1
                     /product="DOR73g"
```

FIG. 8I-2

```
/translation="MDSRRKVRSENLYKTYWLYWRLLGVEGDYPFRRLVDFTITSFIT
ILFPVHLILGMYKKPQIQVFRSLHFTSECLFCSYKFFCFRWKLKEIKTIEGLLQDLDS
RVESEEERNYFNQNPSRVARMLSKSYLVAAISAIITATVAGLFSTGRNLMYLGWFPYD
FQATAAIYWISFSYQAIGSSLLILENLANDSYPPITFCVVSGHVRLLIMRLSRIGHDV
KLSSSENTRKLIEGIQDHRKLMKIIRLLRSTLHLSQLGQFLSSGINISITLINILFFA
ENNFAMLYYAVFFAAMLIELFPSCYYGILMTMEFDKLPYAIFSSNWLKMDKRYNRSLI
ILMQLTLVPVNIKAGGIVGIDMSAFFATVRMAYSFYTLALSFRV"
     intron          723..797
                     /gene="DOR73g"
                     /note="intron"
     exon            798..1212
                     /gene="DOR73g"
                     /note="exon"
BASE COUNT      352 a    231 c    242 g    387 t
ORIGIN
        1 atggattcaa gaaggaaagt ccgaagtgaa aatctttaca aaacctattg gctttactgg
       61 cgacttctgg gagtcgaggg cgattatcct tttcgacggc tagtggattt tacaatcacg
      121 tctttcatta cgattttatt tcccgtgcat cttatactgg gaatgtataa aaagccccag
      181 attcaagtct tcaggagtct gcatttcaca tcggaatgcc ttttctgcag ctataagttt
      241 ttctgttttc gttggaaact taaagaaata aagaccatcg aaggattgct ccaggatctc
      301 gatagtcgag ttgaaagtga agaagaacgc aactacttta atcaaaatcc aagtcgtgtg
      361 gctcgaatgc tttcgaaaag ttacttggta gctgctatat cggccataat cactgcaact
      421 gtagctggtt tatttagtac tggtcgaaat ttaatgtatc tgggttggtt tccctacgat
      481 tttcaagcaa ccgccgcaat ctattggatt agtttttcct atcaggcgat tggctctagt
      541 ctgttgattc tggaaaatct ggccaacgat tcatatccgc cgattacatt ttgtgtggtc
      601 tctggacatg tgagactatt gataatgcgt ttaagtcgaa ttggtcacga tgtaaaatta
      661 tcaagttcgg aaaataccag aaaactcatc gaaggtatcc aggatcacag gaaactaatg
      721 aagtaagaat aaagatttaa gaaccgcatg tttgatagct cagagaactg ataattaatc
      781 aaatgtaact ttccaggat aatacgccta cttcgcagca ctttacatct tagccaactg
      841 ggccagttcc tttctagtgg aatcaacatt tccataacac tcatcaacat cctgttcttt
      901 gcggaaaaca actttgcaat gctttattat gcggtgttct ttgctgcaat gttaatagaa
      961 ctatttccaa gttgttacta tggaattctg atgacaatgg agtttgataa gctaccatat
     1021 gccatcttct ccagcaactg gcttaaaatg gataaaagat acaatcgatc cttgataatt
     1081 ctgatgcaac taacactggt tccagtgaat ataaaagcag gtggtattgt tggcatcgat
     1141 atgagtgcat tttttgccac agttcggatg gcatattcct tttacacttt agccttgtca
     1201 tttcgagtat ag
```

FIG. 8J-1

DOR46

```
LOCUS       XXXXX        1198 bp    DNA             INV       08-FEB-1999
DEFINITION  Drosophila melanogaster odorant receptor DOR46g, GENSCAN
predicted complete cds.
ACCESSION   ;
KEYWORDS    .
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;
            Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1198)
  AUTHORS   Leslie B. Vosshall, Hubert Amrein, Pavel S. Morozov, Andrey
            Rzhetsky and Richard Axel.
  TITLE     A Spatial Map of Olfactory Receptor Expression in the Drosophila
            Antenna
  JOURNAL   Cell (1999) In press
REFERENCE   2  (bases 1 to 1198)
  AUTHORS   Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
            Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
            Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
            Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
            Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
            Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
            Pfeiffer,B., Poon,L., Punch,E., Sequeira,A., Sethi,H., Snir,E.,
            Svirskas,R.R., Twomey,B., Wan,K.H., Weinburg,T., Zhang,R.,
            Zieran,L.L. and Rubin,G.M.
  TITLE     Sequencing of Drosophila melanogaster
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 1198)
  AUTHORS   Vosshall,L.B., Amrein,H., Morozov,P.S., Rzhetsky,A. and Axel,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-FEB-1999) Department of Biochemistry and Molecular
            Biophysics, HHMI-Columbia University, 701 West 168th Street, New
            York, NY 10032, USA
FEATURES             Location/Qualifiers
     Comment:        /note='putative seven transmembrane G protein coupled
                     receptor'
                     /note='sequence annotation for BDGP P1 clone DS07462;
                     predicted using GENSCAN'
     source          1..1198
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /chromosome="II"
                     /map="59D5-7"
     mRNA            join(1..725,784..1198)
                     /gene="DOR46g"
     gene            1..1198
                     /note="odorant receptor"
                     /gene="DOR46g"
     exon            1..725
                     /gene="DOR46g"
     CDS             join(1..725,784..1198)
                     /gene="DOR46g"
                     /note="odorant receptor"
                     /codon_start=1
                     /product="DOR46g"
```

FIG. 8J-2

```
/translation="MAEVRVDSLEFFKSHWTAWRYLGVAHFRVENWKNLYVFYSIVSN
LLVTLCYPVHLGISLFRNRTITEDILNLTTFATCTACSVKCLLYAYNIKDVLEMERLL
RLLDERVVGPEQRSIYGQVRVQLRNVLYVFIGIYMPCALFAELSFLFKEERGLMYPAW
FPFDWLHSTRNYYIANAYQIVGISFQLLQNYVSDCFPAVVLCLISSHIKMLYNRFEEV
GLDPARDAEKDLEACITDHKHILELFRRIEAFISLPMLIQFTVTALNVCIGLAALVFF
VSEPMARMYFIFYSLAMPLQIFPSCFFGTDNEYWFGRLHYAAFSCNWHTQNRSFKRKM
MLFVEQSLKKSTAVAGGMMRIHLDTFFSTLKGAYSLFTIIIRMRK"
     intron          726..783
                     /gene="DOR46g"
     exon            784..1198
                     /gene="DOR46g"
BASE COUNT      251 a     325 c     310 g     312 t
ORIGIN
        1 atggcagagg tcagagtgga cagtctggag tttttcaaga gccattggac cgcctggcgg
       61 tacttgggag tggctcattt tcgggtcgag aactggaaga acctttacgt gttttacagc
      121 attgtgtcga atcttctcgt gaccctgtgc taccccgttc acctgggaat atccctcttt
      181 cgcaaccgca ccatcaccga ggacatcctc aacctgacca cctttgcgac ctgcacagcc
      241 tgttcggtga agtgcctgct ctacgcctac aacatcaagg atgtgctgga gatggagcgg
      301 ctgttgaggc ttttggatga acgcgtcgtg ggtccggagc aacgcagcat ctacggacaa
      361 gtgagggtcc agctgcgaaa tgtgctatac gtgttcatcg gcatctacat gccgtgtgcc
      421 ctgttcgccg agctatcctt tctgttcaag gaggagcgcg gtctgatgta tcccgcctgg
      481 tttcccttcg actggctgca ctccaccagg aactattaca tagcgaacgc ctatcagata
      541 gtgggcatct cgtttcagct gctgcaaaac tatgttagcg actgctttcc ggcggtggtg
      601 ctgtgcctga tctcatccca catcaaaatg ttgtacaaca gattcgagga ggtgggcctg
      661 gatccagcca gagatgcgga gaaggacctg gaggcctgca tcaccgatca caagcatatt
      721 ctagagtggg caggcggctc attgtaacgt tcgtgttcta ttcactttcc aacttttttc
      781 cagactattc cgacgcatcg aggccttcat ttccctgccc atgctaattc agttcacagt
      841 gaccgccttg aatgtgtgca tcggtttagc agccctggtg tttttcgtca gcgagcccat
      901 ggcacggatg tacttcatct tctactccct ggccatgccg ctgcagatct ttccgtcctg
      961 ctttttcggc accgacaacg agtactggtt cggacgcctc cactacgcgg ccttcagttg
     1021 caattggcac acacagaaca ggagctttaa gcggaaaatg atgctgttcg ttgagcaatc
     1081 gttgaagaag agcaccgctg tggctggcgg aatgatgcgt atccacctgg acacgttctt
     1141 ttccacccta aagggggcct actccctctt taccatcatt attcggatga gaaagtag
```

FIG. 8K-1

DOR19g

```
LOCUS       XXXXX        1293 bp    DNA           INV       09-FEB-1999
DEFINITION  Drosophila melanogaster odorant receptor DOR19g, GENSCAN
predicted complete cds.
ACCESSION   ;
KEYWORDS    .
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;
            Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1293)
  AUTHORS   Leslie B. Vosshall, Hubert Amrein, Pavel S. Morozov, Andrey
            Rzhetsky and Richard Axel.
  TITLE     A Spatial Map of Olfactory Receptor Expression in the Drosophila
            Antenna
  JOURNAL   Cell (1999) In press
REFERENCE   2  (bases 1 to 1293)
  AUTHORS   Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
            Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
            Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
            Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
            Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
            Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
            Pfeiffer,B., Poon,L., Punch,E., Sequeira,A., Sethi,H., Snir,E.,
            Svirskas,R.R., Twomey,B., Wan,K.H., Weinburg,T., Zhang,R.,
            Zieran,L.L. and Rubin,G.M.
  TITLE     Sequencing of Drosophila melanogaster
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 1293)
  AUTHORS   Vosshall,L.B., Amrein,H., Morozov,P.S., Rzhetsky,A. and Axel,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-FEB-1999) Department of Biochemistry and Molecular
            Biophysics, HHMI-Columbia University, 701 West 168th Street, New
            York, NY 10032, USA
FEATURES             Location/Qualifiers
     Comment:       /note='putative seven transmembrane G protein coupled
                    receptor'
                    /note='sequence annotation for BDGP P1 clone DS01913;
                    predicted using GENSCAN; join (72218-70926)'
     source            1..1293
                       /organism="Drosophila melanogaster"
                       /db_xref="taxon:7227"
                       /chromosome="II"
                       /map="46F5-6"
     mRNA              1..1293
                       /gene="DOR19g"
     gene              1..1293
                       /note="odorant receptor"
                       /gene="DOR19g"
     exon              1..1293
                       /gene="DOR19g"
     CDS               1..1293
                       /gene="DOR19g"
                       /note="odorant receptor"
                       /codon_start=1
                       /product="DOR19g"
```

FIG. 8K-2

/translation="MVTEDFYKYQVWYFQILGVWQLPTWAADHQRRFQSMRFGFILVI
LFIMLLLFSFEMLNNISQVREILKVFFMFATEISCMAKLLHLKLKSRKLAGLVDAMLS
PEFGVKSEQEMQMLELDRVAVVRMRNSYGIMSLGAASLILIVPCFDNFGELPLAMLEV
CSIEGWICYWSQYLFHSICLLPTCVLNITYDSVAYSLLCFLKVQLQMLVLRLEKLGPV
IEPQDNEKIAMELRECAAYYNRIVRFKDLVELFIKGPGSVQLMCSVLVLVSNLYDMST
MSIANGDAIFMLKTCIYQLVMLWQIFIICYASNEVTVQSSRLCHSIYSSQWTGWNRAN
RRIVLLMMQRFNSPMLLSTFNPTFAFSLEAFGSVGQQKFLYISFITGYALLLSDRQLL
LQLLRTAEARQQLNFETPQHLKIFKPIFKSTQNVMHVH"
BASE COUNT 290 a 322 c 307 g 374 t
ORIGIN

