

(12) PATENT
(19) AUSTRALIAN PATENT OFFICE

(11) Application No. AU 200058722 B2
(10) Patent No. 779985

(54) Title
Gustatory receptors in drosophila

(51)⁶ International Patent Classification(s)
C12N 015/12 C07K 016/18
A01K 067/033 C12Q 001/68
C07K 014/435

(21) Application No: **200058722** (22) Application Date: **2000.06.14**

(87) WIPO No: **WO00/77208**

(30) Priority Data

(31) Number	(32) Date	(33) Country
60/138668	1999.06.14	US
60/181704	2000.02.10	US

(43) Publication Date : **2001.01.02**

(43) Publication Journal Date : **2001.03.15**

(44) Accepted Journal Date : **2005.02.24**

(71) Applicant(s)
Yale University

(72) Inventor(s)
John R. Carlson; Peter J. Clyne; Coral G. Warr

(74) Agent/Attorney
Allens Arthur Robinson, Chifley Towers, 2 Chifley Square, Sydney NSW 2000

(56) Related Art
EBI ACCESSION # AC003945, 30.12.97

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
21 December 2000 (21.12.2000)

PCT

(10) International Publication Number
WO 00/77208 A3

(51) International Patent Classification: C12N 15/12,
C07K 14/435, 16/18, C12Q 1/68, A01K 67/033

(74) Agent: VEITENHEIMER, Erich, E.; Morgan, Lewis
& Bockius LLP, 1800 M. Street, N.W., Washington, DC
20036-5869 (US).

(21) International Application Number: PCT/US00/16211

(22) International Filing Date: 14 June 2000 (14.06.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/138,668 14 June 1999 (14.06.1999) US
60/181,704 10 February 2000 (10.02.2000) US

(81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE,
DK, DM, DZ, 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, MZ, 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.

(84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, MZ, 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).

(71) Applicant (for all designated States except US): YALE
UNIVERSITY [US/US]; 155 Whitney Avenue, New
Haven, CT 06520-8336 (US).

Published:

— With international search report.

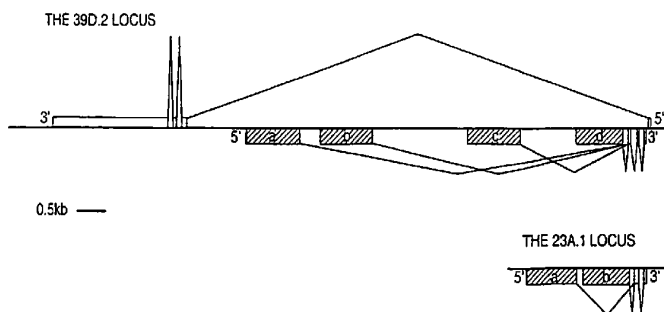
(72) Inventors; and

(75) Inventors/Applicants (for US only): CARLSON, John,
R. [US/US]; 31 Mountain Brook Road, North Haven,
CT 06473 (US). CLYNE, Peter, J. [US/US]; 1939 Turk
Street, San Francisco, CA 94115 (US). WARR, Coral,
G. [US/US]; 24 Bishop Street, 3rd Floor, New Haven, CT
06511 (US).

(88) Date of publication of the international search report:
1 March 2001

For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

(54) Title: GUSTATORY RECEPTORS IN DROSOPHILA



(57) Abstract: The present invention provides nucleic acids and amino acids for novel gustatory receptors as well as methods for identifying gustatory receptors. More specifically, the present invention provides nucleic acids and amino acids for novel gustatory receptors in *Drosophila* as well as methods of using the provided nucleic acids and amino acids. In addition, this invention provides methods of identifying ligands which bind to the novel gustatory receptors as well as a variety of methods for using the receptors and ligands so identified.

WO 00/77208 A3

NOVEL TASTE RECEPTORS IN DROSOPHILA

INVENTORS: John R. Carlson, Peter J. Clyne, Coral G. Warr

5 **RELATED APPLICATIONS**

 This application claims priority to U.S. provisional patent application No. 60/181,704 filed February 10, 2000 and U.S. provisional patent application No. 60/138,668 filed June 14, 1999 both applications herein incorporated by reference in their entirety.

10

U.S. GOVERNMENT SUPPORT

 This work was supported by a grant from the National Institutes of Health (DC-02174).

15 **FIELD OF THE INVENTION**

 This invention pertains to novel taste receptors and to methods of using such receptors. More particularly, this invention pertains to the nucleic acids and amino acids of novel taste receptors in *Drosophila* and to methods of using such nucleic acids and amino acids.

20

BACKGROUND OF THE INVENTION

 Studies in insect gustation have a long history in general physiology. Much more emphasis has been placed on the physiological characteristics of the sensory cells than on the central cellular mechanisms of taste processing (Mitchell *et al.*, (1999) Microsc. Res. Tech. 47,401-415). For a survey review of the organization of the olfactory and gustatory systems in *Drosophila*, see Stocker, (1994) Cell Tissue Res. 275, 3-26.

25

 Both male and female *Drosophila* flies possess gustatory receptors on their legs, but males possess more of these receptors than females (Possidente *et al.*, (1989) Dev.

Biol. 132, 448-457). The labellar hairs of larger flies are not only sensitive to a variety of simple and compound sugars (Dethier, (1955) Q. Rev. Biol. 30, 348), but also to a wide variety of other molecules, such as amino acids (Shiraishi & Kuwabara, (1970) J. Gen. Physiol. 56, 768).

- 5 Behavioral studies have shown that *Drosophila* are sensitive to quinine (Tompkins *et al.*, (1979) Proc. Natl. Acad. Sci. USA 76, 884), which is perceived by humans as bitter, and other insects have been shown to be sensitive to an array of structurally diverse bitter compounds. Moreover, an individual insect taste receptor cell can respond to a broad range of structurally heterogeneous alkaloids and other bitter molecules
- 10 (Glendinning & Hills, (1997) J. Neurophysiol. 78, 734; Chapman *et al.*, (1991) J. Exp. Biol. 158, 241).

- Little is known about the molecular mechanisms of taste perception in animals, particularly the initial events of taste signaling. Several genes known to affect the formation of gustatory sensilla or alter the feeding behavior of *Drosophila* have been
- 15 identified (Singh, (1997) Microsc. Res. Tech. 3, 547-563). For example, a mutation in the *malvolio* (*mvl*) gene affects taste behavior in *Drosophila melanogaster* (Orgad *et al.*, (1998) J. Exp. Biol. 201, 115-120). Also, scalloped (*sd*) mutants show defects in response to a number of taste stimuli (Inamdar *et al.*, (1993) J. Neurogenet. 9, 123-139). Flies containing different Shaker alleles exhibit a variety of defects in their gustatory
- 20 response to sucrose, NaCl and KCl (Balakrishnan *et al.*, (1991) J. Exp. Biol. 157, 161-181).

Although two putative mammalian taste receptors have recently been described (Hoon *et al.*, (1999) Cell 96, 541), remarkably little is understood in general about taste receptors across species.

- 25 In the present invention, a large and diverse family of seven transmembrane domain proteins was identified from the *Drosophila* genome database with a computer algorithm that identifies proteins on the basis of structure. Eighteen of nineteen genes examined were expressed in the *Drosophila* labellum, a gustatory organ of the proboscis.

Expression was not detected in a variety of other tissues. The genes were not expressed in the labellum of a *Drosophila* mutant, *pox-neuro 70*, in which taste neurons are eliminated. Tissue specificity of expression of these genes, along with their structural similarity, supports the possibility that the family encodes a large and divergent family of taste receptors.

SUMMARY OF THE INVENTION

This invention provides isolated nucleic acid molecules including the following:

- (a) isolated nucleic acid molecules that encode the amino acid sequences of *Drosophila* Gustatory Receptor proteins; (b) isolated nucleic acid molecules that encode protein fragments of at least six amino acids of a *Drosophila* Gustatory Receptor proteins; and (c) isolated nucleic acid molecules which hybridize to nucleic acid molecules which include nucleotide sequences encoding *Drosophila* Gustatory Receptor proteins under conditions of sufficient stringency to produce a clear signal.

- This invention also provides such isolated nucleic acid molecules wherein the nucleic acids include at least one exon-intron boundary located in one of the following positions: (a) the nucleotides encoding the amino acids which include the third extracellular loop of a *Drosophila* Gustatory Receptor protein; and (b) the nucleotides encoding the amino acids which include the seventh transmembrane domain of a *Drosophila* Gustatory Receptor protein.

This invention further provides such isolated nucleic acid molecules which have the nucleic acid sequence of one of the following sequences: SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 90 and 91.

- This invention also provides such isolated nucleic acid molecules operably linked to one or more expression control elements.

This invention further provides vectors which include any of the aforementioned nucleic acid molecules and host cells which include such vectors.

This invention also provides host cells transformed so as to contain any of the aforementioned nucleic acid molecules, wherein such host cells can be either prokaryotic host cells or eukaryotic host cells.

5 This invention also provides methods for producing proteins or protein fragments wherein the methods include transforming host cells with any of the aforementioned nucleic acids under conditions in which the protein or protein fragment encoded by said nucleic acid molecule is expressed. This invention also provides such methods wherein the host cells are either prokaryotic host cells or eukaryotic host cells. This invention further provides isolated proteins or protein fragments produced by such methods.

10 This invention provides isolated proteins or protein fragments which include: (a) isolated proteins encoded by one of the following amino acid sequences: SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90 and 92; (b) isolated protein fragments which include at least six amino acids of any of the following
15 sequences: SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90 and 92; (c) isolated proteins which include conservative amino acid substitutions of any of the following sequences: SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90 and 92; and (d) naturally occurring amino
20 acid sequence variants of any of the following sequences: SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90 and 92.

The present invention further provides such isolated proteins or protein fragments
25 which include at least one of the following conserved amino acids: Serine in the amino terminal domain; Phenylalanine in the first transmembrane domain; Arginine in the first extracellular loop; Leucine in the fourth transmembrane domain; Leucine in the third transmembrane domain; Glycine in the fifth transmembrane domain; Tyrosine in the fifth

transmembrane domain; Leucine in the third extracellular loop; Phenylalanine in the third extracellular loop; Alanine in the seventh transmembrane domain; Glycine in the seventh transmembrane domain; Leucine in the seventh transmembrane domain; Aspartate in the seventh transmembrane domain; Alanine in the seventh transmembrane domain;

- 5 Threonine in the seventh transmembrane domain; Tyrosine in the seventh transmembrane domain; Valine in the seventh transmembrane domain; Glutamine in the carboxy terminal domain; and Phenylalanine in the carboxy terminal domain.

- 10 The present invention also provides isolated antibodies that bind to any of the aforementioned polypeptides. The present invention also provides such antibodies which are either monoclonal antibodies or polyclonal antibodies.

- This invention also provides methods of identifying agents which modulate the expression of any of the aforementioned proteins or protein fragments by: (a) exposing cells which express the proteins or protein fragments to the agents; and (b) determining whether the agent modulates expression of said proteins or protein fragments, thereby
15 identifying agents which modulate the expression of the proteins or protein fragments.

- The present invention also provides methods of identifying agents which modulate the activity of any of the aforementioned proteins or protein fragments by: (a) exposing cells which express the proteins or protein fragments to the agents; and (b) determining whether the agents modulate the activity of said proteins or protein fragments, thereby
20 identifying agents which modulate the activity of the proteins or protein fragments.

The present invention also provides such methods where the agent modulates at least one activity of the proteins or protein fragments.

- This invention provides methods of identifying agents which modulate the transcription of any of the aforementioned nucleic acid molecules by: (a) exposing cells
25 which transcribe the nucleic acids to the agents; and (b) determining whether the agents modulate transcription of said nucleic acids, thereby identifying agents which modulate the transcription of the nucleic acid.

This invention further provides methods of identifying binding partners for the

aforementioned proteins or protein fragments by: (a) exposing said proteins or protein fragments to potential binding partners; and (b) determining if the potential binding partners bind to said proteins or protein fragments, thereby identifying binding partners for the proteins or protein fragments.

5 The present invention also provides methods of modulating the expression of nucleic acids encoding the aforementioned proteins or protein fragments by administering an effective amount of agents which modulate the expression of the nucleic acids encoding the proteins or protein fragments.

10 This invention also provides methods of modulating at least one activity of the aforementioned proteins or protein fragments by administering an effective amount of the agents which modulate at least one activity of the proteins or protein fragments.

 This invention provides methods of identifying novel gustatory receptor genes by: (a) selecting candidate gustatory receptor genes by screening nucleic acid databases using an algorithm trained to identify seven transmembrane receptors genes; (b) screening
15 said selected candidate gustatory receptor genes by identifying nucleic acid sequences with conserved amino acid residues and intron-exon boundaries common to gustatory receptors, and having open reading frames of sufficient size so as to encode a seven transmembrane receptor; and (c) identifying the novel gustatory receptor genes and measuring the expression of gustatory receptor genes wherein the detection of expression
20 confirms said candidate gustatory genes as gustatory genes.

 This invention also provides methods of identifying novel gustatory receptor genes by: (a) selecting candidate gustatory receptor genes by screening nucleic acid databases for nucleic acid sequences with sufficient homology to at least one known gustatory receptor gene; (b) screening said selected candidate gustatory receptor genes by
25 identifying nucleic acids with conserved amino acid residues and intron-exon boundaries common to gustatory receptors, and having open reading frames of sufficient size so as to encode a seven transmembrane receptor; and (c) identifying the novel gustatory receptor genes and measuring the expression of gustatory receptor genes wherein the detection of

expression confirms said candidate gustatory genes as gustatory genes.

The present invention also provides transgenic insects modified to contain any of the aforementioned nucleic acid molecules.

This invention also provides such transgenic insects, wherein the nucleic acid
 5 molecules contain mutations that alter expression of the encoded proteins.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 Amino acid sequence alignment of nineteen GR proteins. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu;
 10 F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr. Letters following protein designations identify alternative splicing products of individual genes. Residues conserved in >50% of the predicted proteins are shaded. The approximate locations of the seven predicted transmembrane domains are indicated. Intron-exon boundaries are shown with vertical
 15 lines. The sequences shown are the first 19 full-length proteins identified. All DNA sequences are from the Berkeley *Drosophila* Genome Project (BDGP) database. See the Examples for a complete description.

Figure 2 Representative hydropathy plots of GR proteins. Hydrophobic peaks
 20 predicted by Kyte-Doolittle analysis appear above the center lines. The approximate positions of the seven putative transmembrane domains are indicated above the first hydropathy plot. Similar plots were obtained for all of the GR proteins.

Figure 3 Genomic organization of the 39D.2 and 23A.1 loci. In the 39D.2 locus,
 25 the gray boxes labeled a through d represent four large 5' exons, each of which can be spliced individually to the three 3' exons (indicated in black) to produce alternative transcripts encoding four different proteins. All the exons of the 39D.2 locus are located in an intron of another gene, which is in the opposite orientation and whose exons are

represented by white boxes. This other gene appears to encode a basic helix-loop-helix transcription factor expressed during embryogenesis. In the 23A.1 locus, the gray boxes labeled a and b represent two alternative large 5' exons, either of which can be spliced to the two small 3' exons (indicated in black) to produce transcripts encoding two different proteins.

Figure 4 Tissue specificity of expression of 32D.1 in the labellum. Shown is a gel photograph of an RT-PCR experiment with primers spanning an intron in 32D.1. The size of the predicted PCR product from cDNA is 372 base pairs; any remaining genomic DNA would generate a product of 559 base pairs. A cDNA band is observed in the labellum lane only. In addition, 32D.1 is not expressed in the labellum of the *poxn* 70 mutant. Positive controls are described in the Examples.

The amount of each tissue used to prepare cDNA was that determined to give approximately the same signal with a pair of positive control primers, CG-
 GATCCCTATGTCAAGGTG (SEQ ID NO: 93) and GAAGAGCTTCGTGCTGGTCT
 (SEQ ID NO: 94), representing the *Drosophila* synaptotagmin gene (Perin *et al.*, (1991) J. Biol. Chem. 266, 615). Specifically, the amount of tissue used in each cDNA preparation was as follows: fifty labella, five heads from which taste organs (the labellum, the LSO, the dorsal cibarial sense organ, and the ventral cibarial sense organ) had been surgically removed, twenty thoraces, twenty abdomens, two-hundred legs, and twenty anterior wing margins (the portion of the wing containing chemosensory sensilla).

Complementary DNA preparation and PCR were performed as in Clyne *et al.*, (1999) Neuron 22, 327-338. For all genes, primer pairs (Available as supplementary material at www.sciencemag.org/feature/data/1046815.shl) that span introns were used to distinguish bands amplified from cDNA from those amplified from any remaining genomic DNA. All negative results were confirmed by testing at least one additional primer pair.

Figure 5 GR gene expression in microdissected labral sense organs (LSOs). The shaded areas show the four major taste organs of the *Drosophila* head: the LSO, the dorsal cibarial sense organ (DCSO), the ventral cibarial sense organ (VCSO), and the labellum. The gel track shows an amplification product from RNA extracted from fifty LSOs, amplified with primers N23A.3J and N23A.2D from two exons of gene 23A.1. Specifically, one primer is from the large exon 23A.1a (Figure 3), and the other is from the First common exon at the 3' end. The amplification product is 430 base pairs, which is the expected length for a cDNA product; any remaining genomic DNA would generate a product of 1598 base pairs. The primer pair did not amplify a product from non-gustatory tissue (see Examples, Table 1). The following transcripts were detected in the LSO: 22B.1, 23A.1a, 23A.1b, 32D.1, 39D.2c, 43C.1, and 58A.2.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

I. Specific Embodiments

15 A. *Drosophila* Gustatory Receptor Proteins

The present invention provides a family of isolated proteins, allelic variants of the proteins, and conservative amino acid substitutions of the proteins. As used herein, protein or polypeptide refers to any one of the proteins that has the amino acid sequence depicted in SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90 and 92. The invention also includes naturally occurring allelic variants and proteins that have a slightly different amino acid sequence than that specifically recited above. Allelic variants, though possessing a slightly different amino acid sequence than those recited above, will still have the same or similar biological functions associated with any of the amino acid proteins.

As used herein, the family of proteins related to any one of the amino acid sequences depicted in SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80,

82, 84, 86, 88, 90 and 92 refers to proteins that have been isolated from organisms in addition to *Drosophila*. The methods used to identify and isolate other members of the family of proteins related to these amino acid proteins are described below.

5 The proteins of the present invention are preferably in isolated form. As used herein, a protein is said to be isolated when physical, mechanical or chemical methods are employed to remove the protein from cellular constituents that are normally associated with the protein. A skilled artisan can readily employ standard purification methods to obtain an isolated protein.

10 The proteins of the present invention further include conservative amino acid substitution variants (*i.e.*, conservative) of the proteins herein described. As used herein, a conservative variant refers to at least one alteration in the amino acid sequence that does not adversely affect the biological functions of the protein. A substitution, insertion or deletion is said to adversely affect the protein when the altered sequence prevents or disrupts a biological function associated with the protein. For example, the overall
15 charge, structure or hydrophobic-hydrophilic properties of the protein can be altered without adversely affecting a biological activity. Accordingly, the amino acid sequence can often be altered, for example to render the peptide more hydrophobic or hydrophilic, without adversely affecting the biological activities of the protein.

Ordinarily, the allelic variants, the conservative substitution variants, and the
20 members of the protein family, will have an amino acid sequence having at least 10% amino acid sequence identity with the sequences set forth in SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84 or 86 more preferably at least 35%, even more preferably at least 40% and most preferably at least 45%. Identity or homology
25 with respect to such sequences is defined herein as the percentage of amino acid residues in the candidate sequence that are identical with the known peptides, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent homology, and not considering any conservative substitutions as part of the sequence identity.

N-terminal, C-terminal or internal extensions, deletions or insertions into the peptide sequence shall not be construed as affecting homology.

The proteins of the present invention have seven transmembrane domains as defined by hydropathy analysis (Kyte & Doolittle, (1982) J. Mol. Biol. 157, 105-132).

- 5 Furthermore, the proteins of the present invention have conserved amino acid residues in defined domains of the protein. For example, the proteins of the present invention have at least one of the following conserved amino acids as depicted in Figure 1, including but not limited to, Serine in the amino terminal domain; Phenylalanine in the first transmembrane domain; Arginine in the first extracellular loop; Leucine in the fourth transmembrane domain; Leucine in the third transmembrane domain; Glycine in the fifth transmembrane domain; Tyrosine in the fifth transmembrane domain; Leucine in the third extracellular loop; Phenylalanine in the third extracellular loop; Alanine in the seventh transmembrane domain; Glycine in the seventh transmembrane domain; Leucine in the seventh transmembrane domain; Aspartate in the seventh transmembrane domain; Alanine in the seventh transmembrane domain; Threonine in the seventh transmembrane domain; Tyrosine in the seventh transmembrane domain; Valine in the seventh transmembrane domain; Glutamine in the carboxy terminal domain; and Phenylalanine in the carboxy terminal domain. In addition, the conserved amino acids may be selected from any of the amino acid residues indicated as being conserved among *Drosophila* Gustatory Receptor (DGR) proteins as depicted in Figure 1 (shaded).

- 20 Thus, the proteins of the present invention include molecules having the amino acid sequence disclosed in SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90 and 92; fragments thereof having a consecutive sequence of at least about 3, 4, 5, 6, 10, 15, 20, 25, 30, 35 or more amino acid residues of the proteins, for instance, antigenic fragments such as those found in the extracellular loops of the protein (see Figure 1); amino acid sequence variants wherein an amino acid residue has been inserted N- or C-terminal to, or within, the disclosed sequence; and amino acid

sequence variants of the disclosed sequences, or their fragments as defined above, that have been substituted by another residue. Contemplated variants further include those containing predetermined mutations by, *e.g.*, homologous recombination, site-directed or PCR mutagenesis, and the corresponding proteins of other insect species, including but
 5 not limited to the order *Diptera*, *Lepidoptera*, *Homoptera* and *Coleoptera*, within these orders, preferably the genus *Drosophila*, *Anopheles*, *Aedes*, *Ceratitis*, *Muscidae*, *Culicidae*, *Anagasta* and *Popilla* and the alleles or other naturally occurring variants of the family of proteins; and derivatives wherein the protein has been covalently modified by substitution, chemical, enzymatic, or other appropriate means with a moiety other than
 10 a naturally occurring amino acid (for example a detectable moiety such as an enzyme or radioisotope).

As described below, members of the family of proteins can be used: 1) to identify agents which modulate at least one activity of the protein; 2) to identify binding partners for the protein, 3) as an antigen to raise polyclonal or monoclonal antibodies, and 4) in
 15 methods to modify insect behavior.

B. Nucleic Acid Molecules

The present invention further provides nucleic acid molecules which encode any of the proteins having SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30,
 20 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90 and 92 and the related proteins herein described, preferably in isolated form. As used herein, "nucleic acid" is defined as RNA or DNA that encodes a protein or peptide as defined above, is complementary to a nucleic acid sequence encoding such peptides, hybridizes to such a nucleic acid and remains stably bound to it
 25 under appropriate stringency conditions, or encodes a polypeptide sharing at least 75% sequence identity, preferably at least 80%, and more preferably at least 85%, with the peptide sequences in conserved domains. Specifically contemplated are genomic DNA, cDNA, mRNA and antisense molecules, as well as nucleic acids based on alternative

backbones or including alternative bases whether derived from natural sources or synthesized. Such hybridizing or complementary nucleic acids, however, are defined further as being novel and non-obvious over any prior art nucleic acid including that which encodes, hybridizes under appropriate stringency conditions, or is complementary to nucleic acid encoding a protein according to the present invention.

Homology or identity at the amino acid or nucleotide level is determined by **BLAST** (Basic Local Alignment Search Tool) analysis using the algorithm employed by the programs **blastp**, **blastn**, **blastx**, **tblastn** and **tblastx** (Karlin *et al.*, (1990) Proc. Natl. Acad. Sci. USA 87, 2264-2268 and Altschul, (1993) J. Mol. Evol. 36, 290-300, fully incorporated by reference) which are tailored for sequence similarity searching. The approach used by the **BLAST** program is to first consider similar segments between a query sequence and a database sequence, then to evaluate the statistical significance of all matches that are identified and finally to summarize only those matches which satisfy a preselected threshold of significance. For a discussion of basic issues in similarity searching of sequence databases (see Altschul *et al.*, (1994) Nature Genetics 6, 119-129 which is fully incorporated by reference). The search parameters for **histogram**, **descriptions**, **alignments**, **expect** (*i.e.*, the statistical significance threshold for reporting matches against database sequences), **cutoff**, **matrix** and **filter** are at the default settings. The default scoring matrix used by **blastp**, **blastx**, **tblastn**, and **tblastx** is the **BLOSUM62** matrix (Henikoff *et al.*, (1992) Proc. Natl. Acad. Sci. USA 89, 10915-10919, fully incorporated by reference). For **blastn**, the scoring matrix is set by the ratios of **M** (*i.e.*, the reward score for a pair of matching residues) to **N** (*i.e.*, the penalty score for mismatching residues), wherein the default values for **M** and **N** are 5 and -4, respectively.

"Stringent conditions" are those that (1) employ low ionic strength and high temperature for washing, for example, 0.5 M sodium phosphate buffer at pH 7.2, 1 mM EDTA at pH 8.0 in 7% SDS at either 65°C or 55°C, or (2) employ during hybridization a denaturing agent such as formamide, for example, 50% formamide with 0.1% bovine

serum albumin, 0.1% Ficoll, 0.1% polyvinylpyrrolidone, 0.05 M sodium phosphate buffer at pH 6.5 with 0.75 M NaCl, 0.075 M sodium citrate at 42°C. Another example is use of 50% formamide, 5× SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate at pH 6.8, 0.1% sodium pyrophosphate, 5× Denhardt's solution, sonicated salmon sperm DNA (50 µg/ml), 0.1% SDS and 10% dextran sulfate at 55°C, with washes at 55°C in 0.2× SSC and 0.1% SDS. A skilled artisan can readily determine and vary the stringency conditions appropriately to obtain a clear and detectable hybridization signal. Preferred molecules are those that hybridize under the above conditions to the complements of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 90 and 91, and which encode a functional protein.

As used herein, a nucleic acid molecule is said to be "isolated" when the nucleic acid molecule is substantially separated from contaminant nucleic acid encoding other polypeptides from the source of nucleic acid. For information on how to extract and manipulate nucleic acids from *Drosophila*, see, for example, Roberts, (1998) *Drosophila: A Practical Approach*, IRL Press.

The present invention further provides fragments of any one of the encoding nucleic acids molecules. As used herein, a fragment of an encoding nucleic acid molecule refers to a small portion of the entire protein coding sequence. The size of the fragment will be determined by the intended use. For example, if the fragment is chosen so as to encode an active portion of the protein, the fragment will need to be large enough to encode the functional region(s) of the protein. For instance, fragments of the invention encode antigenic fragments such as the extracellular loops or N-terminal domain of the protein depicted in SEQ ID NO: 9 (21D.1) and as set forth in Figure 1. If the fragment is to be used as a nucleic acid probe or PCR primer, then the fragment length is chosen so as to obtain a relatively small number of false positives during probing and priming.

Fragments of the encoding nucleic acid molecules of the present invention (*i.e.*, synthetic oligonucleotides) that are used as probes or specific primers for the polymerase

chain reaction (PCR), or to synthesize gene sequences encoding proteins of the invention can easily be synthesized by chemical techniques, for example, the phosphotriester method of Matteucci *et al.*, (1981) J. Am. Chem. Soc. 103, 3185-3191) or using automated synthesis methods. In addition, larger DNA segments can readily be prepared by well known methods, such as synthesis of a group of oligonucleotides that define various modular segments of the gene, followed by ligation of oligonucleotides to build the complete modified gene.

The encoding nucleic acid molecules of the present invention may further be modified so as to contain a detectable label for diagnostic and probe purposes. A variety of such labels are known in the art and can readily be employed with the encoding molecules herein described. Suitable labels include, but are not limited to, fluorescent-labeled, biotin-labeled, radio-labeled nucleotides and the like. A skilled artisan can employ any of the art known labels to obtain a labeled encoding nucleic acid molecule.

Modifications to the primary structure itself by deletion, addition, or alteration of the amino acids incorporated into the protein sequence during translation can be made without destroying the activity of the protein. Such substitutions or other alterations result in proteins having an amino acid sequence encoded by a nucleic acid falling within the contemplated scope of the present invention.

C. Isolation of Other Related Nucleic Acid Molecules

As described above, the identification and characterization of the nucleic acid molecules having SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 90 and 91 allows a skilled artisan to isolate nucleic acid molecules that encode other members of the protein family in addition to the sequences herein described. Further, the presently disclosed nucleic acid molecules allow a skilled artisan to isolate nucleic acid molecules that encode other members of the family of proteins in addition to the protein having SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90 and

92.

Essentially, a skilled artisan can readily use any one of the amino acid sequences selected from SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90 and 92, to generate antibody probes to screen expression libraries prepared from appropriate cells. Typically, polyclonal antiserum from mammals such as rabbits immunized with the purified protein (as described below) or monoclonal antibodies can be used to probe a cDNA or genomic expression library to obtain the appropriate coding sequence for other members of the protein family. The cloned cDNA sequence can be expressed as a fusion protein, expressed directly using its own control sequences, or expressed by constructions using control sequences appropriate to the particular host used for expression of the enzyme.

Alternatively, a portion of the coding sequence herein described can be synthesized and used as a probe to retrieve DNA encoding a member of the protein family from any organism. Oligomers containing approximately 18-20 nucleotides (encoding about a six to seven amino acid stretch) are prepared and used to screen genomic DNA or cDNA libraries to obtain hybridization under stringent conditions or conditions of sufficient stringency to eliminate an undue level of false positives.

Additionally, pairs of oligonucleotide primers can be prepared for use in a polymerase chain reaction (PCR) to selectively clone an encoding nucleic acid molecule. A PCR denature/anneal/extend cycle for using such PCR primers is well known in the art and can readily be adapted for use in isolating other encoding nucleic acid molecules. For example, degenerate primers can be used to clone any *Drosophila* Gustatory Receptor (DGR) gene across species. Specifically, based on the sequence information derived from the family of *Drosophila* Gustatory Receptors, degenerate primers can be designed based on conserved sequences among gustatory receptors, which can then be used to clone nucleic acid molecules encoding gustatory receptor proteins from other species of insects.

Applicants have also identified a method for isolating nucleic acid molecules that

encode other members of the protein family in addition to the sequences herein described. Essentially, a two-step strategy is employed to identify gustatory receptor genes from the genomic database. First, a computer algorithm was designed to search genomic sequences for open reading frames (ORFs) from candidate gustatory receptor genes. Second, RT-PCR is
5 used to determine if transcripts from any of these ORFs are expressed in gustatory organs.

The algorithm is used to identify G protein-coupled receptors (GPCR) genes using statistical characterization of amino acid physico-chemical profiles in combination with a non-parametric discriminant function. The algorithm is trained on a set of putative sequences from a database. In the first step, three sets of descriptors are used to summarize the
10 physico-chemical profiles of the sequences. These are GES scale of hydropathy (Engelman *et al.*, (1986) *Annu. Rev. Biophys. Biophys. Chem.* 15, 321-353), polarity (Brown, (1991) *Molecular Biology Labfax*, Academic Press), and amino acid usage frequency. For the first two of these measurements, a computed sliding window profile is employed (White, (1994) *Membrane Protein Structure*, Oxford University Press) using a kernel of a certain number of
15 amino acids as a constant function convoluted with a certain number of amino acids as a Gaussian function. These profiles are then summarized with three statistics; the periodicity, average derivative and the variance of the derivative.

Each sequence is then characterized by multiple variables using a non-parametric linear discriminant function that is optimized to separate the known family proteins from
20 random proteins in the training set. The same linear discriminant function with the scores derived from the training set is used to screen any nucleic acid database for candidate genes. The candidate sequences are given significance values by an odds ratio of the proteins and non-family proteins, computed using the observed empirical distribution of the training set. Those sequences with a sufficiently high odds ratio are considered for further analysis. The
25 algorithm can also be used to identify any protein family by altering the training set of sequences.

The method of identification further includes steps for identifying novel gustatory receptor genes comprising selecting candidate gustatory receptor genes by screening a nucleic

acid database using an algorithm trained to identify seven transmembrane receptors genes; screening said selected candidate gustatory receptor genes by identifying nucleic acid sequences with conserved amino acid residues and intron-exon boundaries common to gustatory receptors, and open reading frames of sufficient size as to encode a seven
5 transmembrane receptor. As an additional step, the expression of gustatory receptor genes is measured to confirm candidate gustatory gene as an gustatory gene. The exon-intron boundaries and conserved amino acid residues may be selected from any of the positions depicted in Figure 3. Alternatively, selecting candidate gustatory receptor genes by screening a nucleic acid database for nucleic acid sequences with sufficient homology to at least one
10 known gustatory receptor gene is also encompassed in the invention. In a preferred embodiment, the nucleic acid database is a genomic database, an EST database or even an gustatory receptor database as previously described (Skoufos *et al.*, (1999) Nucleic Acids Research 27, 343-345).

In one example of the invention, the training set could consist of known gustatory
15 receptors from *Drosophila* and could be used to search genomic sequences for new gustatory receptors in other species. In a similar example, the training set could consist of known sequences coding for receptors from a particular family and could be used to identify homologs across species. Specifically, gustatory receptors of one species could be used as a training set to identify gustatory receptors in another species.

20

D. rDNA molecules containing a DNA molecule

The present invention further provides recombinant DNA molecules (rDNAs) that contain a coding sequence. As used herein, a rDNA molecule is a DNA molecule that has been subjected to molecular manipulation *in vitro*. Methods for generating rDNA molecules
25 are well known in the art, for example, see Sambrook *et al.*, (1989) Molecular Cloning - A Laboratory Manual, Cold Spring Harbor Laboratory Press. In the preferred rDNA molecules, a coding DNA sequence is operably linked to expression control sequences or vector sequences.

The choice of vector and expression control sequences to which one of the protein family encoding sequences of the present invention is operably linked depends directly, as is well known in the art, on the functional properties desired, *e.g.*, protein expression, and the host cell to be transformed. A vector contemplated by the present invention is at least capable of directing the replication or insertion into the host chromosome, and preferably also expression, of the structural gene included in the rDNA molecule.

Expression control elements that are used for regulating the expression of an operably linked protein encoding sequence are known in the art and include, but are not limited to, inducible promoters, constitutive promoters, secretion signals, and other regulatory elements. Preferably, the inducible promoter is readily controlled, such as being responsive to a nutrient in the host cell's medium.

In one embodiment, the vector containing a coding nucleic acid molecule will include a prokaryotic replicon, *i.e.*, a DNA sequence having the ability to direct autonomous replication and maintenance of the recombinant DNA molecule extra-chromosomally in a prokaryotic host cell, such as a bacterial host cell, transformed therewith. Such replicons are well known in the art. In addition, vectors that include a prokaryotic replicon may also include a gene whose expression confers a detectable marker such as a drug resistance. Typical bacterial drug resistance genes are those that confer resistance to ampicillin or tetracycline.

Vectors that include a prokaryotic replicon can further include a prokaryotic or bacteriophage promoter capable of directing the expression (transcription and translation) of the coding gene sequences in a bacterial host cell, such as *E. coli*. A promoter is an expression control element formed by a DNA sequence that permits binding of RNA polymerase and transcription to occur. Promoter sequences compatible with bacterial hosts are typically provided in plasmid vectors containing convenient restriction sites for insertion of a DNA segment of the present invention. Typical of such vector plasmids are pUC8, pUC9, pBR322 and pBR329 available from BioRad Laboratories, pPL and pKK223 available from Pharmacia.

- Expression vectors compatible with eukaryotic cells, preferably those compatible with invertebrate cells such as insect cells, can also be used to form a rDNA molecules that contains a coding sequence. Eukaryotic cell expression vectors are well known in the art and are available from several commercial sources. Typically, such vectors are provided
- 5 containing convenient restriction sites for insertion of the desired DNA segment. Typical of such vectors are pSVL and pKSV-10 (Pharmacia), pBPV-1/pML2d (International Biotechnologies, Inc.), pTDT1 (ATCC, #31255), the vector pCDM8 described herein, and the like eukaryotic expression vectors. Vectors may be modified to include insect cell specific promoters if needed.
- 10 Eukaryotic cell expression vectors used to construct the rDNA molecules of the present invention may further include a selectable marker that is effective in an eukaryotic cell, preferably a drug resistance selection marker. A preferred drug resistance marker is the gene whose expression results in neomycin resistance, *i.e.*, the neomycin phosphotransferase (*neo*) gene (Southern *et al.*, (1982) J. Mol. Appl. Genet. 1, 327-341). Alternatively, the
- 15 selectable marker can be present on a separate plasmid, and the two vectors are introduced by co-transfection of the host cell, and selected by culturing in the appropriate drug for the selectable marker.

E. Host Cells Containing an Exogenously Supplied Coding Nucleic Acid

- 20 The present invention further provides host cells transformed with a nucleic acid molecule that encodes a protein of the present invention. The host cell can be either prokaryotic or eukaryotic. Eukaryotic cells useful for expression of a protein of the invention are not limited, so long as the cell line is compatible with cell culture methods and compatible with the propagation of the expression vector and expression of the gene product. Preferred
- 25 eukaryotic host cells include, but are not limited to, yeast, insect and mammalian cells, preferably insect cells such as those from a *Drosophila* cell line. Preferred *Drosophila* host cells include *Drosophila* Schneider line 2, and the like insect tissue culture cell lines. Any prokaryotic host can be used to express a rDNA molecule encoding a protein of the

invention. The preferred prokaryotic host is *E. coli*.

Transformation of appropriate cell hosts with a rDNA molecule of the present invention is accomplished by well known methods that typically depend on the type of vector used and host system employed. With regard to transformation of prokaryotic host cells, 5 electroporation and salt treatment methods are typically employed, see, for example, Cohen *et al.*, (1972) Proc. Natl. Acad. Sci. USA 69, 2110-2114; and Sambrook *et al.*, (1989) Molecular Cloning - A Laboratory Manual, Cold Spring Harbor Laboratory Press. With regard to transformation of vertebrate cells with vectors containing rDNAs, electroporation, cationic lipid or salt treatment methods are typically employed, see, for example, Graham *et al.*, (1973) 10 Virology 52, 456-467; Wigler *et al.*, (1979) Proc. Natl. Acad. Sci. USA 76, 1373-1376.

Successfully transformed cells, *i.e.*, cells that contain a rDNA molecule of the present invention, can be identified by well known techniques including the selection for a selectable marker. For example, cells resulting from the introduction of an rDNA of the present invention can be cloned to produce single colonies. Cells from those colonies can be 15 harvested, lysed and their DNA content examined for the presence of the rDNA using a method such as that described by Southern, (1975) J. Mol. Biol. 98, 503-517; or Berent *et al.*, (1985) Biotech. Histochem. 3, 208; or the proteins produced from the cell assayed via an immunological method.