```
   1 atggttacgg aggactttta taagtaccag gtgtggtact tccaaatcct tggtgtttgg
  61 cagctcccca cttgggccgc agaccaccag cgtcgttttc agtccatgag gtttggcttc
 121 atcctggtca tcctgttcat catgctgctg cttttctcct tcgaaatgtt gaacaacatt
 181 tcccaagtta gggagatcct aaaggtattc ttcatgttcg ccacggaaat atcctgcatg
 241 gccaaattat tgcatttgaa gttgaagagc cgcaaactcg ctggcttggt tgatgcgatg
 301 ttgtccccag agttcggcgt taaaagtgaa caggaaatgc agatgctgga attggataga
 361 gtggcggttg tccgcatgag gaactcctac ggcatcatgt ccctgggcgc ggcttccctg
 421 atccttatag ttccctgttt cgacaacttt ggcgagctac cactggccat gttggaggta
 481 tgcagcatcg agggatggat ctgctattgg tcgcagtacc ttttccactc gatttgcctg
 541 ctgcccactt gtgtgctgaa tataacctac gactcggtgg cctactcgtt gctctgtttc
 601 ttgaaggttc agctacaaat gctggtcctg cgattagaaa agttgggtcc tgtgatcgaa
 661 ccccaggata atgagaaaat cgcaatggaa ctgcgtgagt gtgccgccta ctacaacagg
 721 attgttcgtt tcaaggacct ggtggagctg ttcataaagg ggccaggatc tgtgcagctc
 781 atgtgttctg ttctggtgct ggtgtccaac ctgtacgaca tgtccaccat gtccattgca
 841 aacggcgatg ccatctttat gctcaagacc tgtatctatc agctggtgat gctctggcag
 901 atcttcatca tttgctacgc ctccaacgag gtaactgtcc agagctctag gttgtgtcac
 961 agcatctaca gctcccaatg gacgggatgg aacagggcaa accgccggat tgtccttctc
1021 atgatgcagc gctttaattc cccgatgctc ctgagcacct ttaaccccac ctttgctttc
1081 agcttggagg cctttggttc tgtagggcag cagaaattcc tttatatatc atttattact
1141 ggttatgctc ttctcctttc agatcgtcaa ctgctcctac agctacttcg cactgctgaa
1201 gcgcgtcaac agttaaattt cgaaacaccg cagcacctaa agattttcaa gccgattttt
1261 aaaagcactc aaaacgttat gcacgtacat taa
```

FIG. 8L-1

DOR24g

```
LOCUS       XXXXX        2075 bp    DNA              INV       09-FEB-1999
DEFINITION  Drosophila melanogaster odorant receptor DOR24g, GENSCAN
predicted complete cds.
ACCESSION   ;
KEYWORDS    .
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea;
            Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 2075)
  AUTHORS   Leslie B. Vosshall, Hubert Amrein, Pavel S. Morozov, Andrey
            Rzhetsky and Richard Axel.
  TITLE     A Spatial Map of Olfactory Receptor Expression in the Drosophila
            Antenna
  JOURNAL   Cell (1999) In press
REFERENCE   2  (bases 1 to 2075)
  AUTHORS   Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,
            Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,
            Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,
            Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,
            Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,
            Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,
            Pfeiffer,B., Poon,L., Punch,E., Sequeira,A., Sethi,H., Snir,E.,
            Svirskas,R.R., Twomey,B., Wan,K.H., Weinburg,T., Zhang,R.,
            Zieran,L.L. and Rubin,G.M.
  TITLE     Sequencing of Drosophila melanogaster
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 2075)
  AUTHORS   Vosshall,L.B., Amrein,H., Morozov,P.S., Rzhetsky,A. and Axel,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-FEB-1999) Department of Biochemistry and Molecular
            Biophysics, HHMI-Columbia University, 701 West 168th Street, New
            York, NY 10032, USA
FEATURES             Location/Qualifiers
     Comment:       /tissue_type='adult antenna'
                    /note='putative seven transmembrane G protein coupled
                    receptor'
                    /note='sequence annotation for BDGP P1 clone DS00724;
                    predicted using GENSCAN; join (62623-64697)'
     source            1..2075
                       /organism="Drosophila melanogaster"
                       /db_xref="taxon:7227"
                       /chromosome="II"
                       /map="47D6-E2"
     mRNA              join(1..185,739..1038,1338..1953,2025..2075)
                       /gene="DOR24g"
     gene              1..2075
                       /note="odorant receptor"
                       /gene="DOR24g"
     CDS               join(1..185,739..1038,1338..1953,2025..2075)
                       /gene="DOR24g"
                       /note="odorant receptor"
                       /codon_start=1
                       /product="DOR24g"
```

FIG. 8L-2