20 F. Production of Recombinant Proteins using a rDNA Molecule

The present invention further provides methods for producing a protein of the invention using nucleic acid molecules herein described. In general terms, the production of a recombinant form of a protein typically involves the following steps: First, a nucleic acid molecule is obtained that encodes a protein of the invention, such as any of the nucleic acid 25 molecule depicted in SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 90 and 91. The nucleic acid molecule is then preferably placed in operable linkage with suitable control sequences, as described above, to form an expression unit containing the

protein open reading frame. The expression unit is used to transform a suitable host and the transformed host is cultured under conditions that allow the production of the recombinant protein. Optionally the recombinant protein is isolated from the medium or from the cells; recovery and purification of the protein may not be necessary in some instances where some
5 impurities may be tolerated.

Each of the foregoing steps can be done in a variety of ways. For example, the desired coding sequences may be obtained from genomic fragments and used directly in appropriate hosts. The construction of expression vectors that are operable in a variety of hosts is accomplished using appropriate replicons and control sequences, as set forth above. The
10 control sequences, expression vectors, and transformation methods are dependent on the type of host cell used to express the gene and were discussed in detail earlier. Suitable restriction sites can, if not normally available, be added to the ends of the coding sequence so as to provide an excisable gene to insert into these vectors. A skilled artisan can readily adapt any host-expression system known in the art for use with the nucleic acid molecules of the
15 invention to produce recombinant protein.

G. Methods to Identify Binding Partners

Another embodiment of the present invention provides methods for use in isolating and identifying binding partners of any of the DGR proteins of the invention. In detail, a
20 protein of the invention is mixed with a potential binding partner or an extract or fraction of a cell under conditions that allow the association of potential binding partners with the protein of the invention. After mixing, peptides, polypeptides, proteins or other molecules that have become associated with a protein of the invention are separated from the mixture. The binding partner that bound to the protein of the invention can then be removed and further
25 analyzed. To identify and isolate a binding partner, the entire protein, for instance a protein comprising the entire amino acid sequence of any of the proteins depicted in SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90 and 92 can be used.

Alternatively, a fragment of any of the proteins can be used.

As used herein, a cellular extract refers to a preparation or fraction which is made from a lysed or disrupted cell. The preferred source of cellular extracts will be cells derived from *Drosophila*, for instance, labellar cellular extract.

5 A variety of methods can be used to obtain an extract of a cell. Cells can be disrupted using either physical or chemical disruption methods. Examples of physical disruption methods include, but are not limited to, sonication and mechanical shearing. Examples of chemical lysis methods include, but are not limited to, detergent lysis and enzyme lysis. A skilled artisan can readily adapt methods for preparing cellular extracts in order to obtain
10 extracts for use in the present methods.

Once an extract of a cell is prepared, the extract is mixed with any of the proteins of the invention under conditions in which association of the protein with the binding partner can occur. A variety of conditions can be used, the most preferred being conditions that closely resemble conditions found in the cytoplasm of a *Drosophila* cell. Features such as osmolarity,
15 pH, temperature and the concentration of cellular extract used, can be varied to optimize the association of the protein with the binding partner.

As used herein, the term "binding partner" refers to any molecule that binds to a DGR protein of the invention. Binding partners to any one of the Gustatory receptors of the invention include, but are not limited to, small molecules, peptides, polypeptides and proteins.
20 In one embodiment, the binding partner is a co-receptor that forms a dimer complex with the Gustatory receptor, such complexes being necessary for efficient signal transduction. In another embodiment, the binding partner can be a G protein or a subunit of a G protein as the Gustatory receptors of the invention are assumed to be G protein-linked because of their seven transmembrane domains.

25 After mixing under appropriate conditions, the bound complex is separated from the mixture. A variety of techniques can be utilized to separate the mixture. For example, antibodies specific to a protein of the invention can be used to immunoprecipitate the binding partner complex. Alternatively, standard chemical separation techniques such as

chromatography and density-sediment centrifugation can be used.

After removal of non-associated cellular constituents found in the extract, the binding partner can be dissociated from the complex using conventional methods. For example, dissociation can be accomplished by altering the salt concentration or pH of the mixture.

- 5 To aid in separating associated binding partner pairs from the mixed extract, the protein of the invention can be immobilized on a solid support. For example, the protein can be attached to a nitrocellulose matrix or acrylic beads. Attachment of the protein to a solid support aids in separating peptide-binding partner pairs from other constituents found in the extract. The identified binding partners can be either a single protein or a complex made up of
- 10 two or more proteins. Alternatively, binding partners may be identified using a Far-Western assay according to the procedures of Takayama *et al.*, (1997) *Methods Mol. Biol.* 69, 171-184 or identified through the use of epitope tagged proteins or GST fusion proteins.

- Alternatively, the nucleic acid molecules of the invention can be used in a yeast two-hybrid system. The yeast two-hybrid system has been used to identify other protein partner
- 15 pairs (Alifragis *et al.*, (1997) *Proc. Natl. Acad. Sci. USA* 94, 13099-13104; Dong *et al.*, (1999) *Gene* 237, 421-428) and can readily be adapted to employ the nucleic acid molecules herein described.

- In another embodiment, binding partners may be identified in insects using single unit recordings as previously described (Kaissling, (1995) *Single unit recordings in insect gustatory organs*, in: Spielman & Brand, (1995) *Experimental Cell Biology of Taste and Olfaction*, CRC Press). Using single unit recordings *in vivo*, response profiles are established
- 20 for potential ligands, these profiles are then categorized into distinct functional classes indicative of distinct receptor-ligand interactions (see, *e.g.*, U.S. Patent No. 5,993,778). Single unit recordings in transgenic insects which contain transgenes resulting in over- or under-
- 25 expression of a gene are also useful for identifying and characterizing ligands which bind to multiple gustatory receptors as well as identifying and characterizing new gustatory receptors.

The nucleic acids of the invention and their corresponding proteins can be used on an array or microarray for high-throughput screening for agents which interact with either the

nucleic acids of the invention or their corresponding proteins.

An "array" or "microarray" generally refers to a grid system which has each position or probe cell occupied by a defined nucleic acid fragments also known as oligonucleotides. The arrays themselves are sometimes referred to as "chips" or "biochips". High-density
5 nucleic acid and protein microarrays often have thousands of probe cells in a variety of grid styles. For DNA microarray protocols particularly suited to studying *Drosophila*, see, for example, Sullivan *et al.*, (2000) *Drosophila* Protocols, Cold Spring Harbor Laboratory Press.

A typical molecular detection chip includes a substrate on which an array of recognition sites, binding sites or hybridization sites are arranged. Each site has a respective
10 molecular receptor which binds or hybridizes with a molecule having a predetermined structure. The solid support substrates which can be used to form surface of the array or chip include organic and inorganic substrates, such as glass, polystyrenes, polyimides, silicon dioxide and silicon nitride. For direct attachment of probes to the electrodes, the electrode surface must be fabricated with materials capable of forming conjugates with the probes.

15 Once the array is fabricated, a sample solution is applied to the molecular detection chip and molecules in the sample bind or hybridize at one or more sites. The sites at which binding occurs are detected, and one or more molecular structures within the sample are subsequently deduced. Detection of labeled batches is a traditional detection strategy and includes radioisotope, fluorescent and biotin labels, but other options are available, including
20 electronic signal transduction.

Polymer arrays of nucleic acid probes can be used to extract information from, for example, nucleic acid samples. These samples are exposed to the probes under conditions that permit binding. The arrays are then scanned to determine to which probes the sample molecules have interacted with the nucleic acids of the polymer array. One can obtain
25 information by careful probe selection and using algorithms to compare patterns of interactions. For example, the method is useful in screening for novel gustatory receptors in multiple organisms. For example, *Drosophila* degenerate gustatory receptor oligonucleotide arrays can be used to examine a nucleic acid sample from another insect species in order to

identify novel gustatory receptors in that species.

In typical applications, a complex solution containing one or more substances to be characterized contacts a polymer array comprising nucleic acids. For example, the array is comprised of nucleic acid probes. The probes of the array can be either DNA or RNA, which
5 may be either single-stranded or double-stranded. In a preferred embodiment of the invention, the probes are arranged (either by immobilization, typically by covalent attachment, of a pre-synthesized probe or by synthesis of the probe on the substrate) on the substrate or chips in lanes stretching across the chip and separated, and these lanes are in
10 turned arranged in blocks of preferably five lanes, although blocks of other sizes will have useful application. The present invention provides individual probes, sets of probes, and arrays of probe sets on chips, in specific patterns which are used to characterize the substances in a complex mixture by producing a distinct image which is representative of the binding interactions between the probes on the chip and the substances in the complex
15 mixture. The pattern of hybridization to the chip allows inferences to be drawn about the substances present in the complex mixture.

The substances in the complex solution will bind to the nucleic acids on the array. The substances of the complex mixture which bind to the nucleic acids of the array may include, but are not limited to, complementary nucleic acids, non-complementary nucleic acids, proteins, antibodies, oligosaccharides, etc. The types of binding may include, but are
20 not limited to, specific and non-specific, competitive and non-competitive, allosteric, cooperative, non-cooperative, complementary and non-complementary, etc. For example, the nucleic acids of the array can bind to complementary nucleic acids in the complex mixture but can also bind in a tertiary manner, independent of base pairing, to non-complementary nucleic acids.

25 The nucleic acids of the array or the substances of the complex mixture may be tagged with a detectable label. The detectable label can be, for example, a luminescent label, a light scattering label or a radioactive label. Accordingly, locations at which substances interact can be identified by either determining if the signal of the label has been quenched by binding or

identifying locations where the signal of the label is present in cases where the substances of the complex mixture have been labeled. Based on the locations where binding is detected, information regarding the complex mixture can be obtained.

5 The methods of this invention will find particular use wherever high through-put of samples is required. In particular, this invention is useful in ligand screening settings and for determining the composition of complex mixtures.

10 Polypeptides are an exemplary system for exploring the relationship between structure and function in biology. When the twenty naturally occurring amino acids are condensed into a polymeric molecule they form a wide variety of three-dimensional configurations, each resulting from a particular amino acid sequence and solvent condition. For example, the number of possible polypeptide configurations using the twenty naturally occurring amino acids for a polymer five amino acids long is over three million. Typical proteins are more than one-hundred amino acids in length.

15 In typical applications, a complex solution containing one or more substances to be characterized contacts a polymer array comprising polypeptides. The polypeptides of the invention can be prepared by classical methods known in the art, for example, by using standard solid phase techniques. The standard methods include exclusive solid phase synthesis, partial solid phase synthesis methods, fragment condensation, classical solution synthesis and recombinant DNA technology (see Merrifield, (1963) Am. Chem. Soc. 85, 2149-2152). On solid phase, the synthesis is typically commenced from the C-terminal end of 20 the peptide using an alpha-amino protected resin. A suitable starting material can be prepared, for instance, by attaching the required alpha-amino acid to a chloromethylated resin, a hydroxy-methyl resin or a benzhydrylamine resin.

25 The alpha-amino protecting groups are those known to be useful in the art of stepwise synthesis of peptides. Included are acyl type protecting groups, aromatic urethane type protecting groups, aliphatic urethane protecting groups and alkyl type protecting groups. The side chain protecting group remains intact during coupling and is not split off during the deprotection of the amino-terminus protecting group or during coupling. The side chain

protecting group must be removable upon the completion of the synthesis of the final peptide and under reaction conditions that will not alter the target peptide.

After removal of the alpha-amino protecting group, the remaining protected amino acids are coupled stepwise in the desired order. An excess of each protected amino acid is generally used with an appropriate carboxyl group activator such as dicyclohexylcarbodiimide (DCC) in solution, for example, in methylene chloride, dimethyl formamide (DMF) mixtures.

In a preferred embodiment, the polypeptides or proteins of the array can bind to other co-receptors to form a heteroduplex on the array. In yet another embodiment, the polypeptides or proteins of the array can bind to peptides or small molecules.

These procedures can also be used to synthesize peptides in which amino acids other than the twenty naturally occurring, genetically encoded amino acids are substituted at one, two, or more positions of any of the compounds of the invention. For instance, naphthylalanine can be substituted for tryptophan, facilitating synthesis. Other synthetic amino acids that can be substituted into the peptides of the present invention include L-hydroxypropyl, L-3, 4-dihydroxyphenylalanyl, d-amino acids such as L-d-hydroxylysyl and D-d-methylalanyl, L- α -methylalanyl and β -amino acids non-naturally occurring synthetic amino acids can also be incorporated into the peptides of the present invention (see Roberts *et al.*, (1983) Peptide Synthesis 5, 341-449).

One can replace the naturally occurring side chains of the twenty genetically encoded amino acids (or D amino acids) with other side chains, for instance with groups such as alkyl, lower alkyl, cyclic four, five, six, to seven-membered alkyl, amide, amide lower alkyl, amide di(lower alkyl), lower alkoxy, hydroxy, carboxy and the lower ester derivatives thereof, and with four, five, six, to seven-membered heterocyclic. In particular, proline analogs in which the ring size of the proline residue is changed from five members to four, six or seven members can be employed. Cyclic groups can be saturated or unsaturated, and if unsaturated, can be aromatic or non-aromatic. Heterocyclic groups preferably contain one or more nitrogen, oxygen, and/or sulphur heteroatoms. Examples of such groups include the

furazanyl, furyl, imidazolidinyl, imidazolyl, imidazoliny, isothiazolyl, isoxazolyl, morpholinyl, oxazolyl, piperazinyl, piperidyl, pyranyl, pyrazinyl, pyrazolidinyl, pyrazolinyl, pyrazolyl, pyridazinyl, pyridyl, pyrimidinyl, pyrrolidinyl, pyrrolinyl, pyrrolyl, thiadiazolyl, thiazolyl, thienyl, thiomorpholinyl and triazolyl. These heterocyclic groups can be substituted
5 or unsubstituted. Where a group is substituted, the substituent can be alkyl, alkoxy, halogen, oxygen, or substituted or unsubstituted phenyl.

One can also readily modify the peptides of the instant invention by phosphorylation (see Bannwarth *et al.*, (1996) *Biorg. Med. Chem. Let.* 6, 2141-2146) and other methods for making peptide derivatives of the compounds of the present invention are described in Hruby
10 *et al.*, (1990) *Biochem. J.* 268, 249-262). Thus, the peptide compounds of the invention also serve as a basis to prepare peptide mimetics with similar biological activity. The array can also comprise peptide mimetics with the same or similar desired biological activity as the corresponding peptide compound but with more favorable activity than the peptide with respect to solubility, stability, and susceptibility to hydrolysis and proteolysis (see Morgan *et*
15 *al.*, (1989) *Ann. Rep. Med. Chem.* 24, 243-252).

Peptides suitable for use in this embodiment generally include those peptides, for example, ligands, that bind to a receptor, such as seven transmembrane proteins. Such peptides typically comprise about 150 amino acid residues or less and, more preferably, about 100 amino acid residues or less. Polypeptides or proteins suitable for use in this embodiment
20 generally include those polypeptides or proteins that interact with a receptor, such as a co-receptor or G protein. Such polypeptides or proteins typically comprise about 150 amino acid residues or more and, more preferably, about 400 amino acids or more.

The peptides of the present invention may exist in a cyclized form with an intramolecular disulfide bond between the thiol groups of the cysteines. Alternatively, an
25 intermolecular disulfide bond between the thiol groups of the cysteines can be produced to yield a dimeric (or higher oligomeric) compound. One or more of the cysteine residues may also be substituted with a homocysteine. Other embodiments of this invention provide for analogs of these disulfide derivatives in which one of the sulfurs has been replaced by a CH₂

group or other isostere for sulfur. These analogs can be made via an intramolecular or intermolecular displacement, using methods known in the art.

H. Methods to Identify Agents that Modulate Expression of DGRs.

5 Another embodiment of the present invention provides methods for identifying agents that modulate the expression of a nucleic acid encoding any one of the DGRproteins of the invention such as any protein having the amino acid sequence depicted in SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90 and 92. Such assays may
10 utilize any available means of monitoring for changes in the expression level of the nucleic acids of the invention. As used herein, an agent is said to modulate the expression of a nucleic acid of the invention, for instance a nucleic acid encoding any one of the proteins having the amino acid sequence depicted in SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78,
15 80, 82, 84, 86, 88, 90 and 92, if it is capable of up- or down-regulating expression of the nucleic acid in a cell.

In one assay format, cell lines that contain reporter gene fusions between the open reading frame of any one of the nucleotides depicted in SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65,
20 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 90 and 91, and any assay fusion partner may be prepared. Numerous assay fusion partners are known and readily available including the firefly luciferase gene and the gene encoding chloramphenicol acetyltransferase (Alam *et al.*, (1990) Anal. Biochem. 188, 245-254). Cell lines containing the reporter gene fusions are then exposed to the agent to be tested under appropriate conditions and time. Differential
25 expression of the reporter gene between samples exposed to the agent and control samples identifies agents which modulate the expression of a nucleic acid encoding at least one of the proteins having the sequence depicted in SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74,

76, 78, 80, 82, 84, 86, 88, 90 and 92.

Additional assay formats may be used to monitor the ability of the agent to modulate the expression of a nucleic acid encoding at least one protein of the invention selected from the group of proteins having SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90 and 92. For instance, mRNA expression may be monitored directly by hybridization to the nucleic acids of the invention. Cell lines are exposed to the agent to be tested under appropriate conditions and time and total RNA or mRNA is isolated by standard procedures such those disclosed in Sambrook *et al.*, (1989) Molecular Cloning - A Laboratory Manual, Cold Spring Harbor Laboratory Press.

Probes to detect differences in RNA expression levels between cells exposed to the agent and control cells may be prepared from the nucleic acids of the invention. It is preferable, but not necessary, to design probes which hybridize only with target nucleic acids under conditions of high stringency. Only highly complementary nucleic acid hybrids form under conditions of high stringency. Accordingly, the stringency of the assay conditions determines the amount of complementary nucleotides which should exist between two nucleic acid strands in order to form a hybrid. Stringency should be chosen to maximize the difference in stability between the probe:target hybrid and potential probe:non-target hybrids.

Probes may be designed from the nucleic acids of the invention through methods known in the art. For instance, the G+C content of the probe and the probe length can affect probe binding to its target sequence. Methods to optimize probe specificity are commonly available in Sambrook *et al.*, (1989) Molecular Cloning - A Laboratory Manual, Cold Spring Harbor Laboratory Press; or Ausubel *et al.*, (1995) Current Protocols in Molecular Biology, Greene Publishing Company.

Hybridization conditions are modified using known methods, such as those described by Sambrook *et al.*, (1989) and Ausubel *et al.*, (1995) as required for each probe. Hybridization of total cellular RNA or RNA enriched for polyA+ RNA can be accomplished in any available format. For instance, total cellular RNA or RNA enriched for polyA RNA

can be affixed to a solid support and the solid support exposed to at least one probe comprising at least one, or part of one of the sequences of the invention under conditions in which the probe will specifically hybridize. Alternatively, nucleic acid fragments comprising at least one, or part of one of the sequences of the invention can be affixed to a solid support, such as a porous glass wafer. The glass wafer can then be exposed to total cellular RNA or polyA RNA from a sample under conditions in which the affixed sequences will specifically hybridize. Such glass wafers and hybridization methods are widely available, for example, those disclosed by Beattie, 1995 (WO 9511755). By examining for the ability of a given probe to specifically hybridize to an RNA sample from an untreated cell population and from a cell population exposed to the agent, agents which up- or down-regulate the expression of a nucleic acid encoding at least one protein having the amino acid sequence depicted in SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90 and 92 are identified.

Hybridization for qualitative and quantitative analysis of mRNA may also be carried out by using a RNase Protection Assay (*i.e.*, RPA, see Ma *et al.*, (1996) Methods 10, 273-238). Briefly, an expression vehicle comprising cDNA encoding the gene product and a phage specific DNA dependent RNA polymerase promoter (*e.g.*, T7, T3 or SP6 RNA polymerase) is linearized at the 3' end of the cDNA molecule, downstream from the phage promoter, wherein such a linearized molecule is subsequently used as a template for synthesis of a labeled antisense transcript of the cDNA by *in vitro* transcription. The labeled transcript is then hybridized to a mixture of isolated RNA (*i.e.*, total or fractionated mRNA) by incubation at 45°C overnight in a buffer comprising 80% formamide, 40 mM Pipes (pH 6.4), 0.4 M NaCl and 1 mM EDTA. The resulting hybrids are then digested in a buffer comprising 40 µg/ml ribonuclease A and 2 µg/ml ribonuclease. After deactivation and extraction of extraneous proteins, the samples are loaded onto urea-polyacrylamide gels for analysis.

In another assay format, agents which effect the expression of the instant gene products, cells or cell lines would first be identified which express said gene products

physiologically. Cells and cell lines so identified would be expected to comprise the necessary cellular machinery such that the fidelity of modulation of the transcriptional apparatus is maintained with regard to exogenous contact of agent with appropriate surface transduction mechanisms and the cytosolic cascades. Further, such cells or cell lines would be

5 transduced or transfected with an expression vehicle (e.g., a plasmid or viral vector) construct comprising an operable non-translated 5'-promoter containing end of the structural gene encoding the instant gene products fused to one or more antigenic fragments, which are peculiar to the instant gene products, wherein said fragments are under the transcriptional control of said promoter and are expressed as polypeptides whose molecular weight can be

10 distinguished from the naturally occurring polypeptides or may further comprise an immunologically distinct tag. Such a process is well known in the art (see Sambrook *et al.*, (1989) Molecular Cloning - A Laboratory Manual, Cold Spring Harbor Laboratory Press).

Cells or cell lines transduced or transfected as outlined above would then be contacted with agents under appropriate conditions; for example, the agent comprises an acceptable

15 excipient and is contacted with cells comprised in an aqueous physiological buffer such as phosphate buffered saline (PBS) at physiological pH, Eagles balanced salt solution (BSS) at physiological pH, PBS or BSS comprising serum or conditioned media comprising PBS or BSS and/or serum incubated at 37°C. Said conditions may be modulated as deemed necessary by one of skill in the art. Subsequent to contacting the cells with the agent, said cells will be

20 disrupted and the polypeptides from disrupted cells are fractionated such that a polypeptide fraction is pooled and contacted with an antibody to be further processed by immunological assay (e.g., ELISA, immunoprecipitation or Western blot). The pool of proteins isolated from the "agent contacted" sample will be compared with a control sample where only the excipient is contacted with the cells and an increase or decrease in the immunologically generated signal

25 from the "agent contacted" sample compared to the control will be used to distinguish the effectiveness of the agent.

I. Methods to Identify Agents that Modulate Activity of DGRs

Another embodiment of the present invention provides methods for identifying agents that modulate at least one activity of a protein of the invention such as any one of the proteins having the amino acid sequence of SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26,
5 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90 and 92. Such methods or assays may utilize any means of monitoring or detecting the desired activity including, but not limited to, behavioral and electrophysiological studies.

In one format, the relative amounts of a protein of the invention are expressed in a cell
10 population that has been exposed to the agent to be tested and is compared to an un-exposed control cell population. In this format, probes such as specific antibodies are used to monitor the differential expression of the protein in the different cell populations. Cell lines or populations are exposed to the agent to be tested under appropriate conditions and time. Cellular lysates may be prepared from the exposed cell line or population and a control,
15 unexposed cell line or population. The cellular lysates are then analyzed with the probe.

Antibody probes are prepared by immunizing suitable mammalian hosts in appropriate immunization protocols using the peptides, polypeptides or proteins of the invention if they are of sufficient length, or if desired, required to enhance immunogenicity, conjugated to suitable carriers. Methods for preparing immunogenic conjugates with carriers
20 such as BSA, KLH, or other carrier proteins are well known in the art. In some circumstances, direct conjugation using, for example, carbodiimide reagents may be effective; in other instances linking reagents such as those supplied by Pierce Chemical Co., may be desirable to provide accessibility to the hapten. The hapten peptides can be extended at either the amino or carboxy terminus with a cysteine residue or interspersed with cysteine residues,
25 for example, to facilitate linking to a carrier. Administration of the immunogens is conducted generally by injection over a suitable time period and with use of suitable adjuvants, as is generally understood in the art. During the immunization schedule, titers of antibodies are taken to determine adequacy of antibody formation.

While the polyclonal antisera produced in this way may be satisfactory for some applications, for some applications, use of monoclonal preparations is preferred.

Immortalized cell lines which secrete the desired monoclonal antibodies may be prepared using the standard method of Kohler & Milstein, (1975) Nature 256, 495-497 or modifications

5 which effect immortalization of lymphocytes or spleen cells, as is generally known. The immortalized cell lines secreting the desired antibodies are screened by immunoassay in which the antigen is the peptide hapten, polypeptide or protein. When the appropriate immortalized cell culture secreting the desired antibody is identified, the cells can be cultured either *in vitro* or by production in ascites fluid.

10 The desired monoclonal antibodies are then recovered from the culture supernatant or from the ascites supernatant. Fragments of the monoclonal or polyclonal antisera which contain the immunologically significant portion can be used as antagonists, as well as the intact antibodies. Use of immunologically reactive fragments, such as the Fab, Fab' of F(ab')₂ fragments is often preferable, as these fragments are generally less immunogenic than the
15 whole immunoglobulin.

The antibodies or fragments may also be produced, using current technology, by recombinant means. Antibody regions that bind specifically to the desired regions of the protein can also be produced in the context of chimeras with multiple species origin, particularly humanized antibodies.

20 Agents that are assayed in the above method can be randomly selected or rationally selected or designed. As used herein, an agent is said to be randomly selected when the agent is chosen randomly without considering the specific sequences involved in the association of the a protein of the invention alone or with its associated substrates, binding partners, etc. An example of randomly selected agents is the use a chemical library or a peptide combinatorial
25 library, or a growth broth of an organism.

As used herein, an agent is said to be rationally selected or designed when the agent is chosen on a non-random basis which takes into account the sequence of the target site and its conformation in connection with the agent's action. Agents can be rationally selected or

rationally designed by utilizing the peptide sequences to identify proposed binding motifs, glycosylation and phosphorylation sites on the protein.

5 The agents of the present invention can be, as examples, peptides, small molecules, vitamin derivatives, as well as carbohydrates. A skilled artisan can readily recognize that there is no limit as to the structural nature of the agents of the present invention. Dominant-negative proteins, DNA encoding these proteins, antibodies to these proteins, peptide fragments of these proteins or mimics of these proteins may be contacted with cells to affect function. "Mimic" as used herein refers to the modification of a region or several regions of a peptide molecule to provide a structure chemically different from the parent peptide but topographically and functionally similar to the parent peptide (see Meyers, (1995) Molecular Biology & Biotechnology, VCH Publishers).

10 The peptide agents of the invention can be prepared using standard solid phase (or solution phase) peptide synthesis methods, as is known in the art. In addition, the DNA encoding these peptides may be synthesized using commercially available oligonucleotide synthesis instrumentation and produced recombinantly using standard recombinant production systems. The production using solid phase peptide synthesis is necessitated if non-gene-encoded amino acids are to be included.

20 Another class of agents of the present invention are antibodies immunoreactive with critical positions of proteins of the invention. Antibody agents are obtained by immunization of suitable mammalian subjects with peptides, containing as antigenic regions, those portions of the protein intended to be targeted by the antibodies.

J. Transgenic Organisms

25 Transgenic insects containing mutant, knock-out or modified genes corresponding to any one of the cDNA sequences depicted in SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 90 and 91 are also included in the invention. Transgenic insects are genetically modified insects into which recombinant, exogenous or cloned genetic

material has been experimentally transferred. Such genetic material is often referred to as a "transgene". The nucleic acid sequence of the transgene, in this case a form of any one of the sequences depicted in SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 90 and 91, may be integrated either at a locus of a genome where that particular nucleic acid sequence is not otherwise normally found or at the normal locus for the transgene. The transgene may consist of nucleic acid sequences derived from the genome of the same species or of a different species than the species of the target insect.

The term "germ cell line transgenic insect" refers to a transgenic insect in which the genetic alteration or genetic information was introduced into a germ line cell, thereby conferring the ability of the transgenic insect to transfer the genetic information to offspring. If such offspring in fact possess some or all of that alteration or genetic information, then they too are transgenic insects.

The alteration or genetic information may be foreign to the species of insect to which the recipient belongs, foreign only to the particular individual recipient, or may be genetic information already possessed by the recipient. In the last case, the altered or introduced gene may be expressed (*i.e.*, over-expression and knock-out) differently than the native gene.

Transgenic insects can be produced by a variety of different methods including P element-mediated transformation by microinjection (see, *e.g.*, Rubin & Spradling, (1982) Science 218, 348-353; Orr & Sohal, (1993) Arch. Biochem. Biophys. 301, 34-40), transformation by microinjection followed by transgene mobilization (Mockett *et al.*, (1999) Arch. Biochem. Biophys. 371, 260-269), electroporation (Huynh & Zieler, (1999) J. Mol. Biol. 288, 13-20) and through the use of baculovirus (Yamao *et al.*, (1999) Genes Dev. 13, 511-516. Furthermore, the use of adenoviral vectors to direct expression of a foreign gene to gustatory neuronal cells can also be used to generate transgenic insects (see, *e.g.*, Holtmaat *et al.*, (1996) Brain. Res. Mol. Brain Res. 41, 148-156).

A number of recombinant or transgenic insects have been produced, including those which over-express superoxide dismutase (Mockett *et al.*, (1999) Arch. Biochem. Biophys.

371, 260-269); express Syrian hamster prion protein (Raeber *et al.*, (1995) Mech. Dev. 51, 317-327); express cell-cycle inhibitory peptide aptamers (Kolonin & Finley (1998) Proc. Natl. Acad. Sci. USA 95, 14266-14271); and those which lack expression of the putative ribosomal protein S3A gene (Reynaud *et al.*, (1997) Mol. Gen. Genet. 256, 462-467).

- 5 While insects remain the preferred choice for most transgenic experimentation, in some instances it is preferable or even necessary to use alternative animal species. Transgenic procedures have been successfully utilized in a variety of animals, including mice, rats, sheep, goats, pigs, dogs, cats, monkeys, chimpanzees, hamsters, rabbits, cows and guinea pigs (see, *e.g.*, Kim *et al.*, (1997) Mol. Reprod. Dev. 46, 515-526; Houdebine, (1995) Reprod. Nutr. Dev. 35, 609-617; Petters, (1994) Reprod. Fertil. Dev. 6, 643-645; Schnieke *et al.*, (1997) Science 278, 2130-2133; and Amoah, (1997) J. Anim. Sci. 75, 578-585).
- 10 The method of introduction of nucleic acid fragments into insect cells can be by any method which favors co-transformation of multiple nucleic acid molecules. For instance, *Drosophila* embryonic Schneider line 2 (S2) cells can be stably transfected as previously

- 15 described (Schneider, (1972) J. Embryol. Exp. Morphol. 27, 353-365). Detailed procedures for producing transgenic insects are readily available to one skilled in the art (see Rubin & Spradling, (1982) Science 218, 348-353; Orr & Sohal, (1993) Arch. Biochem. Biophys. 301, 34-40, herein incorporated by reference in their entirety).

20 **K. Uses for Agents that Modulate at Least One Activity of DGRs**

1. Introduction.

- Organisms, including insects, are continually exposed to a great number of gustatory stimuli released by other organisms as well as by other aspects of their environment. The gustatory receptor genes of the present invention play an important role in the detection and
- 25 processing of these chemical stimuli, some of which have been implicated in initiating and modulating host-seeking and other behaviors, such as mating behaviors (see, for example, Roth, (1951) Ann. Entomol. Soc. Am. 44, 59-74; Jones *et al.*, (1976) Ent. Exp. Appn. 19, 19-22; Gillies, (1980) Bull. Ent. Res. 70, 525-532; Kline *et al.*, (1991) J. Med. Entomol. 28, 254-

258).

Most importantly, the DGR genes of the present invention may be used to track down gustatory receptor genes in insects that damage crops or transmit diseases. The present invention provides the tools and methodologies for finding specific compounds that interfere with the insects' ability to detect tastes.

Of course, the present invention has important implications for improved methods of using pheromones and other semiochemicals for pest control. In addition, recent advancements in many other fields have greatly increased the variety of additional technologies for which the present invention also has significant applications. Examples of such advancements include, but are not limited to the following: (a) the development and application of new techniques of chemical identification and synthesis; (b) new chemical release techniques; (c) more sophisticated application technologies; and (d) more detailed information about the behavior of specific organisms.

While not wishing to be bound by the specific embodiments discussed herein, the following sections provide an overview of the wide variety of applications for which the present invention may be employed.

2. Definitions.

As used herein, the term "allomones" refers to any chemical substance produced or acquired by an organism that, when it contacts an individual of another species, evokes in the receiver a behavioral or developmental reaction adaptively favorable to the transmitter.

As used herein, the term "host" refers to any organism on which another organism depends for some life function. Examples of hosts include, but are not limited to, humans which may serve as a host for the feeding of certain species of mosquito and the leaves of soybeans (*Glycine max* (L.)) which may act as hosts for the oviposit of the green cloverworm (*Plathypena scabra* (F.)).

As used herein, the term "kairomones" refers to any of a heterogeneous group of chemical messengers that are emitted by organisms of one species but benefit members of another species. Examples include, but are not limited to, attractants, phagostimulants, and

other substances that mediate the positive responses of, for example, predators to their prey, herbivores to their food plants, and parasites to their hosts. Kairomones suitable for the purposes of the invention and methods of obtaining them are described, for example, Hedin, (1985) *Bioregulators for Pest Control*, American Chemical Society.

5 As used herein, the term "pheromone" refers to a substance, or characteristic mixture of substances, that is secreted and released by an organism and detected by a second organism of the same or a closely related species, in which it causes a specific reaction, such as a definite behavioral reaction or a developmental process. Examples include, but are not limited to, the mating pheromones of fungi and insects. More than a thousand moth sex pheromones
10 (Toth *et al.*, (1992) *J. Chem. Ecol.* 18, 13-25; Am *et al.*, (1998) *Appl. Entomol. Zoo.* 33, 507-511) and hundreds of other pheromones have now been identified, including aggregation pheromones from beetles and other groups of insects. Various compositions, including resins and composite polymer dispensers, have been developed for the controlled release of pheromones have been developed (see, *e.g.*, U.S. Patent No. 5,750,129 & 5,504,142).

15 As used herein, the term "semiochemical" refers to any chemical substance that delivers a message or signal from one organism to another. Examples of such chemicals include, but are not limited to, pheromones, kairomones, oviposition deterrents, or stimulants, and a wide range of other classes of chemicals (see, for example, Nordlund *et al.*, (1981) *Semiochemicals: Their Role in Pest Control*, John Wiley; Howse *et al.*, (1998) *Insect Pheromones and Their Use in Pest Management*, Chapman & Hall).

20 As used herein, the term "synomones" refers to any chemical substance which benefits both the emitter and receiver. Examples include, but are not limited to, compounds involved in floral attraction of pollinators and species-isolating mechanisms, such as sex pheromones of related species, where an inhibitor often functions to prevent mating among sympatric species.

25 As used herein, the term "volatile" refers to a chemical which evaporates readily at those temperatures and pressures which are considered the relevant temperatures and pressures for the reference organism of interest.

3. As Tools for Further Scientific Research.

Identification of Gustatory Receptor Genes in Other Organisms. The algorithms of the present invention may be used directly to search for gustatory receptor genes in other organisms, as explained elsewhere herein.

5 Alternatively, nucleic acid probes or primers may be designed based on the DGR genes of the present invention. Such probes or primers may be used to identify and isolate gustatory receptor genes in other organisms. Methods of creating and using the necessary nucleic acid probes and primers are discussed elsewhere herein.

 The highest probability of success in locating gustatory genes in other organisms
10 using the DGR genes of the present invention will most likely occur by using a boot-strapping or leap-frogging method. Such methods involve first probing organisms most related to fruit flies and successively progressing to more unrelated organisms, using the most newly identified gustatory receptor genes to identify similar genes in the next, more unrelated, insect of interest. Thus, the first organisms to probe with the DGR genes of the present invention
15 most preferably may be other flies from the order *Diptera* (i.e., the two-winged or true flies). Examples of suitable flies include, but are not limited to, the tsetse fly, horse fly, house fly, bluebottle fly, hover fly and mosquito. *Dipterans* which transmit diseases causing serious health problems are of particular interest (e.g., horse fly, tsetse fly, mosquito).

 After the identification of gustatory receptor genes in various *Diptera* insects, the next
20 organisms to probe most preferably may be from orders within the same subclass as *Diptera*. Finally, the next insects to use would be those from orders not within the same subclass as *Diptera*.

 The insects which cause substantial health risks, crop damage, or other significant damage (e.g., to housing structure or cotton clothing) may be the most desirable targets for
25 such studies. Examples of such insects include, but are not limited to, green cloverworm, Mexican bean beetle, potato leafhopper, corn earworm, green stink bug, northern corn rootworm, western corn rootworm, cutworms, wireworms, thrips, fleas, aphids (e.g., pea aphid, spotted alfalfa aphid), European corn borer, fall armyworm, southwestern corn borer,

grasshoppers, Japanese beetle, termites, leafhoppers (e.g., potato leafhopper, three-cornered alfalfa hopper), stink bugs, crickets, Hessian fly, greenbugs and weevils (e.g., alfalfa weevil, bollweevil).

Gustatory receptor genes identified by this process may then be used to screen non-
5 Insecta organisms for gustatory receptor genes. Organisms of interest may include, but are not limited to, mites, ticks, spiders, nematodes, centipedes, mice, rats, salmon, pigeons, dogs, horses and humans. In a preferred embodiment, the gustatory receptor genes identified by this process would be used to identify gustatory receptor genes in humans.

Genetic Manipulations. The tools and methodologies of the present invention may be
10 used by neurobiologists to probe more complex workings of an organism's response system, including those of a mammal's brain.

Knock-outs. By systematically knocking out and analyzing the expression patterns of the gustatory receptor genes of the present invention and observing the effects on taste sensitivity and behavior, researchers will be able to piece together a wiring diagram of the
15 gustatory system of the fruit fly.

The term "knock-out" generally refers to mutant organisms which contain a null allele of a specific gene. Methods of making knock-out or disruption transgenic animals, especially mice, are generally known by those skilled in the art and are discussed herein and elsewhere (see, for example, the section herein entitled Transgenic Organisms and the following:
20 Manipulating the Mouse Embryo, (1986) Cold Spring Harbor Laboratory Press; Capecchi, (1989) Science 244, 1288-1292; Li *et al.*, (1995) Cell 80, 401-411; U.S. Patent No. 5,981,830 & 5,789,654, each of which is incorporated herein by reference.

Parallel studies may be conducted in other organisms by using the gustatory receptor genes and the methods of the present invention to identify the gustatory receptor genes of
25 other organisms and then creating knock-outs for the gustatory receptor genes of those organisms.

Disabling Genes. Using the gustatory receptor genes of the present invention, it is now possible to selectively disable specific DGR genes and look for changes in taste response

and behavior. Parallel studies may be conducted in other organisms by using the gustatory receptor genes and the methods of the present invention to identify the gustatory receptor genes of other organisms and then disabling gustatory receptor genes of those organisms.

5 Methods of disabling genes are generally known by those skilled in the art. An example of an effective disabling modification would be a single nucleotide deletion occurring at the beginning of a gustatory receptor gene that would produce a translational reading frameshift. Such a frameshift would disable the gene, resulting in non-expressible gene product and thereby disrupting functional protein production by that gene.