```
/translation="MRPTLQVLISVLCLVSAYAWDHTDCNDHYIEFMDYPDERATAYS
NESSEWDFFEFWRQVFGLFLQVQKSTIALLGFDLFSENREMWKRPYRAMNVFSIAAIF
PFILAAVLHNWKNVLLLADAMVALLITILGLFKFSMILYLRRDFKRLIDKFRLLMSNG
EFFPWNIHIIRNYVLSFIWSAFASTGVVLPAVSLDTIFCSFTSNLCAFFKIAQYKVVR
FKGGSLKESQATLNKVFALYQTSLDMCNDLNQCYQPIICAQFFISSLQLCMLGYLFSI
TFAQTEGVYYASFIATIIIQAYIYCYCGENLKTESASFEWAIYDSPWHESLGAGGAST
SICRSLLISMMRAHRGFRITGYFFEANMEAFSSIVRTAMSYITMLRSFS"
     exon            1..185
                     /gene="DOR24g"
     intron          186..738
                     /gene="DOR24g"
     exon            739..1038
                     /gene="DOR24g"
     intron          1039..1337
                     /gene="DOR24g"
     exon            1338..1953
                     /gene="DOR24g"
     intron          1954..2024
                     /gene="DOR24g"
     exon            2025..2075
                     /gene="DOR24g"
BASE COUNT      548 a    460 c    447 g    620 t
ORIGIN
        1 atgcgcccaa cactccaggt cctcatcagt gtcctttgtc tggtctctgc ctacgcctgg
       61 gatcataccg actgtaacga tcactatatc gaattcatgg actatcccga tgagagagct
      121 accgcctata gcaatgaatc ctcagaatgg gatttctttg aattttggag gcaagtgttt
      181 ggcctgtagc aattctgaca tcgctgacca cgcctctggt caaagcaatc acttggcgag
      241 acatgtgacc ctcggatgtc agggcgtgaa taaagaacat taaaatttaa ccggagttct
      301 catttactat gcctaaagag gtgtcaaacg gaggtattat ttgggcgtga tatattatac
      361 cctctaaagg agtatcccca ccaggtgaaa ccctataaaa cccttgacgt cgttaatgga
      421 aagtactgaa ccggaaatga aagtgggcac tttctaatgg tagaaatttg tcgggatgaa
      481 atatataagt agatatatat aatattatat cttgattatc cagcatcaat cacaagataa
      541 aaaaaaaata taattcataa ttcataattc atatgtaata ggcatttgta aatgttgtaa
      601 acaaaggaag ggtattaaag agctgtactt tcgcatctgt atttcattta ttcctactca
      661 attgaaaggt ttaaaacccc tgaactgaac acacttgact tagtgtgagg ccgaattaac
      721 ccttgtcgac atggacagtt ttctgcaagt acagaagagc accattgccc ttctgggctt
      781 tgatctcttt agtgaaaatc gagaaatgtg gaaacgcccc tatagagcaa tgaatgtgtt
      841 tagcatagct gccatttttc cctttatcct ggcagctgtg ctccataatt ggaagaatgt
      901 attgctgctg gccgatgcca tggtggccct actaataacc attctgggcc tattcaagtt
      961 tagcatgata ctttacttac gtcgcgattt caagcgactg attgacaaat ttcgtttgct
     1021 catgtcgaat ggtgagttgt aatccatttc ggccagaatg tgtatcattt catttattat
     1081 tttatagagg cggaacaggg cgaggaatac gccgagattc tcaacgcagc aaacaagcag
     1141 gatcaacgaa tgtgcactct gtttaggact tgtttcctcc tcgcctgggc cttgaatagt
     1201 gttctgcccc tcgtgagaat gggtctcagc tattggttag caggtcatgc agagcccgag
     1261 ttgccttttc cctgtctgta tgtacaaatg atatatatga tatatggtga tcaagttatc
     1321 aggctttgtt cctaaagttt tccctggaat atccacatca ttcgcaatta tgttttgagc
     1381 ttcatctgga gcgctttcgc ctcgacaggt gtggttttac ctgctgtcag cttggatacc
     1441 atattctgtt ccttcaccag caacctgtgc gccttcttca aaattgcgca gtacaaggtg
     1501 gttagattta agggcggatc ccttaaagaa tcacaggcca cattgaacaa agtctttgcc
     1561 ctgtaccaga ccagcttgga tatgtgcaac gatctgaatc agtgctacca accgattatc
     1621 tgcgcccagt tcttcatttc atctctgcaa ctctgcatgc tgggatatct gttctccatt
     1681 acttttgccc agacagaggg cgtctactat gcctcattca tagccacaat cattatacaa
     1741 gcctatatct actgctactg cggggagaac ctgaagacgg agagtgccag cttcgagtgg
     1801 gccatctacg acagtccgtg gcacgagagt ttgggtgctg gtggagcctc tacctcgatc
     1861 tgccgatcct tgctgatcag catgatgcgg gctcatcggg gattccgcat tacgggatac
     1921 tttttcgagg caaacatgga ggccttctca tcggtgggtg aatcatttcc attgtacaat
     1981 acatggattt acatgaatac tcttttctaa ctttccgttt ttagattgtt cgcacggcga
     2041 tgtcctacat cacaatgctg agatcattct cctaa
```

FIG. 9

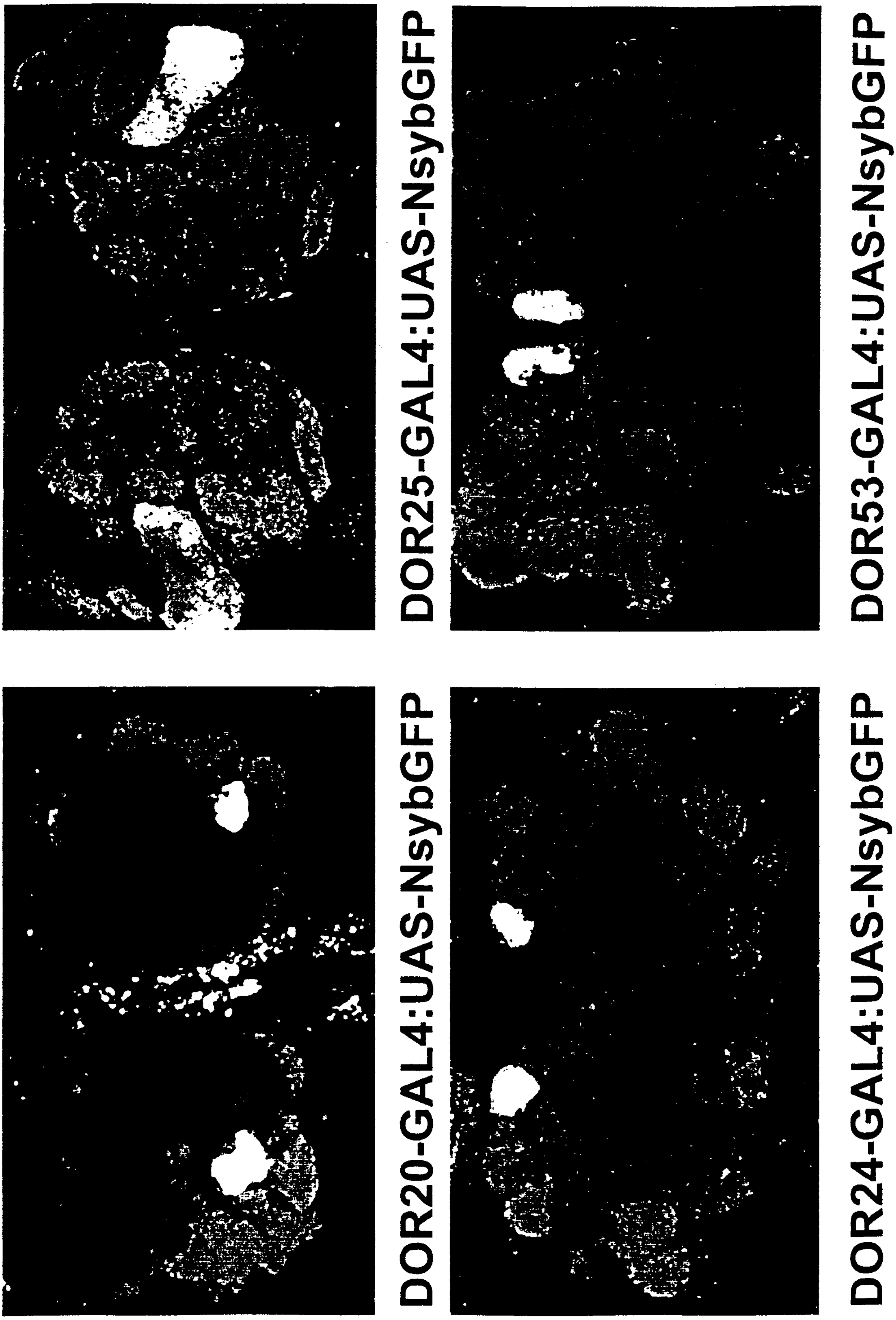
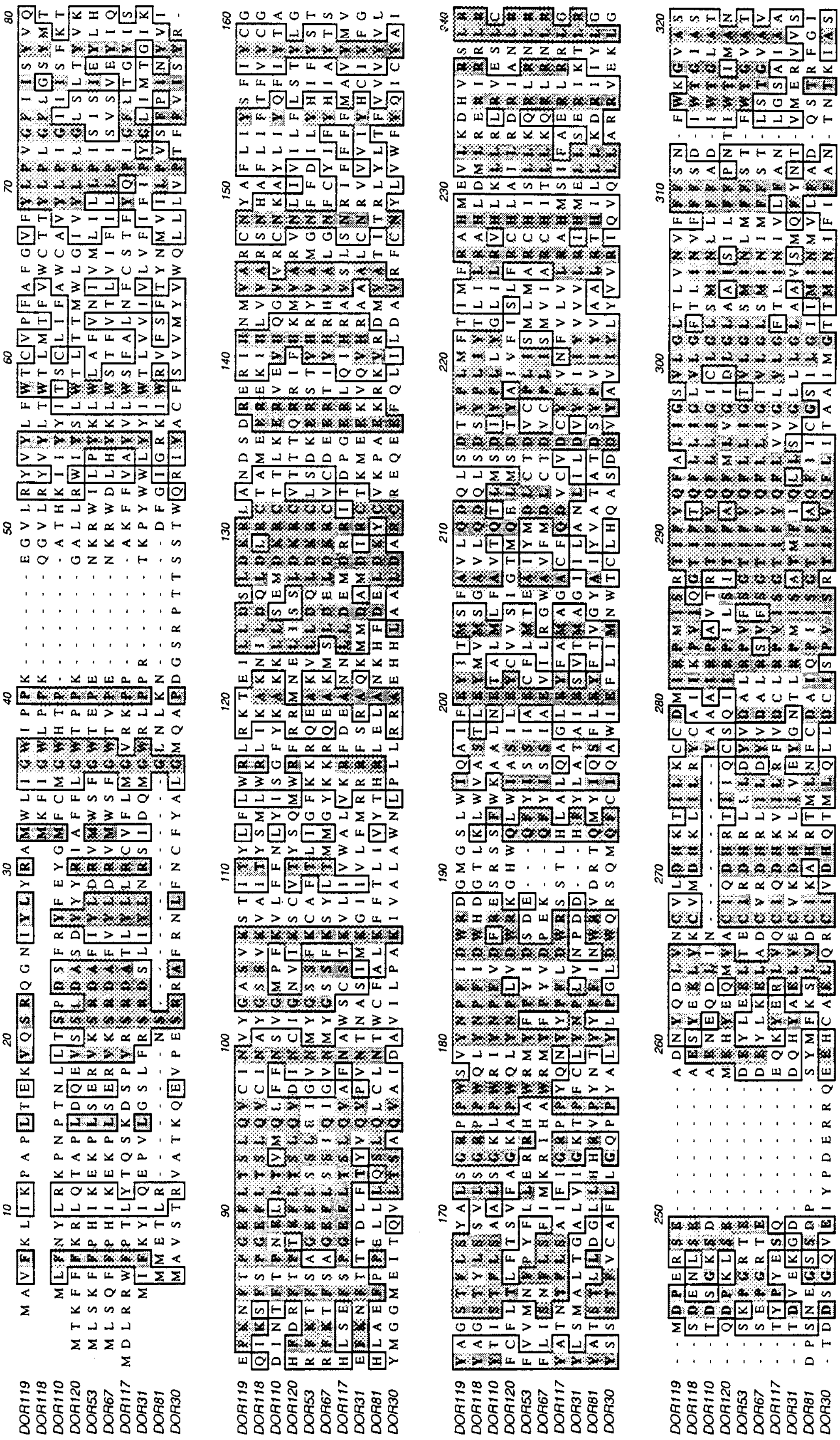

DOR20-GAL4:UAS-NsybGFP
DOR25-GAL4:UAS-NsybGFP
DOR24-GAL4:UAS-NsybGFP
DOR53-GAL4:UAS-NsybGFP

FIG. 10A-1

ClustalW Alignment of the DOR53 Family



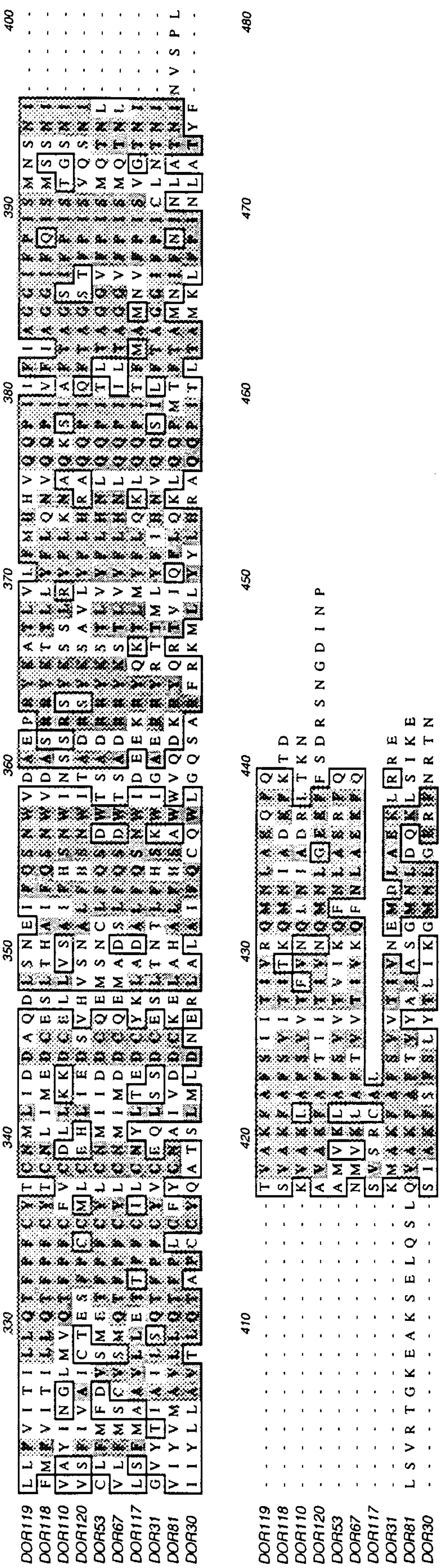
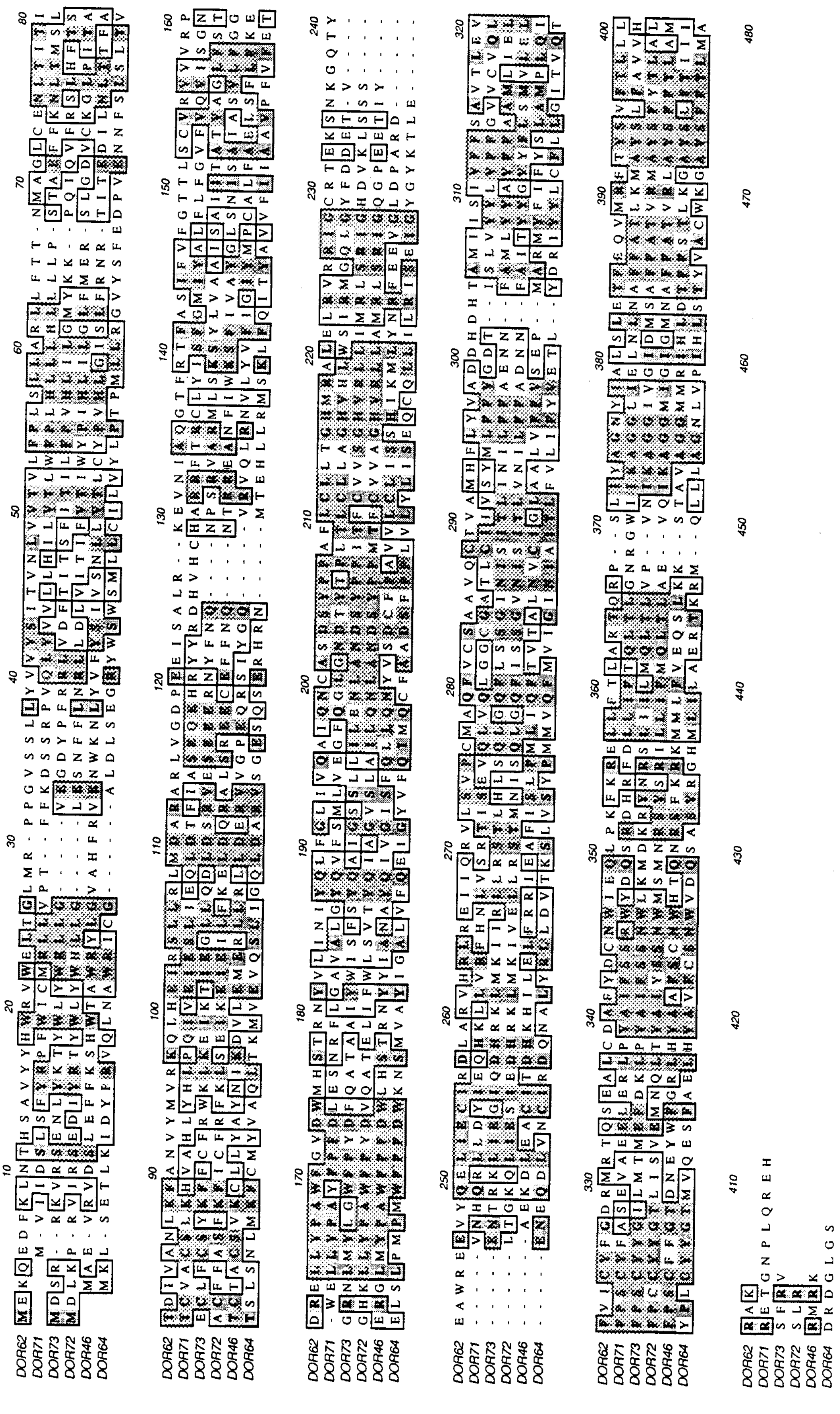


FIG. 10A-2

FIG. 10B

ClustalW Alignment of the DOR64 Family



SEQUENCE LISTING

<110> Vosshall, Leslie B.
The Trustees of Columbia University in the City of

<120> GENES ENCODING INSECT ODORANT RECEPTORS AND USES THEREOF

<130> 0575/58715

<140> 09/257,706
<141> 1999-02-25

<160> 24

<170> PatentIn Ver. 2.0

<210> 1
<211> 553
<212> DNA
<213> Drosophila melanogaster DOR62

<400> 1
caggaactca tcgagtgcat ccgcgatctg gcgcgggtcc atcggctgag ggagatcatt 60
cagcgggtcc tttcagtgcc ctgcatggcc cagttcgtct gctccgccgc cgtccagtgt 120
accgtcgcca tgcacttcct gtacgtagcg gatgaccacg accacaccgc catgatcatc 180
tcgattgtat ttttctcggc cgtcaccttg gaggtgtttg taatctgcta ttttggggac 240
aggatgcgga cacagagcga ggcgctgtgc gatgccttct acgattgcaa ctggatagaa 300
cagctgccca agttcaagcg cgaactgctc ttcaccctgg ccaggacgca gcggccttct 360
cttatttacg caggcaacta catcgcactc tcgctggaga ccttcgagca ggtcatgagg 420
ttcacatact ctgttttcac actcttgctg agggccaagt aagaacttta taatctcttt 480
ttggggagaa aaattttaaa gcacaatagc agaaaaatat atcagataat ataacaaaaa 540
aaaaaaaaaa aaa 553

<210> 2
<211> 153
<212> PRT
<213> Translation DOR62

<400> 2
Gln Glu Leu Ile Glu Cys Ile Arg Asp Leu Ala Arg Val His Arg Leu
1 5 10 15

Arg Glu Ile Ile Gln Arg Val Leu Ser Val Pro Cys Met Ala Gln Phe
20 25 30

Val Cys Ser Ala Ala Val Gln Cys Thr Val Ala Met His Phe Leu Tyr
35 40 45

```
Val Ala Asp Asp His Asp His Thr Ala Met Ile Ile Ser Ile Val Phe
     50                  55                  60

Phe Ser Ala Val Thr Leu Glu Val Phe Val Ile Cys Tyr Phe Gly Asp
 65                  70                  75                  80

Arg Met Arg Thr Gln Ser Glu Ala Leu Cys Asp Ala Phe Tyr Asp Cys
                 85                  90                  95

Asn Trp Ile Glu Gln Leu Pro Lys Phe Lys Arg Glu Leu Leu Phe Thr
            100                 105                 110

Leu Ala Arg Thr Gln Arg Pro Ser Leu Ile Tyr Ala Gly Asn Tyr Ile
        115                 120                 125

Ala Leu Ser Leu Glu Thr Phe Glu Gln Val Met Arg Phe Thr Tyr Ser
    130                 135                 140