10 In addition to disabling genes by deleting nucleotides, causing a transitional reading frameshift, disabling modifications would also be possible by other techniques including insertions, substitutions, inversions or transversions of nucleotides within the gene's DNA that would effectively prevent the formation of the protein coded for by the DNA.

It is also within the capabilities of one skilled in the art to disable genes by the use of less specific methods. Examples of less specific methods would be the use of chemical
15 mutagens such as hydroxylamine or nitrosoguanidine or the use of radiation mutagens such as gamma radiation or ultraviolet radiation to randomly mutate genes, such as the DGR genes of the present invention. Such mutated strains could, by chance, contain disabled gustatory receptor genes such that the genes are no longer capable of producing functional proteins for any one or more of the domains. The presence of the desired disabled genes could be detected
20 by routine screening techniques. For further guidance, see U.S. Patent No. 5,759,538.

Over-expression. Using the gustatory receptor genes of the present invention, it is now possible to selectively over-express specific DGR genes and look for changes in taste response and behavior. Parallel studies may be conducted in other organisms by using the gustatory receptor genes and the methods of the present invention to identify the gustatory
25 receptor genes of other organisms and then overexpress the gustatory receptor genes of those organisms.

Methods of overexpressing genes are generally known by those skilled in the art. For examples of producing cells which overexpress specific genes, see, for example, U.S. Patent

No. 5,905,146; 5,849,999; 5,859,311; 5,602,309; 5,952,169 & 5,772,997 (HER2 receptor).

Modulating or Inhibiting Expression. Using the gustatory receptor genes of the present invention, it is now possible to selectively modulate or inhibit specific DGR genes using antisense oligomers which specifically hybridize with the DNA or RNA encoding the DGR genes. One skilled in the art could so modulate or inhibit the expression of the DGR genes and detect for changes in taste response and behavior. Parallel studies may be conducted in other organisms by using the gustatory receptor genes and the methods of the present invention to identify the gustatory receptor genes in other organisms and then use antisense oligers to the gustatory receptor genes of those organisms. Methods for inhibiting expression of genes, especially genes coding for receptors, using antisense constructs, including generation of antisense sequences *in situ* are described, for example, in U.S. Patent No. 5,856,099; 5,556,956; 5,716,846; 5,135,917 & 6,004,814.

Other methods that can be used to inhibit expression of an endogenous gene are applicable to the present invention. For example, formation of a triple helix at an essential region of a duplex gene serves this purpose. The triplex code, permitting design of the proper single stranded participant is also known in the art. (See Moser *et al.*, (1987) Science 238, 645-650 and Cooney *et al.*, (1988) Science 241, 456-459). Regions in the control sequences containing stretches of purine bases are particularly attractive targets. Triple helix formation along with photo-crosslinking is described, *e.g.*, in Praseuth *et al.*, (1988) Proc. Natl Acad. Sci. USA 85, 1349-1353.

Studying Behavior. Behavioral studies may help organize the gustatory systems in various organisms and may help explain the behavior of various organisms.

The tools and methodologies of the present invention may be used to study the influence of environmental conditions on eating behavior. For example, newly identified gustatory receptor genes may be used to study the effects of different preferences for a particular food source.

In one embodiment, modulation of gustatory receptor activity can be measured by the probosis extension response assay. When gustatory sensilla on either the labellum or the leg

are stimulated with a sugar solution, the fly extends its mouthparts, in a behavior known as the proboscis extension response (PER). A variety of stimuli, including bitter compounds and high concentrations of salts, inhibit the PER when added to the sugar solution. The PER depends on the dose of the sugar solution, and the inhibition by other compounds is

5 dose-dependent as well. The PER is simple to measure, and can be quantitated precisely. It has been characterized in great detail, initially in the classic experiments of Dethier on large flies such as the blowfly *Phormia regina* (Dethier (1955) *Quart. Rev. Biol.* 30, 348-371; Dethier, (1976) *The Hungry Fly*, Harvard Press).

In *Drosophila*, a PER has been shown to be elicited by sugar, and inhibited when

10 NaCl is added to the sugar (Arora, (1987) *Nature* 330, 62-63; Rodrigues & Siddiqi (1978) *Proc. Ind. Acad. Sci.* 87B, 147-160; Tompkins *et al.*, (1979) *Proc. Natl. Acad. Sci. USA* 76, 884-887).

In yet another embodiment, gustatory receptor activation assays may be based on the fact that flies demonstrate strong preferences when presented with two taste stimuli. Using a

15 countercurrent behavioral paradigm in which flies make a series of binary choices between a sucrose medium either with or without quinine sulfate, it has been shown that flies preferentially distribute onto the medium without quinine (Tompkins *et al.*, (1979) *Proc. Natl. Acad. Sci. USA* 76, 884-887), which tastes bitter to humans. Flies manifest preferences for different sugar solutions, as shown in an elegant paradigm in which animals are allowed to

20 feed from the wells of a microtiter dish (Tanimura *et al.*, (1982) *J. Comp. Physiol.* 147, 433-437). Wells of the dish contain agar, with alternate wells containing one of two sugars. Wells containing one sugar are marked with red dye; those containing the other sugar are marked with blue dye. After feeding in the dark, flies are classified according to the color of their abdomen (red, blue, or mixed), which provides a quantitative indication of their feeding

25 preferences, which can be used as a measure for the activity of any particular gustatory receptor.

4. For Organism Detection, Monitoring and Control.

General Pest Management. The gustatory receptor genes identified herein and

identified using the methods of the present invention may be used to identify compounds which may be used for pest management. It is especially desirable to utilize various aspects of the present invention for pest management related to crop protection.

5 The application of pheromones is now firmly established as a key component of pest management and control, especially within the framework of integrated pest management (IPM). An object of organism control is to modulate an organism's behavior or activity so as to reduce the irritation, sickness, or death of the host (e.g., a plant host), or to decrease the general health and proliferation of the organism.

10 For example, the propagation of a mouse population in a given area of actual or potential mice infestation may be prevented or inhibited by a bait containing an effective amount of a first compound which the mice prefer to eat, wherein such compounds could be combined with a second compound, such as a pheromone, which would attract the mice to the bait and would also be combined with a third compound which would have lethal effects on the mice. Thus, in a preferred embodiment, the mice would be attracted to the area by the
15 odor of the second compound, enticed to eat a large amount of the bait because of the taste of the first compound and would die as a result of the presence of the third compound in the bait.

Compositions for attracting insects generally require some physical and/or chemical means for attracting the insects to a bait. In addition, the bait needs to be fully attractive to the taste of the insect so as to induce the attracted insect to ingest the bait. Finally, the bait must
20 be taken in by the insect at a sufficient lethal dose before disgusting the insect in some way or producing a toxic reaction in the insect (see, for example, U.S. Patent No. 4,855,133).

Insect Repellents and Insecticides. The present invention provides the tools and methodologies useful for identifying compounds which modulate insect behavior by exploiting the sensory capabilities of the target insect. For example, attempts have been made
25 to describe and synthesize the complex interactions which underlie host-seeking behavior in mosquitoes. Using the methods and gustatory receptor genes of the present invention, it is possible to design specific compounds which target mosquito gustatory receptor genes. Thus, the present invention provides the ability to alter or to eliminate the orientation and feeding

behaviors of mosquitoes and thereby have a positive impact on world health by controlling mosquito-borne diseases, such as malaria.

Mosquito gustatory receptor genes may be identified and/or targeted using various aspects of the present invention. For example, the gustatory receptor genes of the present invention may be used to design probes as discussed elsewhere herein for the identification and characterization of mosquito gustatory receptor genes. Alternatively, the algorithm of the present invention may be used to identify mosquito gustatory receptor genes in the genetic databases for mosquitoes. Once the mosquito gustatory receptor genes are identified, then various screening methods described elsewhere herein, such as the high throughput assays discussed elsewhere herein, may be used to identify synthetic and natural compounds which may modulate the behavior of the insect.

For general information on insect repellents, see, for example, U.S. Patent No. 4,663,346.

Mating Enhancement and Disruption. The gustatory receptor genes identified herein and identified using the methods of the present invention may be used to identify compounds which interfere with the orientation and mating of a wide range of organisms, including insects. Thus, the present invention enables the identification of compositions which disrupt insect mating by selective inhibition of specific receptor genes involved in mating attraction (see, *e.g.*, U.S. Patent No. 5,064,820).

Animal Repellants. The gustatory receptor genes identified herein and identified using the methods of the present invention may be used to identify compounds which may be used as animal repellants. Such compositions may be used to repel both predatory and non-predatory animals (see, *e.g.*, U.S. Patent No. 4,668,455).

5. Organism Attraction.

Insect Attractants. The gustatory receptor genes identified herein and identified using the methods of the present invention may be used to identify compounds which attract specific insects to a particular location (see, *e.g.*, U.S. Patent No. 4,880,624 & 4,851,218).

For example, aspects of the present invention may be used in various methods which

reduce or eliminate the levels of particular insect pests by selective attraction of a particular insect or pest, such as mosquitoes and tsetse flies. As a particular example, insect traps can be created wherein the taste of a compound selectively attracts a particular insect, like the tsetse fly, and the insect so attracted dies in the trap. Once in the trap, the attraction is maintained by stimulation of a particular gustatory receptor of the invention. In this way, the population of tsetse flies may be reduced or eliminated in a particular area.

The identified compositions which selectively attract and maintain the attraction by stimulation of gustatory receptors may also be combined with an insecticide, for example as an insect bait in microencapsulated form. Alternatively, or in addition, the insect attractant composition may be placed inside an insect trap, or in the vicinity of the entrance to an insect trap.

In addition to killing insects, the trapping of insects is often very important for estimating or calculating how many insects of a particular type are feeding within a specific area. Such estimates are used to determine where and when insecticide spraying should be commenced and terminated.

Insect traps which may be used are, for example, those as described in U.S. Patent No. 5,713,153. Specific examples of insect traps include, but are not limited to, the Gypsy Moth Delta Trap®, Boll Weevil Scout Trap®, Jackson trap, Japanese beetle trap, McPhail trap, Pherocon 1C trap, Pherocon II trap, Pherocon AM trap and Trogo trap.

Kairomones may be used as an attractancy for the enhancement of the pollination of selected plant species.

Attractant compositions which demonstrate biological activity toward one sex which is greater than toward the opposite sex may be useful in trapping one sex of a specific organism over another. For example, a composition may be a highly effective attractant for male apple ermine moths (*Yponomeuta malinellus* (Zeller)) and not so effective an attractant for female apple ermine moths. By attracting and maintaining the attraction of adult males to field traps, the composition provides a means for detecting, monitoring, and controlling this agricultural pest (see, e.g., U.S. Patent No. 5,380,524).

Attracting Predators and Parasitoids. The gustatory receptor genes of the present invention and the gustatory receptor genes identified using the methods of the present invention may also be used to identify chemicals which attract and maintain the attraction of various predators and parasitoids. Attracting the predators and parasitoids which attack
5 certain pests offers an alternative method of pest management.

Animal Attractants. The gustatory receptor genes identified herein and those identified by the methods of the present invention may be used to identify chemicals which attract household domesticated animals. For example, a pheromone-containing litter preparation may attract the animals and absorb liquids and liquid-containing waste released by
10 the attracted animal (see, *e.g.*, U.S. Patent No. 5,415,131).

6. Industrial Applications. The gustatory receptor genes identified by the methods of the present invention may be used for a number of different industrial applications including, but not limited to the following:

(a) Identification of appetite suppressant compounds and using same to suppress
15 and/or control appetite.

(b) As Biosensors.

(1) Explosive and drug detectors. The detectors may be synthetic, such as biologically-inspired robotic sensors, or biological sensors, such as insects which are especially sensitive to certain tastes.

20 (2) Population of gustatory receptor genes expressed in cell culture.

Gustatory receptor genes can be introduced into a cell line and the transformed cells maintained in culture through multiple generations. By creating specific cell lines which express multiple gustatory genes at once, it would be possible to use such cell cultures to investigate how compounds interact with taste receptor genes. Thus, the present invention
25 provides methods for identifying taste fingerprints, wherein such methods include contacting a series of cells containing and expressing known gustatory receptor genes with a desired sample, and determining the type and quantity of the gustatory receptor ligands present in the sample (see, *e.g.*, U.S. Patent No. 5,993,778). As discussed elsewhere herein, the interaction

of substances with the receptors can be identified using appropriate labels, such as those provided by luciferase, the jellyfish green fluorescent protein (GFP) or β -galactosidase.

(3) Biochip Arrays. As discussed elsewhere herein, biochip arrays of gustatory receptor genes can be generated. The arrays may be used to detect gustatory receptor ligands via an appropriate marker or via a chemical or electrical signal. Arrays may be designed for specific purposes, such as, but not limited to, detecting perfumes, explosives, drugs, pollutants, and toxins.

(c) Training organisms to conduct certain tasks. Examples include, but are not limited to, orienting or reorienting the behavior of worker bees of a rearing colony by incorporating a composition which includes one or more pheromones which elicits particular bee behavior towards the larvae. Thus, the beekeeper may orient or reorient the bees towards a particular activity such as, but not limited to, inducing improved acceptance of the larvae at the beginning of rearing, to increase the production of royal jelly, regulate the feeding of the larvae as to favor the development of queen bees, etc. (see, e.g., U.S. Patent No. 5,695,383).

Without further description, it is believed that one of ordinary skill in the art can, using the preceding description and the following illustrative examples, make and utilize the compounds of the present invention and practice the claimed methods. The following working examples therefore, specifically point out the preferred embodiments of the present invention, and are not to be construed as limiting in any way the remainder of the disclosure.

EXAMPLES

Example 1 Identification of candidate taste receptor genes.

With approximately 100% of the *Drosophila* genome sequenced, the *Drosophila* gustatory receptor genes have been sequenced. A multi-step strategy was developed to identify taste receptor genes from the genomic database. First, a computer algorithm was designed to search the *Drosophila* genomic sequence for open reading frames (ORFs) from candidate taste receptor genes. Second, RT-PCR was used to determine if transcripts from

any of these ORFs identified through this approach were expressed in specific tissues and organs, including taste tissue deficient in chemosensory neurons.

Example 2 Algorithm for identification of G protein-coupled receptors (GPCR) genes.

5 A computer algorithm that seeks proteins with particular structural properties, as opposed to proteins with particular sequences, identified a large family of candidate gustatory receptors from the *Drosophila* genomic database (Clyne *et al.*, (1999) Neuron 22, 327-338 incorporated herein in its entirety). The algorithm examines the physicochemical properties of the amino acids in an open reading frame (ORF) and then uses a non-parametric discriminant
10 function to identify ORFs likely to encode multitransmembrane domain proteins.

 The algorithm used to identify G protein-coupled receptors (GPCR) genes used statistical characterization of amino acid physico-chemical profiles in combination with a non-parametric discriminant function. The key approach is to use the information in the interplay between the local structure (transmembrane alpha helix) and the global structure
15 (repeated multiple domains) and characterize this information with concise statistical variables.

 The algorithm was trained on a set of one-hundred putative GPCR sequences from the GPCR database (GPCRDB) at <http://swift.embl-heidelberg.de/7tm> and a set of one-hundred random proteins selected from the SWISSPROT database (this training set was later
20 expanded, but that version was not used for the genes reported in this paper). In the first step, three sets of descriptors were used to summarize the physico-chemical profiles of the sequences. These were GES scale of hydropathy (Engelman *et al.*, (1986) Annu. Rev. Biophys. Biophys. Chem. 15, 321-353), polarity (Brown, (1991) Molecular Biology Labfax, Academic Press), and amino acid usage frequency. For the first two of these measurements, a
25 sliding window profile was employed (White, (1994) Membrane Protein Structure, Oxford University Press) using a kernel of 15 amino acid constant function convoluted with a 16 amino acid Gaussian function.

 These profiles were then summarized with three statistics; the periodicity

(characterizing the quasi-periodic presence of the transmembrane domain), average derivative (characterizing the abrupt change between the transmembrane domain and non-transmembrane domain), and the variance of the derivative (also characterizing the abrupt change). GES periodicity, variance of polarity derivative, polarity periodicity and amino acid frequency were used as the four variables and each sequence was therefore characterized by four variables. These four variables were used in a non-parametric linear discriminant function that was then optimized to separate the known GPCRs from random proteins in the training set. The same linear discriminant function with the scores derived from the training set was then used to screen the genomic database for candidate genes.

The candidate sequences were given significance values by an odds ratio of the GPCRs and non-GPCRs computed using the observed empirical distribution of the training set. More detailed information about the algorithm is available at <http://www.neuron.org/cgi/content/full/22/2/327/dc1>.

The computational screens used the genomic sequence data obtained by FTP from the Berkeley *Drosophila* Genome Project (BDGP, <http://www.fruitfly.org>, version 6/98). First, the ORFs of 300 bases or longer in all six frames were identified. Next, a program written to identify GPCRs statistically by their physico-chemical profile was used to screen for candidate ORFs as described above. The number of possible candidates was reduced by comparing them to *Drosophila* codon usage tables (<http://flybase.bio.indiana.edu>, version 10). Candidate ORFs whose codon usage differed at a significance level of 0.0005 by the chi-square statistic were discarded from the candidate set. Using these screening steps, thirty-four candidate ORFs were obtained.

Example 3 Further analysis of genes identified by the algorithm.

Further analysis of genes identified by this algorithm revealed one gene that led to the definition of a distinct large DGR family of membrane proteins. Forty-three members of this family have been identified in the complete *Drosophila* genome. If the sequenced part of the genome is representative, then extrapolation suggests that the entire genome would encode on

the order of 75 DGR proteins, a figure comparable to previous estimates of 100 candidate *Drosophila* odorant receptors (DOR), as described in Clyne *et al.*, (1999) Neuron 22, 327-338).

This previously unidentified family of proteins shows no sequence similarities to odorant receptors or to other known proteins. This family of proteins has been designated the gustatory receptor (GR) family, with each individual gene named according to its cytogenetic location in the genome. Thus, the GR59D.1 and GR59D.2 genes, which was abbreviated here as 59D.1 and 59D.2, refer to two family members located in cytogenetic region 59D on the second chromosome. This designation of location, however, does not reflect additions to the *Drosophila* genome subsequent to the discovery of the gustatory receptor genes.

The first exon of 23A.1b (Figure 1) was identified by the computer algorithm described in Clyne *et al.*, (1999) Neuron 22, 327-338, as described in Example 2, above.

Examination of the genomic DNA surrounding the first exon of 23A.1b identified other exons, and the genomic structure of this gene was determined with RT-PCR.

Using the sequence of this gene, an extensive series of tBLASTn searches of the Berkeley *Drosophila* Genome Project (BDGP) sequence database (available at <http://www.fruitfly.org>) was performed, which identified ORFs of thirty-eight other genes of the GR family. The full sequences of these genes were identified by an analysis of the genomic DNA flanking these ORFs as described in Clyne *et al.*, (1999) Neuron 22, 327-338 using the *Drosophila* intron-exon consensus splice sequences and RT-PCR analysis. The thirty-nine genes encode a total of forty-three proteins.

The National Center for Biotechnology Information (NCBI) accession number of the BDGP genomic clone on which each transcript is found and the sequence range in the genomic clone for the predicted coding region are given as follows for each GR transcript shown in Figure 1 (NCBI and BDGP data are as of 16 October 1999): transcript GR21D.1, accession number AC004420, range 34784 - 33509; GR22B.1, AC003945, 31740 - 30551; GR23A.1a, AC005558, 108490 - 106118; GR23A.1b, AC005558, 107351 - 106118; GR32D.1, AC005115, 19779 - 21141; GR39D.1, AC007208, 62553 - 64348; GR39D.2a,

AC005130, 9170 - 16119; GR39D.2b, AC005130, 10410 - 16119; GR39D.2c, AC005130, 12989 - 16119; GR39D.2d, AC005130, 14750 - 16119; GR43C.1, AC005452, 50105 - 51583; GR47A.1, AC007352, 114644 - 115920; GR58A.1, AC004368, 62323 - 61087; GR58A.2, AC004368, 62511 - 63791; GR58A.3, AC004368, 65521 - 64229; GR59D.1, 5
AC006245, 68825 - 70050; GR59D.2, AC006245, 70261 - 71505; GR59E.1, AC005639, 30167D31539; and GR59E.2, AC005639, 30036 - 28714.

Accession numbers for the other genes are as follows (data are as of 5 January 2000) (complete sequences are available for the first four and only partial sequences are available for the remaining genes; LU, location unknown): transcript GR1F.1, accession number
10 AL035632, range 7301 - 8711; GR47F.1, AC005653, 42838 - 44204; GR68D.1, AC006492, 46040 - 44916; GR77E.1, AC006490, 104929 - 103117; GR28A.1, AC008354, 66711 - 66973; GR57B.1, AC007837, 102661 - 103185; GR65C.1, AC004251, 23136 - 24215; GR93F.1, AC012873, 35043 - 35228; GR93F.2, AC012892, 2781 - 2650; GR93F.3, AC012892, 4271 - 4143; GR93F.4, AC012892, 6482 - 5559; GR94E.1, AC008200, 72472 - 72308; GR97D.1, AC007984, 121300 - 121977; GR98B.1, AC007817, 45506 - 46916;
15 GR98B.2, AC007817, 10695 - 10784; GR98B.3, AC007817, 45189 - 45284; GR98B.4, AC007817, 39658 - 39765; GRLU.1, AC017438, 22141 - 21398; GRLU.2, AC017138, 10997 - 11122; GRLU.3, AC015395, 43210 - 43612; GRLU.4, BACR28P1-T7, 28 - 129; GRLU.5, BACR28P1-T7, 388 - 734; GRLU.6, BACR06I03-T7, 1028 - 48; and GRLU.7,
20 AC012799, 8212 - 8123.

Example 4 Sequence analysis of DGR genes.

The amino acid sequences of nineteen members of the GR family indicate the high degree of sequence divergence (Figure 1). Sequence alignment revealed only one residue
25 conserved among all members of the family shown and only 24 residues conserved among more than half of the genes shown. Fifteen of these conserved residues lie in the vicinity of the COOH-terminus. Amino acid identity between individual genes ranged from a maximum of 34% to less than 10%. By contrast, other features of the gene family show

substantial conservation. The positions of a number of introns are conserved (Figure 1), suggesting that the family originated from a common ancestral gene. Overall sequence length, ~380 amino acids, is another common feature. All of the genes encode approximately seven predicted transmembrane domains, a feature characteristic of G protein-coupled receptors (GPCRs) (Figure 2).

The GR proteins were identified as GPCRs when the algorithm was modified to distinguish previously described GPCRs from ion channels. The algorithm was set to positively identify 95% of previously described GPCRs, with 4.3% false positives. Most ion channels have six transmembrane domains.

The genes are widely dispersed in the genome, but at the same time, many are found in clusters. The two largest clusters each contain four genes; there are also several clusters of two or three genes. Genes within these clusters are closely spaced, with intergenic distances ranging from 150 to 450 base pairs (bp) in all cases for which the data were available. There is no rule specifying the orientation of genes within clusters, unlike the case of the *Drosophila* odorant receptors, in which genes within a cluster are in the same orientation in all clusters examined (Clyne *et al.*, (1999) Neuron 22, 327-338).

An unusual form of alternative splicing occurs in at least two chromosomal locations. Four large exons in cytogenetic region 39D each contain sequences specifying six predicted transmembrane domains, followed by three small exons that together specify a putative seventh transmembrane domain and the COOH-terminus (Figure 3). Reverse transcription-polymerase chain reaction (RT-PCR) analysis revealed that each of the four large exons is spliced to the smaller exons, thereby generating four predicted seven transmembrane domain proteins. These four proteins are thus distinct through the first six transmembrane domains and identical in the seventh and in the COOH-terminal sequences. Likewise, in cytogenetic region 23A, there are two large exons, each of which specifies six transmembrane domains and is spliced to two small exons that together encode a seventh transmembrane domain and the COOH-terminus (Figure 3). Thus, the gene in region 23A encodes two related proteins. This pattern of splicing, in which alternative large 5' exons encoding most of the protein are

joined to common short 3' exons encoding only a small portion of the protein, is unusual among genes encoding GPCRs and proteins in general. This pattern of splicing provides a mechanism at a single locus for generating products that exhibit a pattern observed for this family in general: extreme diversity among all sequences of the proteins except in a small region in the vicinity of the COOH-terminus.

Example 5 Identification of gustatory receptor genes using RT-PCR.

To assess the tissue specificity of expression, RT-PCR with primers that span introns in the coding regions was performed. Of the 19 transcripts tested, 18 were expressed in the labellum (Figure 4 and Table 1), the major gustatory organ of the fly (Falk *et al.*, (1976) J. Morphol. 150, 327; Dethier, (1976) The Hungry Fly, Harvard University Press; Stocker, (1994) Cell Tissue Res. 275, 3; Nayak & Singh, (1983) Int. J. Insect Morphol. Embryol. 12, 273. Moreover, for most of these genes, expression was labellum-specific in that only 1 of the 19 yielded amplification products from heads depleted of taste organs and only 2 showed expression in the thorax, which contains the thoracic nervous system but no characterized taste sensilla. Likewise, expression in several other tissues, including the abdomen, wings, and legs, was limited to a small fraction of genes (Table 1).

For preparation of RNA, individual flies were frozen in liquid nitrogen, and labella were dissected. On average, 50 labella were used for RNA preparation. Total RNA was prepared as described elsewhere (McKenna *et al.*, (1994) J. Biol. Chem. 269, 16340-16347). The RNA was treated with DNaseI (Gibco-BRL) for thirty minutes at 37°C, phenol/chloroform extracted, and precipitated. The entire RNA preparation was used for oligo dT-primed cDNA synthesis using Superscript II Reverse Transcriptase (Gibco-BRL) according to the manufacturer's directions. PCR was performed using Taq polymerase (Sigma) under standard cycling conditions, with an annealing temperature of 60°C, gene-specific primer concentration of 1 pM, and magnesium concentration of 2.5 mM. For all genes, primer pairs which span introns were used in order to distinguish PCR bands amplified from cDNA from those amplified from any remaining genomic DNA.

Example 6 Tissue specificity of GR gene expression

To further analyze the tissue specificity of GR expression, a microdissection experiment was performed in which the labral sense organ (LSO) (Stocker, (1994) Cell Tissue Res. 275, 3; Nayak & Singh, (1983) Int. J. Insect Morphol. Embryol. 12, 273), a small taste organ that lines the pharynx, was surgically excised from each of fifty animals. The LSO consists of a very limited number of cells and is highly enriched in taste neurons; it does not, for example, contain muscle cells. By RT-PCR amplification, the expression of seven GR transcripts in this taste organ was detected (Figure 5). These results indicate that expression of the GR family extends to include at least one additional taste organ besides the labellum. The data are also fully consistent with the notion that the GR genes are expressed in taste neurons.

To confirm the gene expression in taste receptor neurons, a *Drosophila* mutant, *pox-neuro 70* (*poxn 70*) was used, in which chemosensory bristles are transformed into mechanosensory bristles (Awasaki & Kimura, (1997) J. Neurobiol. 32, 707; Dambly-Chaudiere *et al.*, (1992) Cell 69, 159; Nottebohm *et al.*, (1994) Neuron 12, 25; Nottebohm *et al.*, (1992) Nature 359, 829).

Specifically, in *poxn 70*, which behaves as a null mutation with respect to adult chemosensory organs, chemosensory bristles are transformed into mechanosensory bristles with respect to various morphological and developmental criteria. In particular, most chemosensory bristles in wild-type *Drosophila* are innervated by five neurons: four chemosensory neurons and one mechanosensory neuron. In contrast, wild-type mechanosensory bristles contain a single mechanosensory neuron.

In chemosensory bristles transformed to mechanosensory bristles by *poxn 70* (Awasaki & Kimura, (1997) J. Neurobiol. 32, 707), the number of neurons is reduced from five to one. We predicted that if the GR family is in fact expressed in the chemosensory neurons of taste sensilla, their expression would likely be eliminated in the *poxn 70* mutant. Consistent with this prediction, eighteen of nineteen GR transcripts examined were not

expressed in the labellum of the *poxn 70* mutant (Table 1 and Figure 4). RT-PCR was performed from RNA extracted from the indicated tissues (see description for Figure 4). All primer pairs spanned introns. Positive controls are described in Figure 4. These results indicate that the GR gene family is expressed in labellar chemosensory neurons.

5

Table 1. Tissue-specific expression of GR genes.

	Gene	Labellum	<i>poxn</i> Labellum	Head minus taste organs	Thorax	Abdomen	Leg	Wing
5	21D.1	+	-	-	-	-	-	-
	22B.1	+	-	-	-	+	+	+
	23A.1a	+	-	-	-	-	-	-
	23A.1b	+	-	-	-	-	-	-
	32D.1	+	-	-	-	-	-	-
10	39D.1	+	-	-	-	-	-	-
	39D.2a	+	-	-	-	-	-	-
	39D.2b	+	-	-	-	-	-	-
	39D.2c	+	+	-	+	+	-	-
	39D.2d	+	-	-	+	-	-	+
15	43C.1	+	-	+	-	+	+	+
	47A.1	+	-	-	-	-	-	-
	58A.1	+	-	-	-	-	-	-
	58A.2	+	-	-	+	-	-	-
	58A.3	+	-	-	-	-	-	-
20	59D.1	+	-	-	-	-	-	-
	59D.2	+	-	-	-	-	-	-
	59E.1	-	-	-	-	+	-	+
	59E.2	+	-	-	-	-	-	-

Example 7 Receptor diversity.

- 25 The large size of this protein family likely reflects the diversity of compounds that flies can detect. The extreme diversity of these receptors may not only reflect diversity among

the ligands that they bind, but also diversity in the signal transduction components with which they interact. For example, the lack of conserved intracellular regions suggests the possibility that, during the evolution of this sensory modality, multiple G proteins arose, each interacting with a different subset of receptors. Finally, it seems likely that the *Drosophila* genome en-

5 codes taste receptors in addition to those of the GR family.

Although applicants have detected expression in the labellum and the LSO, few if any family members are expressed in the leg or wing chemosensory hairs (Table 1), some of which are morphologically similar to labellar taste hairs (Stocker, (1994) Cell Tissue Res. 275, 3). The *Drosophila* olfactory system also contains more than one organ, the antenna and

10 maxillary palp, which respond to all, or nearly all, of the same odorants and which derive from the same imaginal discs (Carlson, (1996) Trends Genet. 12, 175). However, most individual members of the DOR gene family are expressed in one or the other but not in both olfactory organs (Clyne, (1999) Neuron. 22, 327; Vosshall *et al.*, (1999) Cell 96, 725). Perhaps the distinction among taste receptor genes is even more extreme in the gustatory system, whose

15 organs derive from different imaginal discs. For example, the legs may express a completely distinct family of genes or a subfamily whose similarities to the present family are sufficiently tenuous as to place it slightly beyond the boundaries that define the GR family.

Example 8 Transgenic *Drosophila*

20 P element mediated germline transformation of *Drosophila* can be carried out as previously described (Rubin & Spradling, (1982) Science 218, 348-353). *Drosophila* embryos are isolated and microinjected with P element expression constructs as previously described (Karess & Rubin, (1984) Cell 38, 135-146) containing a particular DGR nucleotide sequence, at 0.5 mg/ml together with a helper plasmid at 0.1 mg/ml.

25 Non-transformed (Generation 0 or G₀) injected adults are individually back crossed to the recipient strain and the G₁ progeny screened for the w⁺ transformation marker (Klemenz *et al.*, (1987) Nucleic Acids Res. 10, 3947-3959). Transformed lines homozygous for the transgene are established from orange eyed G₁ flies as previously described (Klemenz *et al.*,

(1987) Nucleic Acids Res. 10, 3947-3959).

A line of *Drosophila* in which the 39D.2c gene can be over-expressed is constructed as described above. The 39D.2c coding sequences are joined to an upstream activating sequence (UAS) and introduced by P element-mediated germline transformation into

5 *Drosophila*. A yeast GAL4 transcription factor gene, coupled to a heat shock promoter, is then crossed into the transgenic line. As expected, heat shock of this line results in induction of 39D.2c expression. The heat shock-induced expression of GAL4, also results in binding of GAL4 to the UAS, and subsequent induction of 39D.2c expression. This transgenic line of

10 *Drosophila*, and three other transgenic lines containing other DGR genes, can be tested for elevated responses to any of fifty different tastes. Elevated response to any particular taste is indicative of an ligand which binds and activates the over-expressed receptor (see, e.g., Zhao & Firestein, (1998) Science 279, 237-242).

Although the present invention has been described in detail with reference to

15 examples above, it is understood that various modifications can be made without departing from the spirit of the invention. Accordingly, the invention is limited only by the following claims. All cited patents and publications referred to in this application are herein incorporated by reference in their entirety. The results of the experiments disclosed herein have been published in the journal Science (2000) 287, 1830-1834, this article herein

20 incorporated by reference in its entirety.

EDITORIAL NOTE

APPLICATION NUMBER – 58722/00

The following Sequence Listing pages 1 to 154 are part of the description. The claims pages follow on pages "62" to "66".

SEQUENCE LISTING

<110> Carlson, John R.
Clyne, Peter J.
Warr, Coral G.
Yale University

<120> Novel Taste Receptors in Drosophila

<130> 44574-5072

<140>

<141>

<150> US 60/138,668

<151> 1999-06-14

<150> US 60/181,704

<151> 2000-02-10

<160> 94

<170> PatentIn Ver. 2.1

<210> 1

<211> 1242

<212> DNA

<213> Drosophila melanogaster

<220>

<221> CDS

<222> (1)..(1242)

<223> Coding region GR1F.1

<400> 1

atg gaa ttc ggc atg gac acg ctg aga gct ctg gag ccg ctg cac cgc 48

Met Glu Phe Gly Met Asp Thr Leu Arg Ala Leu Glu Pro Leu His Arg

1 5 10 15

gcc tgc cag gtg tgc aac cta tgg ccc tgg cgc ctc gcc ccc ccg cca 96

Ala Cys Gln Val Cys Asn Leu Trp Pro Trp Arg Leu Ala Pro Pro Pro

20 25 30

gac tcg gag ggc atc ctg ctc cgg cga tcg cgc tgg ctg gag ctg tac 144

Asp Ser Glu Gly Ile Leu Leu Arg Arg Ser Arg Trp Leu Glu Leu Tyr

35 40 45

ggc tgg acc gtg ctg ata gcg gcc acc tcc ttc acc gtg tac ggc ctc 192

Gly Trp Thr Val Leu Ile Ala Ala Thr Ser Phe Thr Val Tyr Gly Leu

50 55 60

ttc cag gag agc agc gta gaa gag aaa cag gac tcg gag tcc acc atc 240

Phe Gln Glu Ser Ser Val Glu Glu Lys Gln Asp Ser Glu Ser Thr Ile

65 70 75 80

tcg agc ata ggt cac acg gtt gac ttc att cag ctg gtg ggc atg cgg 288

1

Ser	Ser	Ile	Gly	His	Thr	Val	Asp	Phe	Ile	Gln	Leu	Val	Gly	Met	Arg		
				85					90					95			
gtg	gct	cat	ctg	gct	gcc	ctg	ctg	gag	gcc	ctg	tgg	cag	cgg	cag	gcg	336	
Val	Ala	His	Leu	Ala	Ala	Leu	Leu	Glu	Ala	Leu	Trp	Gln	Arg	Gln	Ala		
			100					105					110				
cag	cga	ggc	ttc	ttt	gcg	gag	cta	ggc	gag	atc	gat	cgc	ctg	ctg	tcc	384	
Gln	Arg	Gly	Phe	Phe	Ala	Glu	Leu	Gly	Glu	Ile	Asp	Arg	Leu	Leu	Ser		
		115				120					125						
aag	gcg	ttg	agg	gtg	gat	gtg	gag	gct	atg	cgc	atc	aat	atg	cgt	cgc	432	
Lys	Ala	Leu	Arg	Val	Asp	Val	Glu	Ala	Met	Arg	Ile	Asn	Met	Arg	Arg		
		130				135					140						
cag	acg	tcg	cgc	cga	gct	gtg	tgg	atc	ctg	tgg	ggc	tat	gcg	gtt	agc	480	
Gln	Thr	Ser	Arg	Arg	Ala	Val	Trp	Ile	Leu	Trp	Gly	Tyr	Ala	Val	Ser		
145					150					155				160			
caa	ctc	ctt	atc	ctg	ggc	gcc	aag	ctg	ctc	tcc	cgc	gga	gac	aga	ttc	528	
Gln	Leu	Leu	Ile	Leu	Gly	Ala	Lys	Leu	Leu	Ser	Arg	Gly	Asp	Arg	Phe		
			165					170					175				
cgc	atc	tac	tgg	atc	agc	tac	ttg	ctg	cca	ctg	ctt	gtg	tgc	gga	ttg	576	
Pro	Ile	Tyr	Trp	Ile	Ser	Tyr	Leu	Leu	Pro	Leu	Leu	Val	Cys	Gly	Leu		
			180				185						190				
cgt	tac	ttt	cag	atc	ttc	aac	gcc	act	cag	ctc	gtg	cgc	cag	cgc	ctg	624	
Arg	Tyr	Phe	Gln	Ile	Phe	Asn	Ala	Thr	Gln	Leu	Val	Arg	Gln	Arg	Leu		
		195				200					205						
gat	gtg	ctc	cta	gtg	gcc	ttg	cag	cag	ctt	cag	ctg	cac	cag	aaa	ggg	672	
Asp	Val	Leu	Leu	Val	Ala	Leu	Gln	Gln	Leu	Gln	Leu	His	Gln	Lys	Gly		
		210				215					220						
ccc	gca	gtg	gat	act	gtg	ctt	gag	gag	cag	gag	gat	ctg	gaa	gaa	gca	720	
Pro	Ala	Val	Asp	Thr	Val	Leu	Glu	Glu	Gln	Glu	Asp	Leu	Glu	Glu	Ala		
		225				230				235				240			
gca	atg	gat	aga	ctg	atc	gct	gtc	aga	ctc	gtc	tac	caa	cgg	gtg	tgg	768	
Ala	Met	Asp	Arg	Leu	Ile	Ala	Val	Arg	Leu	Val	Tyr	Gln	Arg	Val	Trp		
			245					250					255				
gcc	cta	gtg	gcc	ttg	cta	aac	cgc	tgc	tac	ggg	ctc	tcc	atg	ttg	atg	816	
Ala	Leu	Val	Ala	Leu	Leu	Asn	Arg	Cys	Tyr	Gly	Leu	Ser	Met	Leu	Met		
		260					265						270				
cag	gtg	ggc	aac	gac	ttc	ctg	gct	atc	acc	tcc	aac	tgc	tac	tgg	atg	864	
Gln	Val	Gly	Asn	Asp	Phe	Leu	Ala	Ile	Thr	Ser	Asn	Cys	Tyr	Trp	Met		
		275				280						285					
ttc	ctc	aac	ttc	cgc	cag	tcg	gcg	gcc	tcg	ccc	ttc	gac	atc	ctg	caa	912	
Phe	Leu	Asn	Phe	Arg	Gln	Ser	Ala	Ala	Ser	Pro	Phe	Asp	Ile	Leu	Gln		
		290				295					300						
atc	gtg	gcc	agt	ggc	gta	tgg	tct	gcc	ccc	cac	ttg	ggg	aac	gtg	ctc	960	

```

Ile Val Ala Ser Gly Val Trp Ser Ala Pro His Leu Gly Asn Val Leu
305                310                315                320

gtc ctc tca ctg ctt tgc gac cga acg gcc cag tgc gcg tct cgt ctt 1008
Val Leu Ser Leu Leu Cys Asp Arg Thr Ala Gln Cys Ala Ser Arg Leu
                325                330                335

gcc ctg tgc ctg cac cag gta agc gtg gat tta agg aat gag agc cac 1056
Ala Leu Cys Leu His Gln Val Ser Val Asp Leu Arg Asn Glu Ser His
                340                345                350

aat gct ctg ata acc cag ttc tcg ctg cag cta ctc cac cag cgg ctc 1104
Asn Ala Leu Ile Thr Gln Phe Ser Leu Gln Leu Leu His Gln Arg Leu
                355                360                365

cac ttc agc gcc gct ggc ttc ttc aac gtg gac tgc acc ctt ctc tat 1152
His Phe Ser Ala Ala Gly Phe Phe Asn Val Asp Cys Thr Leu Leu Tyr
                370                375                380

acg atc gtg gga gcc act acc acg tac ttg ata atc ctg att cag ttt 1200
Thr Ile Val Gly Ala Thr Thr Tyr Leu Ile Ile Leu Ile Gln Phe
                385                390                395                400

cac atg agc gaa tcc acc atc ggc agt gat tcc aac gga cag 1242
His Met Ser Glu Ser Thr Ile Gly Ser Asp Ser Asn Gly Gln
                405                410

<210> 2
<211> 414
<212> PRT
<213> Drosophila melanogaster

<400> 2
Met Glu Phe Gly Met Asp Thr Leu Arg Ala Leu Glu Pro Leu His Arg
1          5          10          15

Ala Cys Gln Val Cys Asn Leu Trp Pro Trp Arg Leu Ala Pro Pro Pro
20          25          30

Asp Ser Glu Gly Ile Leu Leu Arg Arg Ser Arg Trp Leu Glu Leu Tyr
35          40          45

Gly Trp Thr Val Leu Ile Ala Ala Thr Ser Phe Thr Val Tyr Gly Leu
50          55          60

Phe Gln Glu Ser Ser Val Glu Glu Lys Gln Asp Ser Glu Ser Thr Ile
65          70          75          80

Ser Ser Ile Gly His Thr Val Asp Phe Ile Gln Leu Val Gly Met Arg
85          90          95

Val Ala His Leu Ala Ala Leu Leu Glu Ala Leu Trp Gln Arg Gln Ala
100         105         110

Gln Arg Gly Phe Phe Ala Glu Leu Gly Glu Ile Asp Arg Leu Leu Ser