Val Phe Thr Leu Leu Leu Arg Ala Lys
145                 150


<210> 3
<211> 1493
<212> DNA
<213> Drosophila melanogaster DOR104

<400> 3
ggcacgagca gtcgatggcc agtcttcagt tccacggcaa cgtcgatgcg gacatcaggt 60
atgatattag cctggatccg gctagggaat cgaatctctt ccgtctgcta atgggactcc 120
agttggcgaa tggcacgaag ccatcgccgc ggttacccaa atggtggcca aagcggctgg 180
aaatgattgg taaagtgctg cccaaagcct attgttccat ggtgattttc acctccctgc 240
atttgggtgt cctgttcacg aaaaccacac tggatgtcct gccgacgggg gagctgcagg 300
ccataacgga tgccctcacc atgaccataa tatacttttt cacgggctac ggcaccatct 360
actggtgcct gcgctcccgg cgcctcttgg cctacatgga gcacatgaac cgggagtatc 420
gccatcattc gctggccggg gtgacctttg tgagtagcca tgcggccttt aggatgtcca 480
gaaacttcac ggtggtgtgg ataatgtcct gcctgctggg cgtgatttcc tggggcgttt 540
cgccactgat gctgggcatc cggatgctgc cgctccaatg ttggtatccc ttcgacgccc 600
tgggtcccgg cacatatacg gcggtctatg ctacacaact tttcggtcag atcatggtgg 660
gcatgacctt tggattcggg ggatcactgt ttgtcaccct gagcctgcta ctcctgggac 720
aattcgatgt gctctactgc agcctgaaga acctggatgc ccataccaag ttgctgggcg 780
gggagtctgt aaatggcctg agttcgctgc aagaggagtt gctgctgggg gactcgaaga 840
gggaattaaa tcagtacgtt ttgctccagg agcatccgac ggatctgctg agattgtcgg 900
caggacgaaa atgtcctgac caaggaaatg cgtttcacaa cgccttggtg gaatgcattc 960
gcttgcatcg cttcattctg cactgctcac aggagttgga gaatctattc agtccatatt 1020
gtctggtcaa gtcactgcag atcacctttc agctttgcct gctggtcttt gtgggcgttt 1080
cgggtactcg agaggtcctg cggattgtca accagctaca gtacttggga ctgaccatct 1140
```

tcgagctcct aatgttcacc tattgtggcg aactcctcag tcggcatagt attcgatctg 1200
gcgacgcctt ttggaggggt gcgtggtgga agcacgccca tttcatccgc caggacatcc 1260
tcatctttct ggtcaatagt agacgtgcag ttcacgtgac tgccggcaag ttttatgtga 1320
tggatgtgaa tcgtctaaga tcggttataa cgcaggcgtt cagcttcttg actttgctgc 1380
aaaagttggc tgccaagaag acggaatcgg agctctaaac tggtaccacg catcgatatt 1440
tatttagcgc attaaaaaaa agtcgagtaa aagcaaaaaa aaaaaaaaaa aaa 1493

<210> 4
<211> 467
<212> PRT
<213> Translation DOR104

<400> 4
Met Ala Ser Leu Gln Phe His Gly Asn Val Asp Ala Asp Ile Arg Tyr
1   5   10   15

Asp Ile Ser Leu Asp Pro Ala Arg Glu Ser Asn Leu Phe Arg Leu Leu
            20   25   30

Met Gly Leu Gln Leu Ala Asn Gly Thr Lys Pro Ser Pro Arg Leu Pro
        35   40   45

Lys Trp Trp Pro Lys Arg Leu Glu Met Ile Gly Lys Val Leu Pro Lys
    50   55   60

Ala Tyr Cys Ser Met Val Ile Phe Thr Ser Leu His Leu Gly Val Leu
65   70   75   80

Phe Thr Lys Thr Thr Leu Asp Val Leu Pro Thr Gly Glu Leu Gln Ala
                85   90   95

Ile Thr Asp Ala Leu Thr Met Thr Ile Ile Tyr Phe Phe Thr Gly Tyr
            100   105   110

Gly Thr Ile Tyr Trp Cys Leu Arg Ser Arg Arg Leu Leu Ala Tyr Met
        115   120   125

Glu His Met Asn Arg Glu Tyr Arg His His Ser Leu Ala Gly Val Thr
    130   135   140

Phe Val Ser Ser His Ala Ala Phe Arg Met Ser Arg Asn Phe Thr Val
145   150   155   160

Val Trp Ile Met Ser Cys Leu Leu Gly Val Ile Ser Trp Gly Val Ser
                165   170   175

Pro Leu Met Leu Gly Ile Arg Met Leu Pro Leu Gln Cys Trp Tyr Pro
            180   185   190

Phe Asp Ala Leu Gly Pro Gly Thr Tyr Thr Ala Val Tyr Ala Thr Gln
195 200 205

Leu Phe Gly Gln Ile Met Val Gly Met Thr Phe Gly Phe Gly Gly Ser
210 215 220

Leu Phe Val Thr Leu Ser Leu Leu Leu Leu Gly Gln Phe Asp Val Leu
225 230 235 240

Tyr Cys Ser Leu Lys Asn Leu Asp Ala His Thr Lys Leu Leu Gly Gly
245 250 255

Glu Ser Val Asn Gly Leu Ser Ser Leu Gln Glu Glu Leu Leu Leu Gly
260 265 270

Asp Ser Lys Arg Glu Leu Asn Gln Tyr Val Leu Leu Gln Glu His Pro
275 280 285

Thr Asp Leu Leu Arg Leu Ser Ala Gly Arg Lys Cys Pro Asp Gln Gly
290 295 300

Asn Ala Phe His Asn Ala Leu Val Glu Cys Ile Arg Leu His Arg Phe
305 310 315 320

Ile Leu His Cys Ser Gln Glu Leu Glu Asn Leu Phe Ser Pro Tyr Cys
325 330 335

Leu Val Lys Ser Leu Gln Ile Thr Phe Gln Leu Cys Leu Leu Val Phe
340 345 350

Val Gly Val Ser Gly Thr Arg Glu Val Leu Arg Ile Val Asn Gln Leu
355 360 365

Gln Tyr Leu Gly Leu Thr Ile Phe Glu Leu Leu Met Phe Thr Tyr Cys
370 375 380

Gly Glu Leu Leu Ser Arg His Ser Ile Arg Ser Gly Asp Ala Phe Trp
385 390 395 400

Arg Gly Ala Trp Trp Lys His Ala His Phe Ile Arg Gln Asp Ile Leu
405 410 415

Ile Phe Leu Val Asn Ser Arg Arg Ala Val His Val Thr Ala Gly Lys
420 425 430

Phe Tyr Val Met Asp Val Asn Arg Leu Arg Ser Val Ile Thr Gln Ala
435 440 445

Phe Ser Phe Leu Thr Leu Leu Gln Lys Leu Ala Ala Lys Lys Thr Glu
450 455 460

Ser Glu Leu
465

<210> 5
<211> 1556
<212> DNA
<213> Drosophila melanogaster DOR87

<400> 5
ggcacgaggc ttatagaaag tgccgagcaa tgacaatcga ggatatcggc ctggtgggca 60
tcaacgtgcg gatgtggcga cacttggccg tgctgtaccc cactccgggc tccagctggc 120
gcaagttcgc cttcgtgctg ccggtgactg cgatgaatct gatgcagttc gtctacctgc 180
tgcggatgtg gggcgacctg cccgccttca ttctgaacat gttcttcttc tcggccattt 240
tcaacgccct gatgcgcacg tggctggtca taatcaagcg gcgccagttc gaggagtttc 300
tcggccaact ggccactctg ttccattcga ttctcgactc caccgacgag tgggggcgtg 360
gcatcctgcg gagggcggaa cgggaggctc ggaacctggc catccttaat ttgagtgcct 420
ccttcctgga cattgtcggt gctctggtat cgccgctttt cagggaggag agagctcatc 480
ccttcggcgt agctctacca ggagtgagca tgaccagttc acccgtctac gaggttatct 540
acttggccca actgcctacg cccctgctgc tgtccatgat gtacatgcct ttcgtcagcc 600
tttttgccgg cctggccatc tttgggaagg ccatgctgca gatcctggta cacaggctgg 660
gccagattgg cggagaagag cagtcggagg aggagcgctt ccaaaggctg gcctcctgca 720
ttgcgtacca cacgcaggtg atgcgctatg tgtggcagct caacaaactg gtggccaaca 780
ttgtggcggt ggaagcaatt atttttggct cgataatctg ctcactgctc ttctgtctga 840
atattataac ctcacccacc caggtgatct cgatagtgat gtacattctg accatgctgt 900
acgttctctt cacctactac aatcgggcca atgaaatatg cctcgagaac aaccgggtgg 960
cggaggctgt ttacaatgtg ccctggtacg aggcaggaac tcggtttcgc aaaaccctcc 1020
tgatcttctt gatgcaaaca caacacccga tggagataag agtcggcaac gtttacccca 1080
tgacattggc catgttccag agtctgttga atgcgtccta ctcctacttt accatgctgc 1140
gtggcgtcac cggcaaatga gctgaaagac cgaaaaaacc ggagtatccc cttccatatt 1200
cccccctgctc ctttattttc ctttcctttt ccctttccgt tttcccattc gcttttccag 1260
caatccgggt aatgcaaaaa gttgttgctg gctgtggtcc tggctgcttg tttggcattt 1320
gcatatgctt gtcgtttgaa aggatttaat cggactgctg gcacggagtc ggcatcctgg 1380
ctcctggatc ctggcatgca aatagttggc ttcttagatt gttacacaaa atagattgta 1440
gattgcagct gaatgttgtg cttggaataa agtcaaaagg atgtggagtc ggcccaaggc 1500
tctgcccatt ctgtttgctc gggatgcccg aaagtatgaa aaaaaaaaaa aaaaaa 1556

<210> 6
<211> 376
<212> PRT
<213> Translation DOR87

<400> 6
Met Thr Ile Glu Asp Ile Gly Leu Val Gly Ile Asn Val Arg Met Trp

1 5 10 15

Arg His Leu Ala Val Leu Tyr Pro Thr Pro Gly Ser Ser Trp Arg Lys
20 25 30

Phe Ala Phe Val Leu Pro Val Thr Ala Met Asn Leu Met Gln Phe Val
35 40 45

Tyr Leu Leu Arg Met Trp Gly Asp Leu Pro Ala Phe Ile Leu Asn Met
50 55 60

Phe Phe Phe Ser Ala Ile Phe Asn Ala Leu Met Arg Thr Trp Leu Val
65 70 75 80

Ile Ile Lys Arg Arg Gln Phe Glu Glu Phe Leu Gly Gln Leu Ala Thr
85 90 95

Leu Phe His Ser Ile Leu Asp Ser Thr Asp Glu Trp Gly Arg Gly Ile
100 105 110

Leu Arg Arg Ala Glu Arg Glu Ala Arg Asn Leu Ala Ile Leu Asn Leu
115 120 125

Ser Ala Ser Phe Leu Asp Ile Val Gly Ala Leu Val Ser Pro Leu Phe
130 135 140

Arg Glu Glu Arg Ala His Pro Phe Gly Val Ala Leu Pro Gly Val Ser
145 150 155 160

Met Thr Ser Ser Pro Val Tyr Glu Val Ile Tyr Leu Ala Gln Leu Pro
165 170 175

Thr Pro Leu Leu Leu Ser Met Met Tyr Met Pro Phe Val Ser Leu Phe
180 185 190

Ala Gly Leu Ala Ile Phe Gly Lys Ala Met Leu Gln Ile Leu Val His
195 200 205

Arg Leu Gly Gln Ile Gly Gly Glu Glu Gln Ser Glu Glu Glu Arg Phe
210 215 220

Gln Arg Leu Ala Ser Cys Ile Ala Tyr His Thr Gln Val Met Arg Tyr
225 230 235 240

Val Trp Gln Leu Asn Lys Leu Val Ala Asn Ile Val Ala Val Glu Ala
245 250 255

Ile Ile Phe Gly Ser Ile Ile Cys Ser Leu Leu Phe Cys Leu Asn Ile

260 265 270

Ile Thr Ser Pro Thr Gln Val Ile Ser Ile Val Met Tyr Ile Leu Thr
275 280 285

Met Leu Tyr Val Leu Phe Thr Tyr Tyr Asn Arg Ala Asn Glu Ile Cys
290 295 300

Leu Glu Asn Asn Arg Val Ala Glu Ala Val Tyr Asn Val Pro Trp Tyr
305 310 315 320

Glu Ala Gly Thr Arg Phe Arg Lys Thr Leu Leu Ile Phe Leu Met Gln
325 330 335

Thr Gln His Pro Met Glu Ile Arg Val Gly Asn Val Tyr Pro Met Thr
340 345 350

Leu Ala Met Phe Gln Ser Leu Leu Asn Ala Ser Tyr Ser Tyr Phe Thr
355 360 365

Met Leu Arg Gly Val Thr Gly Lys
370 375

<210> 7
<211> 1305
<212> DNA
<213> Drosophila melanogaster DOR53

<400> 7
tttttcccc acataaaaga aaagccattg agcgagcggg ttaagtcccg agatgccttc 60
atttacttgg atcgggtgat gtggtccttt ggctggacag agcctgaaaa caaaaggtgg 120
atccttcctt ataaactgtg gttagcgttc gtgaacatag taatgctcat ccttctgccg 180
atctcgataa gcatcgagta cctccaccga tttaaaacct tctcggcggg ggagttcctt 240
agttccctcg agattggagt caacatgtac ggaagctctt ttaagtgcgc cttcaccttg 300
attggattca agaaaagaca ggaagctaag gttttactgg atcagctgga caagagatgc 360
cttagcgata aggagaggtc cactgttcat cgctatgtcg ccatgggaaa ctttttcgat 420
attttgtatc acattttta ctccaccttc gtggtaatga acttcccgta ttttctgctt 480
gagagacgcc atgcttggcg catgtacttt ccatatatcg attccgacga acagttttac 540
atctccagca tcgccgagtg ttttctgatg acggaggcca tctacatgga tctctgtacg 600
gacgtgtgtc ccttgatctc catgcttatg gctcgatgcc acatcagcct cctgaaacag 660
cgactgagaa atctccgatc gaagccagga aggaccgaag atgagtactt ggaggagctc 720
accgagtgca ttcgggatca tcgattgcta ttggactatg ttgacgcatt gcgacccgtc 780
ttttcgggaa ccatttttgt gcagttcctc ctgatcggta ctgtactggg tctctcaatg 840
ataaatctaa tgttcttctc gacattttgg actggtgtcg ccacttgcct ttttatgttc 900
gacgtgtcca tggagacgtt cccctttgc tatttgtgca acatgattat cgatgactgc 960
caggaaatgt ccaattgcct ctttcaatcg gactggacct ctgccgatcg tcgctacaaa 1020
tccactttgg tatactttct tcacaatctt cagcaaccca ttactctcac ggctggtgga 1080

gtgtttccta tttccatgca aacaaatttg gctatggtga agctggcatt ttctgtggtt 1140
acggtaatta agcaatttaa cttggccgaa aggtttcaat aagttgagag ggacgagctc 1200
tgctactatt atattatata ttatattata ttatatatat attattttat attatatatt 1260
gctgtaccct aataaatatt tagtaataaa aaaaaaaaaa aaaaa 1305

<210> 8
<211> 367
<212> PRT
<213> Translation DOR53

<400> 8
Met Trp Ser Phe Gly Trp Thr Glu Pro Glu Asn Lys Arg Trp Ile Leu
1 5 10 15

Pro Tyr Lys Leu Trp Leu Ala Phe Val Asn Ile Val Met Leu Ile Leu
20 25 30

Leu Pro Ile Ser Ile Ser Ile Glu Tyr Leu His Arg Phe Lys Thr Phe
35 40 45

Ser Ala Gly Glu Phe Leu Ser Ser Leu Glu Ile Gly Val Asn Met Tyr
50 55 60

Gly Ser Ser Phe Lys Cys Ala Phe Thr Leu Ile Gly Phe Lys Lys Arg
65 70 75 80

Gln Glu Ala Lys Val Leu Leu Asp Gln Leu Asp Lys Arg Cys Leu Ser
85 90 95

Asp Lys Glu Arg Ser Thr Val His Arg Tyr Val Ala Met Gly Asn Phe
100 105 110

Phe Asp Ile Leu Tyr His Ile Phe Tyr Ser Thr Phe Val Val Met Asn
115 120 125

Phe Pro Tyr Phe Leu Leu Glu Arg Arg His Ala Trp Arg Met Tyr Phe
130 135 140

Pro Tyr Ile Asp Ser Asp Glu Gln Phe Tyr Ile Ser Ser Ile Ala Glu
145 150 155 160

Cys Phe Leu Met Thr Glu Ala Ile Tyr Met Asp Leu Cys Thr Asp Val
165 170 175

Cys Pro Leu Ile Ser Met Leu Met Ala Arg Cys His Ile Ser Leu Leu
180 185 190

Lys Gln Arg Leu Arg Asn Leu Arg Ser Lys Pro Gly Arg Thr Glu Asp

195 200 205

Glu Tyr Leu Glu Glu Leu Thr Glu Cys Ile Arg Asp His Arg Leu Leu
210 215 220

Leu Asp Tyr Val Asp Ala Leu Arg Pro Val Phe Ser Gly Thr Ile Phe
225 230 235 240

Val Gln Phe Leu Leu Ile Gly Thr Val Leu Gly Leu Ser Met Ile Asn
245 250 255

Leu Met Phe Phe Ser Thr Phe Trp Thr Gly Val Ala Thr Cys Leu Phe
260 265 270

Met Phe Asp Val Ser Met Glu Thr Phe Pro Phe Cys Tyr Leu Cys Asn
275 280 285

Met Ile Ile Asp Asp Cys Gln Glu Met Ser Asn Cys Leu Phe Gln Ser
290 295 300

Asp Trp Thr Ser Ala Asp Arg Arg Tyr Lys Ser Thr Leu Val Tyr Phe
305 310 315 320

Leu His Asn Leu Gln Gln Pro Ile Thr Leu Thr Ala Gly Gly Val Phe
325 330 335

Pro Ile Ser Met Gln Thr Asn Leu Ala Met Val Lys Leu Ala Phe Ser
340 345 350

Val Val Thr Val Ile Lys Gln Phe Asn Leu Ala Glu Arg Phe Gln
355 360 365

<210> 9
<211> 1321
<212> DNA
<213> Drosophila melanogaster DOR67

<400> 9
ggcacgagga aatgttaagc cagttctttc cccacattaa agaaaagcca ttgagcgagc 60
gggttaagtc ccgagatgcc ttcgtttact tagatcgggt gatgtggtcc tttggctgga 120
cagtgcctga aaacaaaagg tgggatctac attacaaact gtggtcaact ttcgtgacat 180
tggtgatatt tatccttctg ccgatatcgg taagcgttga gtatattcag cggttcaaga 240
ccttctcggc gggtgagttt cttagctcaa tccagattgg cgttaacatg tacggaagca 300
gctttaaaag ttatttgacc atgatgggat ataagaagag acaggaggct aagatgtcac 360
tggatgagct ggacaagaga tgcgtttgtg atgaggagag gaccattgta catcgacatg 420
tcgccctggg aaacttttgc tatattttct atcacattgc gtacactagc tttttgattt 480
caaacttttt gtcatttata atgaagagaa tccatgcctg gcgcatgtac tttccctacg 540