```

115					120					125					
Lys	Ala	Leu	Arg	Val	Asp	Val	Glu	Ala	Met	Arg	Ile	Asn	Met	Arg	Arg
130					135					140					
Gln	Thr	Ser	Arg	Arg	Ala	Val	Trp	Ile	Leu	Trp	Gly	Tyr	Ala	Val	Ser
145					150					155					
Gln	Leu	Leu	Ile	Leu	Gly	Ala	Lys	Leu	Leu	Ser	Arg	Gly	Asp	Arg	Phe
165					170					175					
Pro	Ile	Tyr	Trp	Ile	Ser	Tyr	Leu	Leu	Pro	Leu	Leu	Val	Cys	Gly	Leu
180					185					190					
Arg	Tyr	Phe	Gln	Ile	Phe	Asn	Ala	Thr	Gln	Leu	Val	Arg	Gln	Arg	Leu
195					200					205					
Asp	Val	Leu	Leu	Val	Ala	Leu	Gln	Gln	Leu	Gln	Leu	His	Gln	Lys	Gly
210					215					220					
Pro	Ala	Val	Asp	Thr	Val	Leu	Glu	Glu	Gln	Glu	Asp	Leu	Glu	Glu	Ala
225					230					235					
Ala	Met	Asp	Arg	Leu	Ile	Ala	Val	Arg	Leu	Val	Tyr	Gln	Arg	Val	Trp
245					250					255					
Ala	Leu	Val	Ala	Leu	Leu	Asn	Arg	Cys	Tyr	Gly	Leu	Ser	Met	Leu	Met
260					265					270					
Gln	Val	Gly	Asn	Asp	Phe	Leu	Ala	Ile	Thr	Ser	Asn	Cys	Tyr	Trp	Met
275					280					285					
Phe	Leu	Asn	Phe	Arg	Gln	Ser	Ala	Ala	Ser	Pro	Phe	Asp	Ile	Leu	Gln
290					295					300					
Ile	Val	Ala	Ser	Gly	Val	Trp	Ser	Ala	Pro	His	Leu	Gly	Asn	Val	Leu
305					310					315					
Val	Leu	Ser	Leu	Leu	Cys	Asp	Arg	Thr	Ala	Gln	Cys	Ala	Ser	Arg	Leu
325					330					335					
Ala	Leu	Cys	Leu	His	Gln	Val	Ser	Val	Asp	Leu	Arg	Asn	Glu	Ser	His
340					345					350					
Asn	Ala	Leu	Ile	Thr	Gln	Phe	Ser	Leu	Gln	Leu	Leu	His	Gln	Arg	Leu
355					360					365					
His	Phe	Ser	Ala	Ala	Gly	Phe	Asn	Val	Asp	Cys	Thr	Leu	Leu	Tyr	
370					375					380					
Thr	Ile	Val	Gly	Ala	Thr	Thr	Tyr	Leu	Ile	Ile	Leu	Ile	Gln	Phe	
385					390					395					
His	Met	Ser	Glu	Ser	Thr	Ile	Gly	Ser	Asp	Ser	Asn	Gly	Gln		
405					410										

```

<210> 3
<211> 1335
<212> DNA
<213> Drosophila melanogaster

<220>
<221> CDS
<222> (1)..(1332)
<223> Coding region GR5A.1

<400> 3
atg cga cag ctg aaa ggg cga aat cgg tgc aat cgt gct gtt cgg cat 48
Met Arg Gln Leu Lys Gly Arg Asn Arg Cys Asn Arg Ala Val Arg His
1 5 10 15

ctg aaa gtt cag ggc aag atg tgg ttg aaa aat ctt aaa agc ggt ttg 96
Leu Lys Val Gln Gly Lys Met Trp Leu Lys Asn Leu Lys Ser Gly Leu
20 25 30

gaa caa att cga gaa agc cag gtg cgg gga acg cga aag aac ttc cta 144
Glu Gln Ile Arg Glu Ser Gln Val Arg Gly Thr Arg Lys Asn Phe Leu
35 40 45

cac gat ggc tcc ttc cat gag gca gtt gct cca gtt ctg gcg gtg gcc 192
His Asp Gly Ser Phe His Glu Ala Val Ala Pro Val Leu Ala Val Ala
50 55 60

cag tgc ttt tgc cta atg ccc gtg tgt gga atc agt gct cca acg tac 240
Gln Cys Phe Cys Leu Met Pro Val Cys Gly Ile Ser Ala Pro Thr Tyr
65 70 75 80

agg gga tta agc ttc aat cgg cgc agc tgg cga ttc tgg tac agc tcc 288
Arg Gly Leu Ser Phe Asn Arg Arg Ser Trp Arg Phe Trp Tyr Ser Ser
85 90 95

ttg tac ctc tgc tcc acg tcg gtg gat ctg gcc ttc agc att cgc cgt 336
Leu Tyr Leu Cys Ser Thr Ser Val Asp Leu Ala Phe Ser Ile Arg Arg
100 105 110

gtg gcc cac agt gtg ctg gat gtg cgc agt gtg gag ccc atc gtt ttt 384
Val Ala His Ser Val Leu Asp Val Arg Ser Val Glu Pro Ile Val Phe
115 120 125

cac gtt agc atc ctg atc gcc tcc tgg cag ttc ctg aac ctg gcc caa 432
His Val Ser Ile Leu Ile Ala Ser Trp Gln Phe Leu Asn Leu Ala Gln
130 135 140

ctc tgg cct gga ttg atg cgc cac tgg gcg gcg gtg gag cga cgg cta 480
Leu Trp Pro Gly Leu Met Arg His Trp Ala Ala Val Glu Arg Arg Leu
145 150 155 160

ccg ggc tac acc tgt tgc ctg caa cgt gct cgt cct gcc cgt cgc ctc 528
Pro Gly Tyr Thr Cys Cys Leu Gln Arg Ala Arg Pro Ala Arg Arg Leu
165 170 175

```

```

aag ttg gtg gcc ttc gtg ctg ctg gta gtt tca ctg atg gaa cac ctg 576
Lys Leu Val Ala Phe Val Leu Leu Val Val Ser Leu Met Glu His Leu
180 185 190

ctg agc atc att tcg gtg gtt tac tac gac ttc tgc cca cga aga agc 624
Leu Ser Ile Ile Ser Val Val Tyr Tyr Asp Phe Cys Pro Arg Arg Ser
195 200 205

gat ccc gtg gaa tcc tat ttg ctt ggc gcc agt gcc cag tta ttc gaa 672
Asp Pro Val Glu Ser Tyr Leu Leu Gly Ala Ser Ala Gln Leu Phe Glu
210 215 220

gtg ttc ccc tac tcc aac tgg ctg gct tgg ctg ggc aag atc cag aat 720
Val Phe Pro Tyr Ser Asn Trp Leu Ala Trp Leu Gly Lys Ile Gln Asn
225 230 235 240

gtg ctg ctc acc ttc ggc tgg agt tac atg gac ata ttc cta atg atg 768
Val Leu Leu Thr Phe Gly Trp Ser Tyr Met Asp Ile Phe Leu Met Met
245 250 255

ctg ggc atg ggt ctc agc gag atg ttg gcc agg ctg aac cgc agc ctg 816
Leu Gly Met Gly Leu Ser Glu Met Leu Ala Arg Leu Asn Arg Ser Leu
260 265 270

gag cag cag gtg cga cag ccc atg ccg gaa gcc tat tgg acg tgg tca 864
Glu Gln Gln Val Arg Gln Pro Met Pro Glu Ala Tyr Trp Thr Trp Ser
275 280 285

cgc act ctt tat cgc tcc ata gta gag ctc ata cgg gag gtg gat gac 912
Arg Thr Leu Tyr Arg Ser Ile Val Glu Leu Ile Arg Glu Val Asp Asp
290 295 300

gcc gtg tcc ggt ata atg ctg ata tcg ttt ggc agc aat ctg tac ttc 960
Ala Val Ser Gly Ile Met Leu Ile Ser Phe Gly Ser Asn Leu Tyr Phe
305 310 315

atc tgc ctg cag ctc ctc aag agc atc aac acc atg ccc tcc tcc gcc 1008
Ile Cys Leu Gln Leu Leu Lys Ser Ile Asn Thr Met Pro Ser Ser Ala
325 330 335

cac gcc gtc tac ttt tac ttc tcg ctg ctg ttc ctg ctc agc cga tca 1056
His Ala Val Tyr Phe Tyr Phe Ser Leu Leu Phe Leu Leu Ser Arg Ser
340 345 350

acg gcc gtg ctc ctc ttc gtt tcg gcg atc aac gac cag gcc agg gag 1104
Thr Ala Val Leu Leu Phe Val Ser Ala Ile Asn Asp Gln Ala Arg Glu
355 360 365

cca ctg cgt ctg ctg cgc ctt gtc ccg ctc aaa gga tat cat ccg gag 1152
Pro Leu Arg Leu Leu Arg Leu Val Pro Leu Lys Gly Tyr His Pro Glu
370 375 380

gtc ttc cgc ttc gcc gcc gaa ttg gcc agc gat cag gtg gcg ctt acg 1200
Val Phe Arg Phe Ala Ala Glu Leu Ala Ser Asp Gln Val Ala Leu Thr
385 390 395 400

```

ggg ctc aag ttc ttc aat gtc acc aga aag ttg ttc ctg gcg atg gcg 1248
 Gly Leu Lys Phe Phe Asn Val Thr Arg Lys Leu Phe Leu Ala Met Ala
 405 410 415

ggc acg gtg gcc act tat gaa cta gtg ctc att cag ttt cac gag gat 1296
 Gly Thr Val Ala Thr Tyr Glu Leu Val Leu Ile Gln Phe His Glu Asp
 420 425 430

aaa aag acc tgg gac tgc tca ccc ttc aat ctt gat tag 1335
 Lys Lys Thr Trp Asp Cys Ser Pro Phe Asn Leu Asp
 435 440

<210> 4

<211> 444

<212> FRT

<213> *Drosophila melanogaster*

<400> 4

Met Arg Gln Leu Lys Gly Arg Asn Arg Cys Asn Arg Ala Val Arg His
 1 5 10 15

Leu Lys Val Gln Gly Lys Met Trp Leu Lys Asn Leu Lys Ser Gly Leu
 20 25 30

Glu Gln Ile Arg Glu Ser Gln Val Arg Gly Thr Arg Lys Asn Phe Leu
 35 40 45

His Asp Gly Ser Phe His Glu Ala Val Ala Pro Val Leu Ala Val Ala
 50 55 60

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

Arg Gly Leu Ser Phe Asn Arg Arg Ser Trp Arg Phe Trp Tyr Ser Ser
 85 90 95

Leu Tyr Leu Cys Ser Thr Ser Val Asp Leu Ala Phe Ser Ile Arg Arg
 100 105 110

Val Ala His Ser Val Leu Asp Val Arg Ser Val Glu Pro Ile Val Phe
 115 120 125

His Val Ser Ile Leu Ile Ala Ser Trp Gln Phe Leu Asn Leu Ala Gln
 130 135 140

Leu Trp Pro Gly Leu Met Arg His Trp Ala Ala Val Glu Arg Arg Leu
 145 150 155 160

Pro Gly Tyr Thr Cys Cys Leu Gln Arg Ala Arg Pro Ala Arg Arg Leu
 165 170 175

Lys Leu Val Ala Phe Val Leu Leu Val Val Ser Leu Met Glu His Leu
 180 185 190

Leu Ser Ile Ile Ser Val Val Tyr Tyr Asp Phe Cys Pro Arg Arg Ser

195					200					205					
Asp	Pro	Val	Glu	Ser	Tyr	Leu	Leu	Gly	Ala	Ser	Ala	Gln	Leu	Phe	Glu
210						215				220					
Val	Phe	Pro	Tyr	Ser	Asn	Trp	Leu	Ala	Trp	Leu	Gly	Lys	Ile	Gln	Asn
225					230					235					240
Val	Leu	Leu	Thr	Phe	Gly	Trp	Ser	Tyr	Met	Asp	Ile	Phe	Leu	Met	Met
				245					250					255	
Leu	Gly	Met	Gly	Leu	Ser	Glu	Met	Leu	Ala	Arg	Leu	Asn	Arg	Ser	Leu
		260						265					270		
Glu	Gln	Gln	Val	Arg	Gln	Pro	Met	Pro	Glu	Ala	Tyr	Trp	Thr	Trp	Ser
		275					280					285			
Arg	Thr	Leu	Tyr	Arg	Ser	Ile	Val	Glu	Leu	Ile	Arg	Glu	Val	Asp	Asp
		290				295					300				
Ala	Val	Ser	Gly	Ile	Met	Leu	Ile	Ser	Phe	Gly	Ser	Asn	Leu	Tyr	Phe
305					310					315					320
Ile	Cys	Leu	Gln	Leu	Leu	Lys	Ser	Ile	Asn	Thr	Met	Pro	Ser	Ser	Ala
			325						330					335	
His	Ala	Val	Tyr	Phe	Tyr	Phe	Ser	Leu	Leu	Phe	Leu	Leu	Ser	Arg	Ser
			340					345					350		
Thr	Ala	Val	Leu	Leu	Phe	Val	Ser	Ala	Ile	Asn	Asp	Gln	Ala	Arg	Glu
		355					360					365			
Pro	Leu	Arg	Leu	Leu	Arg	Leu	Val	Pro	Leu	Lys	Gly	Tyr	His	Pro	Glu
		370				375					380				
Val	Phe	Arg	Phe	Ala	Ala	Glu	Leu	Ala	Ser	Asp	Gln	Val	Ala	Leu	Thr
385					390					395					400
Gly	Leu	Lys	Phe	Phe	Asn	Val	Thr	Arg	Lys	Leu	Phe	Leu	Ala	Met	Ala
			405						410					415	
Gly	Thr	Val	Ala	Thr	Tyr	Glu	Leu	Val	Leu	Ile	Gln	Phe	His	Glu	Asp
			420					425					430		
Lys	Lys	Thr	Trp	Asp	Cys	Ser	Pro	Phe	Asn	Leu	Asp				
		435					440								

<210> 5

<211> 1158

<212> DNA

<213> *Drosophila melanogaster*

<220>

<221> CDS

<222> (1)..(1155)

<223> Coding region 8D.1

<400> 5

```

atg agc ggc cat ctg ggt cgg gtc ctg cag ttc cac ctg cgg ctc tac 48
Met Ser Gly His Leu Gly Arg Val Leu Gln Phe His Leu Arg Leu Tyr
  1           5           10           15

cag gtg ctc ggc ttc cat ggg ctg ccg ttg ccg ggc gat ggg aat ccg 96
Gln Val Leu Gly Phe His Gly Leu Pro Leu Pro Gly Asp Gly Asn Pro
          20           25           30

gcc agg acc agg agg cgt ctg atg gca tgg agc ctg ttc ctg ctc att 144
Ala Arg Thr Arg Arg Arg Leu Met Ala Trp Ser Leu Phe Leu Leu Ile
          35           40           45

tcg ctg agt gcc ctc gta ctc gcg tgc ctc ttt agc ggc gag gag ttc 192
Ser Leu Ser Ala Leu Val Leu Ala Cys Leu Phe Ser Gly Glu Glu Phe
          50           55           60

ctc tat cgc ggc gac atg ttc ggc tgt gcc aat gat gcc ctt aaa tac 240
Leu Tyr Arg Gly Asp Met Phe Gly Cys Ala Asn Asp Ala Leu Lys Tyr
          65           70           75           80

gta ttc gcc gaa ttg ggc gtg ctg gcc ata tat ctg gag acg ctg agc 288
Val Phe Ala Glu Leu Gly Val Leu Ala Ile Tyr Leu Glu Thr Leu Ser
          85           90           95

agc cag cgg cat ttg gcc aac ttc tgg tgg ctg cac ttc aag ttg ggc 336
Ser Gln Arg His Leu Ala Asn Phe Trp Trp Leu His Phe Lys Leu Gly
          100          105          110

ggc caa aaa acg ggc ttg gtg agc ctg cgc agt gag ttc cag cag ttt 384
Gly Gln Lys Thr Gly Leu Val Ser Leu Arg Ser Glu Phe Gln Gln Phe
          115          120          125

tgt cgc tat ctg ata ttc ctg tac gcc atg atg gcc gcc gaa gtg gcg 432
Cys Arg Tyr Leu Ile Phe Leu Tyr Ala Met Met Ala Ala Glu Val Ala
          130          135          140

atc cat ttg gga ttg tgg cag ttc caa gcg ctc acc caa cat atg ttg 480
Ile His Leu Gly Leu Trp Gln Phe Gln Ala Leu Thr Gln His Met Leu
          145          150          155          160

ctc ttt tgg agc acc tat gag ccg ctc gtg tgg ctg acg tat ctg cgc 528
Leu Phe Trp Ser Thr Tyr Glu Pro Leu Val Trp Leu Thr Tyr Leu Arg
          165          170          175

aat ctg cag ttc gta ctg cac ttg gag ctg ctc agg gag cag ctg acc 576
Asn Leu Gln Phe Val Leu His Leu Glu Leu Leu Arg Glu Gln Leu Thr
          180          185          190

ggc ttg gaa cgc gaa atg ggt ctg ctg gcg gag tac tcg cga ttt gct 624
Gly Leu Glu Arg Glu Met Gly Leu Leu Ala Glu Tyr Ser Arg Phe Ala
          195          200          205