```
tcgaccccga aaagcaattt tacatctcta gcatcgccga agtcattctt agggggtggg 600
ccgtcttcat ggatctctgc acggatgtgt gtcctttgat ctccatggta atagcacgat 660
gccacatcac ccttctgaaa cagcgcctgc gaaatctacg atcggaacca ggaaggacgg 720
aagatgagta cttgaaggag ctcgccgact gcgttcgaga tcaccgcttg atattggact 780
atgtcgacgc attgcgatcc gtcttttcgg ggacaatttt tgtgcagttc ctcttgatcg 840
gtattgtact gggtctgtca atgataaata taatgttttt ctcaacactt tcgactggtg 900
tcgccgttgt cctttttatg tcctgcgtat ctatgcagac gttccccttt tgctatttgt 960
gtaacatgat tatggatgac tgccaagaga tggccgactc cctttttcaa tcggactgga 1020
catctgccga tcgtcgctac aaatccactt tggtatactt tcttcacaat cttcagcagc 1080
ccattattct tacggctggt ggagtctttc ctatttccat gcaaacaaat ttaaatatgg 1140
tgaagctggc ctttactgtg gttacaatag tgaaacaatt taacttggca gaaaagtttc 1200
aataagttaa gatatgcaag ctctgctatt ataaacctac actcgagaaa atatttcttc 1260
acattaataa accttcagta cttactgctt gtggcgcccc cggaaaaaaa aaaaaaaaaa 1320
a                                                                 1321
```

<210> 10
<211> 367
<212> PRT
<213> Translation DOR67

<400> 10

```
Met Trp Ser Phe Gly Trp Thr Val Pro Glu Asn Lys Arg Trp Asp Leu
1               5                   10                  15

His Tyr Lys Leu Trp Ser Thr Phe Val Thr Leu Val Ile Phe Ile Leu
            20                  25                  30

Leu Pro Ile Ser Val Ser Val Glu Tyr Ile Gln Arg Phe Lys Thr Phe
        35                  40                  45

Ser Ala Gly Glu Phe Leu Ser Ser Ile Gln Ile Gly Val Asn Met Tyr
    50                  55                  60

Gly Ser Ser Phe Lys Ser Tyr Leu Thr Met Met Gly Tyr Lys Lys Arg
65                  70                  75                  80

Gln Glu Ala Lys Met Ser Leu Asp Glu Leu Asp Lys Arg Cys Val Cys
                85                  90                  95

Asp Glu Glu Arg Thr Ile Val His Arg His Val Ala Leu Gly Asn Phe
            100                 105                 110

Cys Tyr Ile Phe Tyr His Ile Ala Tyr Thr Ser Phe Leu Ile Ser Asn
        115                 120                 125

Phe Leu Ser Phe Ile Met Lys Arg Ile His Ala Trp Arg Met Tyr Phe
    130                 135                 140
```

Pro Tyr Val Asp Pro Glu Lys Gln Phe Tyr Ile Ser Ser Ile Ala Glu
145 150 155 160

Val Ile Leu Arg Gly Trp Ala Val Phe Met Asp Leu Cys Thr Asp Val
165 170 175

Cys Pro Leu Ile Ser Met Val Ile Ala Arg Cys His Ile Thr Leu Leu
180 185 190

Lys Gln Arg Leu Arg Asn Leu Arg Ser Glu Pro Gly Arg Thr Glu Asp
195 200 205

Glu Tyr Leu Lys Glu Leu Ala Asp Cys Val Arg Asp His Arg Leu Ile
210 215 220

Leu Asp Tyr Val Asp Ala Leu Arg Ser Val Phe Ser Gly Thr Ile Phe
225 230 235 240

Val Gln Phe Leu Leu Ile Gly Ile Val Leu Gly Leu Ser Met Ile Asn
245 250 255

Ile Met Phe Phe Ser Thr Leu Ser Thr Gly Val Ala Val Val Leu Phe
260 265 270

Met Ser Cys Val Ser Met Gln Thr Phe Pro Phe Cys Tyr Leu Cys Asn
275 280 285

Met Ile Met Asp Asp Cys Gln Glu Met Ala Asp Ser Leu Phe Gln Ser
290 295 300

Asp Trp Thr Ser Ala Asp Arg Arg Tyr Lys Ser Thr Leu Val Tyr Phe
305 310 315 320

Leu His Asn Leu Gln Gln Pro Ile Ile Leu Thr Ala Gly Gly Val Phe
325 330 335

Pro Ile Ser Met Gln Thr Asn Leu Asn Met Val Lys Leu Ala Phe Thr
340 345 350

Val Val Thr Ile Val Lys Gln Phe Asn Leu Ala Glu Lys Phe Gln
355 360 365

<210> 11
<211> 1308
<212> DNA
<213> Drosophila melanogaster DOR64

<400> 11

```
ggcacgagcc aagaattcaa aatgaaactc agcgaaaccc taaaaatcga ctattttcga 60
gtccagttga atgcctggcg aatttgtggt gccttggatc tcagcgaggg taggtactgg 120
agttggtcga tgctattgtg catcttggtg tacctgccga cacccatgct actgagagga 180
gtatacagtt tcgaggatcc ggtggaaaat aatttcagct tgagcctgac ggtcacatcg 240
ctgtccaatc tcatgaagtt ctgcatgtac gtggcccaac taacaaagat ggtcgaggtc 300
cagagtctta ttggtcagct ggatgcccgg gtttctggcg agagccagtc tgagcgtcat 360
agaaatatga ccgagcacct gctaaggatg tccaagctgt tccagatcac ctacgctgta 420
gtcttcatca ttgctgcagt tcccttcgtt ttcgaaactg agctaagctt acccatgccc 480
atgtggtttc ccttcgactg gaagaactcg atggtggcct acatcggagc tctggttttc 540
caggagattg gctatgtctt tcaaattatg caatgctttg cagctgactc gtttcccccg 600
ctcgtactgt acctgatctc cgagcaatgt caattgctga tcctgagaat ctctgaaatc 660
ggatatggtt acaagactct ggaggagaac gaacaggatc tggtcaactg catcagggat 720
caaaacgcgc tgtatagatt actcgatgtg accaagagtc tcgtttcgta tcccatgatg 780
gtgcagttta tggttattgg catcaacatc gccatcaccc tatttgtcct gatattttac 840
gtggagacct tgtacgatcg catctattat ctttgctttc tcttgggcat caccgtgcag 900
acatatccat tgtgctacta tggaaccatg gtgcaggaga gttttgctga gcttcactat 960
gcggtattct gcagcaactg ggtggatcaa agtgccagct atcgtgggca catgctcatc 1020
ctggcggagc gcactaagcg gatgcagctt ctcctcgccg gcaacctggt gcccatccac 1080
ctgagcacct acgtggcctg ttggaaggga gcctactcct tcttcaccct gatggccgat 1140
cgagatggcc tgggttctta gtagcccagt catttcactc acattctaca tcaagtagta 1200
ctaccactga acacgaacac gaatatttca aaagtaaaca cataatattc acaatagtgt 1260
atcactttaa taaaattttt ggttaccatg aaaaaaaaaa aaaaaaaa 1308
```

<210> 12
<211> 379
<212> PRT
<213> Translation DOR64

<400> 12

```
Met Lys Leu Ser Glu Thr Leu Lys Ile Asp Tyr Phe Arg Val Gln Leu
1               5                   10                  15

Asn Ala Trp Arg Ile Cys Gly Ala Leu Asp Leu Ser Glu Gly Arg Tyr
            20                  25                  30

Trp Ser Trp Ser Met Leu Leu Cys Ile Leu Val Tyr Leu Pro Thr Pro
        35                  40                  45

Met Leu Leu Arg Gly Val Tyr Ser Phe Glu Asp Pro Val Glu Asn Asn
    50                  55                  60

Phe Ser Leu Ser Leu Thr Val Thr Ser Leu Ser Asn Leu Met Lys Phe
65                  70                  75                  80

Cys Met Tyr Val Ala Gln Leu Thr Lys Met Val Glu Val Gln Ser Leu
                85                  90                  95
```

Ile Gly Gln Leu Asp Ala Arg Val Ser Gly Glu Ser Gln Ser Glu Arg
100 105 110

His Arg Asn Met Thr Glu His Leu Leu Arg Met Ser Lys Leu Phe Gln
115 120 125

Ile Thr Tyr Ala Val Val Phe Ile Ile Ala Ala Val Pro Phe Val Phe
130 135 140

Glu Thr Glu Leu Ser Leu Pro Met Pro Met Trp Phe Pro Phe Asp Trp
145 150 155 160

Lys Asn Ser Met Val Ala Tyr Ile Gly Ala Leu Val Phe Gln Glu Ile
165 170 175

Gly Tyr Val Phe Gln Ile Met Gln Cys Phe Ala Ala Asp Ser Phe Pro
180 185 190

Pro Leu Val Leu Tyr Leu Ile Ser Glu Gln Cys Gln Leu Leu Ile Leu
195 200 205

Arg Ile Ser Glu Ile Gly Tyr Gly Tyr Lys Thr Leu Glu Glu Asn Glu
210 215 220

Gln Asp Leu Val Asn Cys Ile Arg Asp Gln Asn Ala Leu Tyr Arg Leu
225 230 235 240

Leu Asp Val Thr Lys Ser Leu Val Ser Tyr Pro Met Met Val Gln Phe
245 250 255

Met Val Ile Gly Ile Asn Ile Ala Ile Thr Leu Phe Val Leu Ile Phe
260 265 270

Tyr Val Glu Thr Leu Tyr Asp Arg Ile Tyr Tyr Leu Cys Phe Leu Leu
275 280 285

Gly Ile Thr Val Gln Thr Tyr Pro Leu Cys Tyr Tyr Gly Thr Met Val
290 295 300

Gln Glu Ser Phe Ala Glu Leu His Tyr Ala Val Phe Cys Ser Asn Trp
305 310 315 320

Val Asp Gln Ser Ala Ser Tyr Arg Gly His Met Leu Ile Leu Ala Glu
325 330 335

Arg Thr Lys Arg Met Gln Leu Leu Leu Ala Gly Asn Leu Val Pro Ile
340 345 350

His Leu Ser Thr Tyr Val Ala Cys Trp Lys Gly Ala Tyr Ser Phe Phe
355 360 365

Thr Leu Met Ala Asp Arg Asp Gly Leu Gly Ser
370 375

<210> 13
<211> 1252
<212> DNA
<213> Drosophila melanogaster DOR71g

<400> 13
atggtcatta tcgacagtct tagtttttat cgtccattct ggatctgcat gcgattgctg 60
gtaccgactt tcttcaagga ttcctcacgt cctgtccagc tgtacgtggt gttgctgcac 120
atcctggtca ccttgtggtt tccactgcat ctgctgctgc atcttctgct acttccatct 180
accgctgagt tctttaagaa cctgaccatg tctctgactt gtgtggcctg cagtctgaag 240
catgtggccc acttgtatca cttgccgcag attgtggaaa tcgaatcact gatcgagcaa 300
ttagacacat ttattgccag cgaacaggag catcgttact atcgggatca cgtacattgc 360
catgctaggc gctttacaag atgtctctat attagctttg gcatgatcta tgcgcttttc 420
ctgttcggcg tcttcgttca ggttattagc ggaaattggg aacttctcta tccagcctat 480
ttcccattcg acttggagag caatcgcttt ctcggcgcag tagccttggg ctatcaggta 540
ttcagcatgt tagttgaagg cttccagggg ctgggcaacg atacctatac cccactgacc 600
ctatgccttc tggccggaca tgtccatttg tggtccatac gaatgggtca actgggatac 660
ttcgatgacg agacggtggt gaatcatcag cgtttgctgg attacattga gcagcataaa 720
ctcttggtgc ggtaagcttt gattaactaa cttttgacaa gaagtttatt cactttaact 780
ggttccaaaa acgatgcact caatgtgcag attccacaac ctggtgagcc ggaccatcag 840
cgaagtgcaa ctggtgcagc tgggcggatg tggagccact ctgtgcatca ttgtctccta 900
catgctcttc tttgtgggcg acacaatctc gctggtctac tacttggtgt tctttggagt 960
ggtctgcgtg cagctctttc ccagctgcta ttttgccagc gaagtagccg aggagttgga 1020
acggctgcca tatgcgatct tctccagcag atggtacgat caatcgcggg atcatcgatt 1080
cgatttgctc atctttacac aattaacact gggaaaccgg gggtggatca tcaaggcagg 1140
aggtcttatc gagctgaatt tgaatgcctt tttcgccacc ctgaagatgg cctattccct 1200
ttttgcagtt gtggtgcggg caaagggtat atagagagtc tgtttaatta aa 1252

<210> 14
<211> 384
<212> PRT
<213> Translation DOR71g

<400> 14
Met Val Ile Ile Asp Ser Leu Ser Phe Tyr Arg Pro Phe Trp Ile Cys
1 5 10 15

Met Arg Leu Leu Val Pro Thr Phe Phe Lys Asp Ser Ser Arg Pro Val
20 25 30

Gln Leu Tyr Val Val Leu Leu His Ile Leu Val Thr Leu Trp Phe Pro

35 40 45

Leu His Leu Leu Leu His Leu Leu Leu Leu Pro Ser Thr Ala Glu Phe
50 55 60

Phe Lys Asn Leu Thr Met Ser Leu Thr Cys Val Ala Cys Ser Leu Lys
65 70 75 80

His Val Ala His Leu Tyr His Leu Pro Gln Ile Val Glu Ile Glu Ser
85 90 95

Leu Ile Glu Gln Leu Asp Thr Phe Ile Ala Ser Glu Gln Glu His Arg
100 105 110

Tyr Tyr Arg Asp His Val His Cys His Ala Arg Arg Phe Thr Arg Cys
115 120 125

Leu Tyr Ile Ser Phe Gly Met Ile Tyr Ala Leu Phe Leu Phe Gly Val
130 135 140

Phe Val Gln Val Ile Ser Gly Asn Trp Glu Leu Leu Tyr Pro Ala Tyr
145 150 155 160

Phe Pro Phe Asp Leu Glu Ser Asn Arg Phe Leu Gly Ala Val Ala Leu
165 170 175

Gly Tyr Gln Val Phe Ser Met Leu Val Glu Gly Phe Gln Gly Leu Gly
180 185 190

Asn Asp Thr Tyr Thr Pro Leu Thr Leu Cys Leu Leu Ala Gly His Val
195 200 205

His Leu Trp Ser Ile Arg Met Gly Gln Leu Gly Tyr Phe Asp Asp Glu
210 215 220

Thr Val Val Asn His Gln Arg Leu Leu Asp Tyr Ile Glu Gln His Lys
225 230 235 240

Leu Leu Val Arg Phe His Asn Leu Val Ser Arg Thr Ile Ser Glu Val
245 250 255

Gln Leu Val Gln Leu Gly Gly Cys Gly Ala Thr Leu Cys Ile Ile Val
260 265 270

Ser Tyr Met Leu Phe Phe Val Gly Asp Thr Ile Ser Leu Val Tyr Tyr
275 280 285

Leu Val Phe Phe Gly Val Val Cys Val Gln Leu Phe Pro Ser Cys Tyr

290 295 300

Phe Ala Ser Glu Val Ala Glu Glu Leu Glu Arg Leu Pro Tyr Ala Ile
305 310 315 320

Phe Ser Ser Arg Trp Tyr Asp Gln Ser Arg Asp His Arg Phe Asp Leu
325 330 335

Leu Ile Phe Thr Gln Leu Thr Leu Gly Asn Arg Gly Trp Ile Ile Lys
340 345 350

Ala Gly Gly Leu Ile Glu Leu Asn Leu Asn Ala Phe Phe Ala Thr Leu
355 360 365

Lys Met Ala Tyr Ser Leu Phe Ala Val Val Val Arg Ala Lys Gly Ile
370 375 380

<210> 15
<211> 1321
<212> DNA
<213> Drosophila melanogaster DOR72g

<400> 15
atggacttaa aaccgcgagt cattcgaagt gaagatatct acagaaccta ttggttatat 60
tggcatcttt tgggcctgga aagcaatttc tttctgaatc gcttgttgga tttggtgatt 120
acaattttcg taaccatttg gtatccaatt cacctgattc tgggactgtt tatggaaaga 180
tctttggggg atgtctgcaa gggtctacca attacggcag catgcttttt cgccagcttt 240
aaatttattt gttttcgctt caagctatct gaaattaaag aaatcgaaat attatttaaa 300
gagctggatc agcgagcttt aagtcgagag gaatgcgagt tttcaatca aaatacgaga 360
cgtgaggcga atttcatttg gaaaagtttc attgtggcct atggactgtc gaatatctcg 420
gctattgcat cagttctttt cggcggtgga cataagctat tatatcccgc ctggtttcca 480
tacgatgtgc aggccacgga actaatattt tggctaagtg taacatacca aattgccgga 540
gtaagtttgg ccatacttca gaatttggcc aatgattcct atccaccgat gacattttgc 600
gtggttgccg gtcatgtaag acttttggcg atgcgcttga gtagaattgg ccaaggtcca 660
gaggaaacaa tatacttaac cggaaagcaa ttaatcgaaa gcatcgagga tcaccgaaaa 720
ctaatgaagt aatgtacata tatagaatgg tttttagtta ttatcattaa atgaacgtgt 780
tgtaggaaaa ccattctgtt tgtcgggtgt cacggaaatc gattttcctt aatttacata 840
tgatattaaa tacttccttg caaacaatta tcatattagt aatttagaat ctttattatt 900
tatttccaga atagtggaat tactgcgcag caccatgaat atttcgcagc tcggccagtt 960
tatttcaagt ggtgttaata tttccataac actagtcaac attctcttct ttgcggataa 1020
taatttcgct ataacctact acggagtgta cttcctatcg atggtgttgg aattattccc 1080
gtgctgctat tacggcaccc tgatatccgt ggagatgaac cagctgacct atgcgattta 1140
ctcaagtaac tggatgagta tgaatcggag ctacagccgc atcctactga tcttcatgca 1200
actcaccctg gcggaagtgc agatcaaggc cggtgggatg attggcatcg gaatgaacgc 1260

cttctttgcc accgtgcgat tggcctactc cttcttcact ttggccatgt cgctgcgtta 1320
a 1321

<210> 16
<211> 379
<212> PRT
<213> Translation DOR72g

<400> 16
Met Asp Leu Lys Pro Arg Val Ile Arg Ser Glu Asp Ile Tyr Arg Thr
1 5 10 15

Tyr Trp Leu Tyr Trp His Leu Leu Gly Leu Glu Ser Asn Phe Phe Leu
20 25 30

Asn Arg Leu Leu Asp Leu Val Ile Thr Ile Phe Val Thr Ile Trp Tyr
35 40 45

Pro Ile His Leu Ile Leu Gly Leu Phe Met Glu Arg Ser Leu Gly Asp
50 55 60

Val Cys Lys Gly Leu Pro Ile Thr Ala Ala Cys Phe Phe Ala Ser Phe
65 70 75 80

Lys Phe Ile Cys Phe Arg Phe Lys Leu Ser Glu Ile Lys Glu Ile Glu
85 90 95

Ile Leu Phe Lys Glu Leu Asp Gln Arg Ala Leu Ser Arg Glu Glu Cys
100 105 110

Glu Phe Phe Asn Gln Asn Thr Arg Arg Glu Ala Asn Phe Ile Trp Lys
115 120 125

Ser Phe Ile Val Ala Tyr Gly Leu Ser Asn Ile Ser Ala Ile Ala Ser
130 135 140

Val Leu Phe Gly Gly Gly His Lys Leu Leu Tyr Pro Ala Trp Phe Pro
145 150 155 160

Tyr Asp Val Gln Ala Thr Glu Leu Ile Phe Trp Leu Ser Val Thr Tyr
165 170 175

Gln Ile Ala Gly Val Ser Leu Ala Ile Leu Gln Asn Leu Ala Asn Asp
180 185 190

Ser Tyr Pro Pro Met Thr Phe Cys Val Val Ala Gly His Val Arg Leu
195 200 205

Leu Ala Met Arg Leu Ser Arg Ile Gly Gln Gly Pro Glu Glu Thr Ile
210 215 220

Tyr Leu Thr Gly Lys Gln Leu Ile Glu Ser Ile Glu Asp His Arg Lys
225 230 235 240

Leu Met Lys Ile Val Glu Leu Leu Arg Ser Thr Met Asn Ile Ser Gln
245 250 255

Leu Gly Gln Phe Ile Ser Ser Gly Val Asn Ile Ser Ile Thr Leu Val
260 265 270

Asn Ile Leu Phe Phe Ala Asp Asn Asn Phe Ala Ile Thr Tyr Tyr Gly
275 280 285

Val Tyr Phe Leu Ser Met Val Leu Glu Leu Phe Pro Cys Cys Tyr Tyr
290 295 300

Gly Thr Leu Ile Ser Val Glu Met Asn Gln Leu Thr Tyr Ala Ile Tyr
305 310 315 320

Ser Ser Asn Trp Met Ser Met Asn Arg Ser Tyr Ser Arg Ile Leu Leu
325 330 335

Ile Phe Met Gln Leu Thr Leu Ala Glu Val Gln Ile Lys Ala Gly Gly
340 345 350

Met Ile Gly Ile Gly Met Asn Ala Phe Phe Ala Thr Val Arg Leu Ala
355 360 365

Tyr Ser Phe Phe Thr Leu Ala Met Ser Leu Arg
370 375

<210> 17
<211> 1212
<212> DNA
<213> Drosophila melanogaster DOR73g

<400> 17
atggattcaa gaaggaaagt ccgaagtgaa aatctttaca aaacctattg gctttactgg 60
cgacttctgg gagtcgaggg cgattatcct tttcgacggc tagtggattt tacaatcacg 120
tctttcatta cgattttatt tcccgtgcat cttatactgg gaatgtataa aaagccccag 180
attcaagtct tcaggagtct gcatttcaca tcggaatgcc ttttctgcag ctataagttt 240
ttctgttttc gttggaaact taaagaaata aagaccatcg aaggattgct ccaggatctc 300
gatagtcgag ttgaaagtga agaagaacgc aactacttta atcaaaatcc aagtcgtgtg 360
gctcgaatgc tttcgaaaag ttacttggta gctgctatat cggccataat cactgcaact 420
gtagctggtt tatttagtac tggtcgaaat ttaatgtatc tgggttggtt tccctacgat 480

```
tttcaagcaa ccgccgcaat ctattggatt agtttttcct atcaggcgat tggctctagt  540
ctgttgattc tggaaaatct ggccaacgat tcatatccgc cgattacatt ttgtgtggtc  600
tctggacatg tgagactatt gataatgcgt ttaagtcgaa ttggtcacga tgtaaaatta  660
tcaagttcgg aaaataccag aaaactcatc gaaggtatcc aggatcacag gaaactaatg  720
aagtaagaat aaagatttaa gaaccgcatg tttgatagct cagagaactg ataattaatc  780
aaatgtaact tttccaggat aatacgccta cttcgcagca ctttacatct tagccaactg  840
ggccagttcc tttctagtgg aatcaacatt tccataacac tcatcaacat cctgttcttt  900
gcggaaaaca actttgcaat gctttattat gcggtgttct ttgctgcaat gttaatagaa  960
ctatttccaa gttgttacta tggaattctg atgacaatgg agtttgataa gctaccatat 1020
gccatcttct ccagcaactg gcttaaaatg gataaaagat acaatcgatc cttgataatt 1080
ctgatgcaac taacactggt tccagtgaat ataaaagcag gtggtattgt tggcatcgat 1140
atgagtgcat tttttgccac agttcggatg gcatattcct tttacacttt agccttgtca 1200
tttcgagtat ag                                                     1212
```

<210> 18
<211> 378
<212> PRT
<213> Translation DOR73g

<400> 18

```
Met Asp Ser Arg Arg Lys Val Arg Ser Glu Asn Leu Tyr Lys Thr Tyr
1               5                   10                  15

Trp Leu Tyr Trp Arg Leu Leu Gly Val Glu Gly Asp Tyr Pro Phe Arg
            20                  25                  30

Arg Leu Val Asp Phe Thr Ile Thr Ser Phe Ile Thr Ile Leu Phe Pro
        35                  40                  45

Val His Leu Ile Leu Gly Met Tyr Lys Lys Pro Gln Ile Gln Val Phe
    50                  55                  60

Arg Ser Leu His Phe Thr Ser Glu Cys Leu Phe Cys Ser Tyr Lys Phe
65                  70                  75                  80

Phe Cys Phe Arg Trp Lys Leu Lys Glu Ile Lys Thr Ile Glu Gly Leu
                85                  90                  95

Leu Gln Asp Leu Asp Ser Arg Val Glu Ser Glu Glu Glu Arg Asn Tyr
            100                 105                 110

Phe Asn Gln Asn Pro Ser Arg Val Ala Arg Met Leu Ser Lys Ser Tyr
        115                 120                 125

Leu Val Ala Ala Ile Ser Ala Ile Ile Thr Ala Thr Val Ala Gly Leu
    130                 135                 140

Phe Ser Thr Gly Arg Asn Leu Met Tyr Leu Gly Trp Phe Pro Tyr Asp
```

145 150 155 160

Phe Gln Ala Thr Ala Ala Ile Tyr Trp Ile Ser Phe Ser Tyr Gln Ala
165 170 175

Ile Gly Ser Ser Leu Leu Ile Leu Glu Asn Leu Ala Asn Asp Ser Tyr
180 185 190

Pro Pro Ile Thr Phe Cys Val Val Ser Gly His Val Arg Leu Leu Ile
195 200 205

Met Arg Leu Ser Arg Ile Gly His Asp Val Lys Leu Ser Ser Ser Glu
210 215 220

Asn Thr Arg Lys Leu Ile Glu Gly Ile Gln Asp His Arg Lys Leu Met
225 230 235 240

Lys Ile Ile Arg Leu Leu Arg Ser Thr Leu His Leu Ser Gln Leu Gly
245 250 255

Gln Phe Leu Ser Ser Gly Ile Asn Ile Ser Ile Thr Leu Ile Asn Ile
260 265 270

Leu Phe Phe Ala Glu Asn Asn Phe Ala Met Leu Tyr Tyr Ala Val Phe
275 280 285

Phe Ala Ala Met Leu Ile Glu Leu Phe Pro Ser Cys Tyr Tyr Gly Ile
290 295 300

Leu Met Thr Met Glu Phe Asp Lys Leu Pro Tyr Ala Ile Phe Ser Ser
305 310 315 320

Asn Trp Leu Lys Met Asp Lys Arg Tyr Asn Arg Ser Leu Ile Ile Leu
325 330 335

Met Gln Leu Thr Leu Val Pro Val Asn Ile Lys Ala Gly Gly Ile Val
340 345 350

Gly Ile Asp Met Ser Ala Phe Phe Ala Thr Val Arg Met Ala Tyr Ser
355 360 365

Phe Tyr Thr Leu Ala Leu Ser Phe Arg Val
370 375

<210> 19
<211> 1198
<212> DNA

<213> Drosophila melanogaster DOR46

<400> 19

```
atggcagagg tcagagtgga cagtctggag tttttcaaga gccattggac cgcctggcgg 60
tacttgggag tggctcattt tcgggtcgag aactggaaga acctttacgt gttttacagc 120
attgtgtcga atcttctcgt gaccctgtgc taccccgttc acctgggaat atccctcttt 180
cgcaaccgca ccatcaccga ggacatcctc aacctgacca cctttgcgac ctgcacagcc 240
tgttcggtga agtgcctgct ctacgcctac aacatcaagg atgtgctgga gatggagcgg 300
ctgttgaggc ttttggatga acgcgtcgtg ggtccggagc aacgcagcat ctacggacaa 360
gtgagggtcc agctgcgaaa tgtgctatac gtgttcatcg gcatctacat gccgtgtgcc 420
ctgttcgccg agctatcctt tctgttcaag gaggagcgcg gtctgatgta tcccgcctgg 480
tttcccttcg actggctgca ctccaccagg aactattaca tagcgaacgc ctatcagata 540
gtgggcatct cgtttcagct gctgcaaaac tatgttagcg actgctttcc ggcggtggtg 600
ctgtgcctga tctcatccca catcaaaatg ttgtacaaca gattcgagga ggtgggcctg 660
gatccagcca gagatgcgga gaaggacctg gaggcctgca tcaccgatca caagcatatt 720
ctagagtggg caggcggctc attgtaacgt tcgtgttcta ttcactttcc aactttttc 780
cagactattc cgacgcatcg aggccttcat ttccctgccc atgctaattc agttcacagt 840
gaccgccttg aatgtgtgca tcggtttagc agccctggtg tttttcgtca gcgagcccat 900
ggcacggatg tacttcatct tctactccct ggccatgccg ctgcagatct ttccgtcctg 960
ctttttcggc accgacaacg agtactggtt cggacgcctc cactacgcgg ccttcagttg 1020
caattggcac acacagaaca ggagctttaa gcggaaaatg atgctgttcg ttgagcaatc 1080
gttgaagaag agcaccgctg tggctggcgg aatgatgcgt atccacctgg acacgttctt 1140
ttccacccta aagggggcct actccctctt taccatcatt attcggatga gaaagtag   1198
```

<210> 20
<211> 379
<212> PRT
<213> Translation DOR46

<400> 20

```
Met Ala Glu Val Arg Val Asp Ser Leu Glu Phe Phe Lys Ser His Trp
1               5                   10                  15

Thr Ala Trp Arg Tyr Leu Gly Val Ala His Phe Arg Val Glu Asn Trp
            20                  25                  30

Lys Asn Leu Tyr Val Phe Tyr Ser Ile Val Ser Asn Leu Leu Val Thr
        35                  40                  45

Leu Cys Tyr Pro Val His Leu Gly Ile Ser Leu Phe Arg Asn Arg Thr
    50                  55                  60

Ile Thr Glu Asp Ile Leu Asn Leu Thr Thr Phe Ala Thr Cys Thr Ala
65                  70                  75                  80

Cys Ser Val Lys Cys Leu Leu Tyr Ala Tyr Asn Ile Lys Asp Val Leu
                85                  90                  95
```

Glu Met Glu Arg Leu Leu Arg Leu Leu Asp Glu Arg Val Val Gly Pro
100 105 110

Glu Gln Arg Ser Ile Tyr Gly Gln Val Arg Val Gln Leu Arg Asn Val
115 120 125

Leu Tyr Val Phe Ile Gly Ile Tyr Met Pro Cys Ala Leu Phe Ala Glu
130 135 140

Leu Ser Phe Leu Phe Lys Glu Glu Arg Gly Leu Met Tyr Pro Ala Trp
145 150 155 160

Phe Pro Phe Asp Trp Leu His Ser Thr Arg Asn Tyr Tyr Ile Ala Asn
165 170 175

Ala Tyr Gln Ile Val Gly Ile Ser Phe Gln Leu Leu Gln Asn Tyr Val
180 185 190

Ser Asp Cys Phe Pro Ala Val Val Leu Cys Leu Ile Ser Ser His Ile
195 200 205

Lys Met Leu Tyr Asn Arg Phe Glu Glu Val Gly Leu Asp Pro Ala Arg
210 215 220

Asp Ala Glu Lys Asp Leu Glu Ala Cys Ile Thr Asp His Lys His Ile
225 230 235 240

Leu Glu Leu Phe Arg Arg Ile Glu Ala Phe Ile Ser Leu Pro Met Leu
245 250 255

Ile Gln Phe Thr Val Thr Ala Leu Asn Val Cys Ile Gly Leu Ala Ala
260 265 270

Leu Val Phe Phe Val Ser Glu Pro Met Ala Arg Met Tyr Phe Ile Phe
275 280 285

Tyr Ser Leu Ala Met Pro Leu Gln Ile Phe Pro Ser Cys Phe Phe Gly
290 295 300

Thr Asp Asn Glu Tyr Trp Phe Gly Arg Leu His Tyr Ala Ala Phe Ser
305 310 315 320

Cys Asn Trp His Thr Gln Asn Arg Ser Phe Lys Arg Lys Met Met Leu
325 330 335

Phe Val Glu Gln Ser Leu Lys Lys Ser Thr Ala Val Ala Gly Gly Met
340 345 350

Met Arg Ile His Leu Asp Thr Phe Phe Ser Thr Leu Lys Gly Ala Tyr
355 360 365

Ser Leu Phe Thr Ile Ile Ile Arg Met Arg Lys
370 375

<210> 21
<211> 1293
<212> DNA
<213> Drosophila melanogaster DOR19g

<400> 21
atggttacgg aggactttta taagtaccag gtgtggtact tccaaatcct tggtgtttgg 60
cagctcccca cttgggccgc agaccaccag cgtcgttttc agtccatgag gtttggcttc 120
atcctggtca tcctgttcat catgctgctg cttttctcct tcgaaatgtt gaacaacatt 180
tcccaagtta gggagatcct aaaggtattc ttcatgttcg ccacggaaat atcctgcatg 240
gccaaattat tgcatttgaa gttgaagagc cgcaaactcg ctggcttggt tgatgcgatg 300
ttgtccccag agttcggcgt taaaagtgaa caggaaatgc agatgctgga attggataga 360
gtggcggttg tccgcatgag gaactcctac ggcatcatgt ccctgggcgc ggcttccctg 420
atccttatag ttccctgttt cgacaacttt ggcgagctac cactggccat gttggaggta 480
tgcagcatcg agggatggat ctgctattgg tcgcagtacc ttttccactc gatttgcctg 540
ctgcccactt gtgtgctgaa tataacctac gactcggtgg cctactcgtt gctctgtttc 600
ttgaaggttc agctacaaat gctggtcctg cgattagaaa agttgggtcc tgtgatcgaa 660
ccccaggata atgagaaaat cgcaatggaa ctgcgtgagt gtgccgccta ctacaacagg 720
attgttcgtt tcaaggacct ggtggagctg ttcataaagg ggccaggatc tgtgcagctc 780
atgtgttctg ttctggtgct ggtgtccaac ctgtacgaca tgtccaccat gtccattgca 840
aacggcgatg ccatctttat gctcaagacc tgtatctatc agctggtgat gctctggcag 900
atcttcatca tttgctacgc ctccaacgag gtaactgtcc agagctctag gttgtgtcac 960
agcatctaca gctcccaatg gacgggatgg aacagggcaa accgccggat tgtccttctc 1020
atgatgcagc gctttaattc cccgatgctc ctgagcacct ttaaccccac ctttgctttc 1080
agcttggagg cctttggttc tgtagggcag cagaaattcc tttatatatc atttattact 1140
ggttatgctc ttctcctttc agatcgtcaa ctgctcctac agctacttcg cactgctgaa 1200
gcgcgtcaac agttaaattt cgaaacaccg cagcacctaa agattttcaa gccgattttt 1260
aaaagcactc aaaacgttat gcacgtacat taa 1293

<210> 22
<211> 430
<212> PRT
<213> Translation DOR19g

<400> 22
Met Val Thr Glu Asp Phe Tyr Lys Tyr Gln Val Trp Tyr Phe Gln Ile
1 5 10 15

Leu Gly Val Trp Gln Leu Pro Thr Trp Ala Ala Asp His Gln Arg Arg
20 25 30

Phe Gln Ser Met Arg Phe Gly Phe Ile Leu Val Ile Leu Phe Ile Met
35 40 45

Leu Leu Leu Phe Ser Phe Glu Met Leu Asn Asn Ile Ser Gln Val Arg
50 55 60

Glu Ile Leu Lys Val Phe Phe Met Phe Ala Thr Glu Ile Ser Cys Met
65 70 75 80

Ala Lys Leu Leu His Leu Lys Leu Lys Ser Arg Lys Leu Ala Gly Leu
85 90 95

Val Asp Ala Met Leu Ser Pro Glu Phe Gly Val Lys Ser Glu Gln Glu
100 105 110

Met Gln Met Leu Glu Leu Asp Arg Val Ala Val Val Arg Met Arg Asn
115 120 125

Ser Tyr Gly Ile Met Ser Leu Gly Ala Ala Ser Leu Ile Leu Ile Val
130 135 140

Pro Cys Phe Asp Asn Phe Gly Glu Leu Pro Leu Ala Met Leu Glu Val
145 150 155 160

Cys Ser Ile Glu Gly Trp Ile Cys Tyr Trp Ser Gln Tyr Leu Phe His
165 170 175

Ser Ile Cys Leu Leu Pro Thr Cys Val Leu Asn Ile Thr Tyr Asp Ser
180 185 190

Val Ala Tyr Ser Leu Leu Cys Phe Leu Lys Val Gln Leu Gln Met Leu
195 200 205

Val Leu Arg Leu Glu Lys Leu Gly Pro Val Ile Glu Pro Gln Asp Asn
210 215 220

Glu Lys Ile Ala Met Glu Leu Arg Glu Cys Ala Ala Tyr Tyr Asn Arg
225 230 235 240

Ile Val Arg Phe Lys Asp Leu Val Glu Leu Phe Ile Lys Gly Pro Gly
245 250 255

Ser Val Gln Leu Met Cys Ser Val Leu Val Leu Val Ser Asn Leu Tyr
260 265 270

Asp Met Ser Thr Met Ser Ile Ala Asn Gly Asp Ala Ile Phe Met Leu
275 280 285

Lys Thr Cys Ile Tyr Gln Leu Val Met Leu Trp Gln Ile Phe Ile Ile
290 295 300

Cys Tyr Ala Ser Asn Glu Val Thr Val Gln Ser Ser Arg Leu Cys His
305 310 315 320

Ser Ile Tyr Ser Ser Gln Trp Thr Gly Trp Asn Arg Ala Asn Arg Arg
325 330 335

Ile Val Leu Leu Met Met Gln Arg Phe Asn Ser Pro Met Leu Leu Ser
340 345 350

Thr Phe Asn Pro Thr Phe Ala Phe Ser Leu Glu Ala Phe Gly Ser Val
355 360 365

Gly Gln Gln Lys Phe Leu Tyr Ile Ser Phe Ile Thr Gly Tyr Ala Leu
370 375 380

Leu Leu Ser Asp Arg Gln Leu Leu Leu Gln Leu Leu Arg Thr Ala Glu
385 390 395 400

Ala Arg Gln Gln Leu Asn Phe Glu Thr Pro Gln His Leu Lys Ile Phe
405 410 415

Lys Pro Ile Phe Lys Ser Thr Gln Asn Val Met His Val His
420 425 430

<210> 23
<211> 2075
<212> DNA
<213> Drosophila melanogaster DOR24g

<400> 23
atgcgcccaa cactccaggt cctcatcagt gtcctttgtc tggtctctgc ctacgcctgg 60
gatcataccg actgtaacga tcactatatc gaattcatgg actatcccga tgagagagct 120
accgcctata gcaatgaatc ctcagaatgg gatttctttg aattttggag gcaagtgttt 180
ggcctgtagc aattctgaca tcgctgacca cgcctctggt caaagcaatc acttggcgag 240
acatgtgacc ctcggatgtc agggcgtgaa taaagaacat taaaatttaa ccggagttct 300
catttactat gcctaaagag gtgtcaaacg gaggtattat ttgggcgtga tatattatac 360
cctctaaagg agtatcccca ccaggtgaaa ccctataaaa cccttgacgt cgttaatgga 420
aagtactgaa ccggaaatga aagtgggcac tttctaatgg tagaaatttg tcgggatgaa 480
atatataagt agatatatat aatattatat cttgattatc cagcatcaat cacaagataa 540
aaaaaaaata taattcataa ttcataattc atatgtaata ggcatttgta aatgttgtaa 600
acaaaggaag ggtattaaag agctgtactt tcgcatctgt atttcattta ttcctactca 660
attgaaaggt ttaaaacccc tgaactgaac acacttgact tagtgtgagg ccgaattaac 720
ccttgtcgac atggacagtt ttctgcaagt acagaagagc accattgccc ttctgggctt 780
tgatctcttt agtgaaaatc gagaaatgtg gaaacgcccc tatagagcaa tgaatgtgtt 840

tagcatagct gccatttttc cctttatcct ggcagctgtg ctccataatt ggaagaatgt 900
attgctgctg gccgatgcca tggtggccct actaataacc attctgggcc tattcaagtt 960
tagcatgata ctttacttac gtcgcgattt caagcgactg attgacaaat ttcgtttgct 1020
catgtcgaat ggtgagttgt aatccatttc ggccagaatg tgtatcattt catttattat 1080
tttatagagg cggaacaggg cgaggaatac gccgagattc tcaacgcagc aaacaagcag 1140
gatcaacgaa tgtgcactct gtttaggact tgtttcctcc tcgcctgggc cttgaatagt 1200
gttctgcccc tcgtgagaat gggtctcagc tattggttag caggtcatgc agagcccgag 1260
ttgccttttc cctgtctgta tgtacaaatg atatatatga tatatggtga tcaagttatc 1320
aggctttgtt cctaaagttt tccctggaat atccacatca ttcgcaatta tgttttgagc 1380
ttcatctgga gcgctttcgc ctcgacaggt gtggttttac ctgctgtcag cttggatacc 1440
atattctgtt ccttcaccag caacctgtgc gccttcttca aaattgcgca gtacaaggtg 1500
gttagattta agggcggatc ccttaaagaa tcacaggcca cattgaacaa agtctttgcc 1560
ctgtaccaga ccagcttgga tatgtgcaac gatctgaatc agtgctacca accgattatc 1620
tgcgcccagt tcttcatttc atctctgcaa ctctgcatgc tgggatatct gttctccatt 1680
acttttgccc agacagaggg cgtctactat gcctcattca tagccacaat cattatacaa 1740
gcctatatct actgctactg cggggagaac ctgaagacgg agagtgccag cttcgagtgg 1800
gccatctacg acagtccgtg gcacgagagt ttgggtgctg gtggagcctc tacctcgatc 1860
tgccgatcct tgctgatcag catgatgcgg gctcatcggg gattccgcat tacgggatac 1920
tttttcgagg caaacatgga ggccttctca tcggtgggtg aatcatttcc attgtacaat 1980
acatggattt acatgaatac tcttttctaa ctttccgttt ttagattgtt cgcacggcga 2040
tgtcctacat cacaatgctg agatcattct cctaa 2075

<210> 24
<211> 383
<212> PRT
<213> Drosophila melanogaster DOR24g

<400> 24
Met Arg Pro Thr Leu Gln Val Leu Ile Ser Val Leu Cys Leu Val Ser
1 5 10 15

Ala Tyr Ala Trp Asp His Thr Asp Cys Asn Asp His Tyr Ile Glu Phe
20 25 30

Met Asp Tyr Pro Asp Glu Arg Ala Thr Ala Tyr Ser Asn Glu Ser Ser
35 40 45

Glu Trp Asp Phe Phe Glu Phe Trp Arg Gln Val Phe Gly Leu Phe Leu
50 55 60

Gln Val Gln Lys Ser Thr Ile Ala Leu Leu Gly Phe Asp Leu Phe Ser
65 70 75 80

Glu Asn Arg Glu Met Trp Lys Arg Pro Tyr Arg Ala Met Asn Val Phe
85 90 95

Ser Ile Ala Ala Ile Phe Pro Phe Ile Leu Ala Ala Val Leu His Asn
100 105 110

Trp Lys Asn Val Leu Leu Leu Ala Asp Ala Met Val Ala Leu Leu Ile
115 120 125

Thr Ile Leu Gly Leu Phe Lys Phe Ser Met Ile Leu Tyr Leu Arg Arg
130 135 140

Asp Phe Lys Arg Leu Ile Asp Lys Phe Arg Leu Leu Met Ser Asn Gly
145 150 155 160

Glu Phe Phe Pro Trp Asn Ile His Ile Ile Arg Asn Tyr Val Leu Ser
165 170 175

Phe Ile Trp Ser Ala Phe Ala Ser Thr Gly Val Val Leu Pro Ala Val
180 185 190

Ser Leu Asp Thr Ile Phe Cys Ser Phe Thr Ser Asn Leu Cys Ala Phe
195 200 205

Phe Lys Ile Ala Gln Tyr Lys Val Val Arg Phe Lys Gly Gly Ser Leu
210 215 220

Lys Glu Ser Gln Ala Thr Leu Asn Lys Val Phe Ala Leu Tyr Gln Thr
225 230 235 240

Ser Leu Asp Met Cys Asn Asp Leu Asn Gln Cys Tyr Gln Pro Ile Ile
245 250 255

Cys Ala Gln Phe Phe Ile Ser Ser Leu Gln Leu Cys Met Leu Gly Tyr
260 265 270

Leu Phe Ser Ile Thr Phe Ala Gln Thr Glu Gly Val Tyr Tyr Ala Ser
275 280 285

Phe Ile Ala Thr Ile Ile Ile Gln Ala Tyr Ile Tyr Cys Tyr Cys Gly
290 295 300

Glu Asn Leu Lys Thr Glu Ser Ala Ser Phe Glu Trp Ala Ile Tyr Asp
305 310 315 320

Ser Pro Trp His Glu Ser Leu Gly Ala Gly Gly Ala Ser Thr Ser Ile
325 330 335

Cys Arg Ser Leu Leu Ile Ser Met Met Arg Ala His Arg Gly Phe Arg
340 345 350

Ile Thr Gly Tyr Phe Phe Glu Ala Asn Met Glu Ala Phe Ser Ser Ile
355 360 365

```
Val Arg Thr Ala Met Ser Tyr Ile Thr Met Leu Arg Ser Phe Ser
    370                 375                 380
```