```

```

agc gaa acg ggt cgg agt ttt cct gga ttc gaa agt ttc ctg cgc cga 672
Ser Glu Thr Gly Arg Ser Phe Pro Gly Phe Glu Ser Phe Leu Arg Arg
210 215 220

cga cta gtg cag aag cag cgc atc tat agc cat gtg tat gac atg ctc 720
Arg Leu Val Gln Lys Gln Arg Ile Tyr Ser His Val Tyr Asp Met Leu
225 230 235 240

aaa tgt ttc cag ggt gcc ttc aac ttc tcc att ctc gcc gtc ctg ctg 768
Lys Cys Phe Gln Gly Ala Phe Asn Phe Ser Ile Leu Ala Val Leu Leu
245 250 255

acc atc aac ata cgc atc gcc gtg gac tgc tac ttc atg tac tac agc 816
Thr Ile Asn Ile Arg Ile Ala Val Asp Cys Tyr Phe Met Tyr Tyr Ser
260 265 270

atc tac aac aat gtg att aac aac gat tac tac cta atc gtt ccc gcc 864
Ile Tyr Asn Asn Val Ile Asn Asn Asp Tyr Tyr Leu Ile Val Pro Ala
275 280 285

ttg ctc gag att ccc gcc ttc ata tac gct tgc cag agc tgc atg gtc 912
Leu Leu Glu Ile Pro Ala Phe Ile Tyr Ala Ser Gln Ser Cys Met Val
290 295 300

gtg gtg ccc agg atc gcc cac cag ctg cat aat ata gtc acc gat tcc 960
Val Val Pro Arg Ile Ala His Gln Leu His Asn Ile Val Thr Asp Ser
305 310 315 320

ggt tgc tgc agc tgt ccc gat ctc tcc ctg cag att cag aac ttt tca 1008
Gly Cys Cys Ser Cys Pro Asp Leu Ser Leu Gln Ile Gln Asn Phe Ser
325 330 335

ctg caa ctc ctg cat cag ccg ata cga atc gat tgc ctc ggc ctg acc 1056
Leu Gln Leu Leu His Gln Pro Ile Arg Ile Asp Cys Leu Gly Leu Thr
340 345 350

atc ctg gat tgc agt ctt cta act cgg atg gcc tgt tcc gtg ggc acc 1104
Ile Leu Asp Cys Ser Leu Leu Thr Arg Met Ala Cys Ser Val Gly Thr
355 360 365

tac atg atc tat agc atc cag ttt ata cca aag ttc agc aat acc tat 1152
Tyr Met Ile Tyr Ser Ile Gln Phe Ile Pro Lys Phe Ser Asn Thr Tyr
370 375 380

atg tag 1158
Met
385

```

```

<210> 6
<211> 385
<212> PRT
<213> Drosophila melanogaster

```

```

<400> 6
Met Ser Gly His Leu Gly Arg Val Leu Gln Phe His Leu Arg Leu Tyr

```

1	5	10	15
Gln Val Leu Gly Phe His Gly Leu Pro Leu Pro Gly Asp Gly Asn Pro	20	25	30
Ala Arg Thr Arg Arg Arg Leu Met Ala Trp Ser Leu Phe Leu Leu Ile	35	40	45
Ser Leu Ser Ala Leu Val Leu Ala Cys Leu Phe Ser Gly Glu Glu Phe	50	55	60
Leu Tyr Arg Gly Asp Met Phe Gly Cys Ala Asn Asp Ala Leu Lys Tyr	65	70	75
Val Phe Ala Glu Leu Gly Val Leu Ala Ile Tyr Leu Glu Thr Leu Ser	85	90	95
Ser Gln Arg His Leu Ala Asn Phe Trp Trp Leu His Phe Lys Leu Gly	100	105	110
Gly Gln Lys Thr Gly Leu Val Ser Leu Arg Ser Glu Phe Gln Gln Phe	115	120	125
Cys Arg Tyr Leu Ile Phe Leu Tyr Ala Met Met Ala Ala Glu Val Ala	130	135	140
Ile His Leu Gly Leu Trp Gln Phe Gln Ala Leu Thr Gln His Met Leu	145	150	155
Leu Phe Trp Ser Thr Tyr Glu Pro Leu Val Trp Leu Thr Tyr Leu Arg	165	170	175
Asn Leu Gln Phe Val Leu His Leu Glu Leu Leu Arg Glu Gln Leu Thr	180	185	190
Gly Leu Glu Arg Glu Met Gly Leu Leu Ala Glu Tyr Ser Arg Phe Ala	195	200	205
Ser Glu Thr Gly Arg Ser Phe Pro Gly Phe Glu Ser Phe Leu Arg Arg	210	215	220
Arg Leu Val Gln Lys Gln Arg Ile Tyr Ser His Val Tyr Asp Met Leu	225	230	235
Lys Cys Phe Gln Gly Ala Phe Asn Phe Ser Ile Leu Ala Val Leu Leu	245	250	255
Thr Ile Asn Ile Arg Ile Ala Val Asp Cys Tyr Phe Met Tyr Tyr Ser	260	265	270
Ile Tyr Asn Asn Val Ile Asn Asn Asp Tyr Tyr Leu Ile Val Pro Ala	275	280	285
Leu Leu Glu Ile Pro Ala Phe Ile Tyr Ala Ser Gln Ser Cys Met Val	290	295	300

Val Val Pro Arg Ile Ala His Gln Leu His Asn Ile Val Thr Asp Ser
 305 310 315 320
 Gly Cys Cys Ser Cys Pro Asp Leu Ser Leu Gln Ile Gln Asn Phe Ser
 325 330 335
 Leu Gln Leu Leu His Gln Pro Ile Arg Ile Asp Cys Leu Gly Leu Thr
 340 345 350
 Ile Leu Asp Cys Ser Leu Leu Thr Arg Met Ala Cys Ser Val Gly Thr
 355 360 365
 Tyr Met Ile Tyr Ser Ile Gln Phe Ile Pro Lys Phe Ser Asn Thr Tyr
 370 375 380
 Met
 385

<210> 7
 <211> 1119
 <212> DNA
 <213> Drosophila melanogaster

<220>
 <221> CDS
 <222> (1)..(1119)
 <223> Coding region GR10B.1

<400> 7
 atg cga gtg ggc aag ttg tgc cgc ctg gct ttg cgc ttc tgg atg gga 48
 Met Arg Val Gly Lys Leu Cys Arg Leu Ala Leu Arg Phe Trp Met Gly
 1 5 10 15
 ttg att ctg gtc ctg ggt ttc tcc agc cac tac tac aat ccc act cgg 96
 Leu Ile Leu Val Leu Gly Phe Ser Ser His Tyr Tyr Asn Pro Thr Arg
 20 25 30
 cgg cgt ttg gtt tac tgc aga att ttg cag aca tac gac tgg ctg tta 144
 Arg Arg Leu Val Tyr Ser Arg Ile Leu Gln Thr Tyr Asp Trp Leu Leu
 35 40 45
 atg gtc atc aac ttg ggt gct ttt tac ctc tac tat aga tac gcc atg 192
 Met Val Ile Asn Leu Gly Ala Phe Tyr Leu Tyr Tyr Arg Tyr Ala Met
 50 55 60
 acg tat ttt ctg gag ggc atg ttt agg cgg caa gga ttc gta aat cag 240
 Thr Tyr Phe Leu Glu Gly Met Phe Arg Arg Gln Gly Phe Val Asn Gln
 65 70 75 80
 gtg tcc aca tgc aac gta ttc cag cag cta ctg atg gca gtg acg ggc 288
 Val Ser Thr Cys Asn Val Phe Gln Gln Leu Leu Met Ala Val Thr Gly
 85 90 95
 aca tgg ttg cac ttt cta ttt gaa cgg cac gtg tgt cag acg tac aac 336

Thr	Trp	Leu	His	Phe	Leu	Phe	Glu	Arg	His	Val	Cys	Gln	Thr	Tyr	Asn	
			100					105					110			
gag	ttg	tcc	aga	att	ttg	aag	cat	gac	ctg	aag	ctc	aag	gag	cac	agt	384
Glu	Leu	Ser	Arg	Ile	Leu	Lys	His	Asp	Leu	Lys	Leu	Lys	Glu	His	Ser	
		115					120				125					
cgc	ttc	tac	tgc	ctg	gcc	ttt	ctg	gcc	aaa	gtc	tac	aat	ttt	ttc	cac	432
Arg	Phe	Tyr	Cys	Leu	Ala	Phe	Leu	Ala	Lys	Val	Tyr	Asn	Phe	Phe	His	
		130				135				140						
aac	ttc	aac	ttt	gcc	cta	agt	gcc	atc	atg	cat	tgg	ggc	ctg	cga	ccc	480
Asn	Phe	Asn	Phe	Ala	Leu	Ser	Ala	Ile	Met	His	Trp	Gly	Leu	Arg	Pro	
145					150				155						160	
ttc	aat	gtt	tgg	gac	tta	ctg	gcc	aat	ctg	tac	ttt	gtc	tac	aac	tca	528
Phe	Asn	Val	Trp	Asp	Leu	Leu	Ala	Asn	Leu	Tyr	Phe	Val	Tyr	Asn	Ser	
			165					170						175		
ctg	gct	agg	gat	gcc	att	ctg	gtg	gcc	tat	gtt	ctg	ttg	ctg	ctc	aac	576
Leu	Ala	Arg	Asp	Ala	Ile	Leu	Val	Ala	Tyr	Val	Leu	Leu	Leu	Leu	Asn	
			180					185						190		
ctg	tcg	gag	gca	ctg	cgt	ctg	aat	ggc	cag	cag	gag	cac	gac	acc	tac	624
Leu	Ser	Glu	Ala	Leu	Arg	Leu	Asn	Gly	Gln	Gln	Glu	His	Asp	Thr	Tyr	
		195					200						205			
agc	gac	ctc	atg	aag	cag	ctg	cgc	aga	cgg	gag	cga	ttg	ctg	agg	att	672
Ser	Asp	Leu	Met	Lys	Gln	Leu	Arg	Arg	Arg	Glu	Arg	Leu	Leu	Arg	Ile	
		210				215					220					
gga	agg	cga	gtg	cat	cgc	atg	ttc	gcc	tgg	ctg	gtg	gcc	ata	gcc	tta	720
Gly	Arg	Arg	Val	His	Arg	Met	Phe	Ala	Trp	Leu	Val	Ala	Ile	Ala	Leu	
225					230					235					240	
ata	tat	ctg	gtg	ttt	ttc	aac	acg	gcc	acc	att	tac	ctg	ggc	tac	acc	768
Ile	Tyr	Leu	Val	Phe	Phe	Asn	Thr	Ala	Thr	Ile	Tyr	Leu	Gly	Tyr	Thr	
			245					250						255		
atg	ttc	att	cag	aag	cac	gac	gcc	ttg	ggc	ctg	cgt	ggc	cga	ggc	ttg	816
Met	Phe	Ile	Gln	Lys	His	Asp	Ala	Leu	Gly	Leu	Arg	Gly	Arg	Gly	Leu	
			260					265					270			
aag	atg	ctg	ctc	acg	gtc	gtc	tcc	ttt	ttg	gtc	atc	cta	tgg	gac	gtc	864
Lys	Met	Leu	Leu	Thr	Val	Val	Ser	Phe	Leu	Val	Ile	Leu	Trp	Asp	Val	
		275					280						285			
gtg	cta	ctg	cag	gtc	atc	tgt	gaa	aag	ctt	ttg	gcc	gag	gaa	aac	aaa	912
Val	Leu	Leu	Gln	Val	Ile	Cys	Glu	Lys	Leu	Leu	Ala	Glu	Glu	Asn	Lys	
		290				295					300					
att	tgc	gac	tgt	cca	gag	gat	gtg	gcc	tct	tcc	agg	acg	act	tac	aga	960
Ile	Cys	Asp	Cys	Pro	Glu	Asp	Val	Ala	Ser	Ser	Arg	Thr	Thr	Tyr	Arg	
305					310					315				320		
cag	tgg	gaa	atg	tcc	gct	tta	aga	cga	gct	att	aca	cga	tca	tcg	ccg	1008

Gln Trp Glu Met Ser Ala Leu Arg Arg Ala Ile Thr Arg Ser Ser Pro
 325 330 335

gaa aac aat gtt ttg ggc atg ttt cga atg gat atg cga tgc gca ttc 1056
 Glu Asn Asn Val Leu Gly Met Phe Arg Met Asp Met Arg Cys Ala Phe
 340 345 350

gct ttg atc agc tgc agt ttg tcc tat ggc att ata atc att cag att 1104
 Ala Leu Ile Ser Cys Ser Leu Ser Tyr Gly Ile Ile Ile Ile Gln Ile
 355 360 365

gga tat ata cca ggc 1119
 Gly Tyr Ile Pro Gly
 370

<210> 8
 <211> 373
 <212> PRT
 <213> *Drosophila melanogaster*

<400> 8
 Met Arg Val Gly Lys Leu Cys Arg Leu Ala Leu Arg Phe Trp Met Gly
 1 5 10 15

Leu Ile Leu Val Leu Gly Phe Ser Ser His Tyr Tyr Asn Pro Thr Arg
 20 25 30

Arg Arg Leu Val Tyr Ser Arg Ile Leu Gln Thr Tyr Asp Trp Leu Leu
 35 40 45

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

Thr Tyr Phe Leu Glu Gly Met Phe Arg Arg Gln Gly Phe Val Asn Gln
 65 70 75 80

Val Ser Thr Cys Asn Val Phe Gln Gln Leu Leu Met Ala Val Thr Gly
 85 90 95

Thr Trp Leu His Phe Leu Phe Glu Arg His Val Cys Gln Thr Tyr Asn
 100 105 110

Glu Leu Ser Arg Ile Leu Lys His Asp Leu Lys Leu Lys Glu His Ser
 115 120 125

Arg Phe Tyr Cys Leu Ala Phe Leu Ala Lys Val Tyr Asn Phe Phe His
 130 135 140

Asn Phe Asn Phe Ala Leu Ser Ala Ile Met His Trp Gly Leu Arg Pro
 145 150 155 160

Phe Asn Val Trp Asp Leu Leu Ala Asn Leu Tyr Phe Val Tyr Asn Ser
 165 170 175

Leu Ala Arg Asp Ala Ile Leu Val Ala Tyr Val Leu Leu Leu Leu Asn

180 185 190
Leu Ser Glu Ala Leu Arg Leu Asn Gly Gln Gln Glu His Asp Thr Tyr
195 200 205
Ser Asp Leu Met Lys Gln Leu Arg Arg Arg Glu Arg Leu Leu Arg Ile
210 215 220
Gly Arg Arg Val His Arg Met Phe Ala Trp Leu Val Ala Ile Ala Leu
225 230 235 240
Ile Tyr Leu Val Phe Phe Asn Thr Ala Thr Ile Tyr Leu Gly Tyr Thr
245 250 255
Met Phe Ile Gln Lys His Asp Ala Leu Gly Leu Arg Gly Arg Gly Leu
260 265 270
Lys Met Leu Leu Thr Val Val Ser Phe Leu Val Ile Leu Trp Asp Val
275 280 285
Val Leu Leu Gln Val Ile Cys Glu Lys Leu Leu Ala Glu Glu Asn Lys
290 295 300
Ile Cys Asp Cys Pro Glu Asp Val Ala Ser Ser Arg Thr Thr Tyr Arg
305 310 315 320
Gln Trp Glu Met Ser Ala Leu Arg Arg Ala Ile Thr Arg Ser Ser Pro
325 330 335
Glu Asn Asn Val Leu Gly Met Phe Arg Met Asp Met Arg Cys Ala Phe
340 345 350
Ala Leu Ile Ser Cys Ser Leu Ser Tyr Gly Ile Ile Ile Ile Gln Ile
355 360 365
Gly Tyr Ile Pro Gly
370

<210> 9
<211> 1092
<212> DNA
<213> Drosophila melanogaster

<220>
<221> CDS
<222> (1)..(1092)
<223> Coding region GR21D.1

<220>
<221> misc_feature
<222> (505)..(766)
<223> Splice sites before bases 505, 562 and 766.

<400> 9

```

atg ggc gta atg ccc att cac cgc aat ccg ccc gag aag aac ctg cct 48
Met Gly Val Met Pro Ile His Arg Asn Pro Pro Glu Lys Asn Leu Pro
1 5 10 15

aga acg ggc tac tcc tgg ggc tcc aag cag gtc atg tgg gcc atc ttc 96
Arg Thr Gly Tyr Ser Trp Gly Ser Lys Gln Val Met Trp Ala Ile Phe
20 25 30

atc tac agc tgc cag acg acc att gtt gtt ctg gtg ctg cgt gag cgc 144
Ile Tyr Ser Cys Gln Thr Thr Ile Val Val Leu Val Leu Arg Glu Arg
35 40 45

gtg aaa aaa ttc gtc acc agc ccg gac aag cgc ttc gat gag gcc atc 192
Val Lys Lys Phe Val Thr Ser Pro Asp Lys Arg Phe Asp Glu Ala Ile
50 55 60

tac aat gtc att ttc atc agt ctg ctc ttc acc aac ttt ctg ctc ccg 240
Tyr Asn Val Ile Phe Ile Ser Leu Leu Phe Thr Asn Phe Leu Leu Pro
65 70 75 80

gtg gcc agc tgg cgg cac ggt ccc cag gtg gct atc ttc aag aac atg 288
Val Ala Ser Trp Arg His Gly Pro Gln Val Ala Ile Phe Lys Asn Met
85 90 95

tgg acc aac tac cag tat aag ttc ttc aaa act acc gcc tcg ccg atc 336
Trp Thr Asn Tyr Gln Tyr Lys Phe Phe Lys Thr Thr Gly Ser Pro Ile
100 105 110

gtc ttt ccg aat ctc tac cca ctc acc tgg tcg ctg tgc gtc ttc tcc 384
Val Phe Pro Asn Leu Tyr Pro Leu Thr Trp Ser Leu Cys Val Phe Ser
115 120 125

tgg ctc cta agc atc gcc atc aac ctg tcg cag tac ttc ctg cag ccg 432
Trp Leu Leu Ser Ile Ala Ile Asn Leu Ser Gln Tyr Phe Leu Gln Pro
130 135 140

gac ttc ccg ctg tgg tac aca ttc gcc tac tac ccc atc atc gcc atg 480
Asp Phe Arg Leu Trp Tyr Thr Phe Ala Tyr Tyr Pro Ile Ile Ala Met
145 150 155 160

ctc aac tgc ttc tgc agc ttg tgg tac atc aac tgc aat gca ttc gga 528
Leu Asn Cys Phe Cys Ser Leu Trp Tyr Ile Asn Cys Asn Ala Phe Gly
165 170 175

act gca agt cgc gct ctt tcc gac gct ctg cag aca acc atc cgg gcc 576
Thr Ala Ser Arg Ala Leu Ser Asp Ala Leu Gln Thr Thr Ile Arg Gly
180 185 190

gag aag ccc gcc caa aaa ctc act gag tac cgc cat cta tgg gtg gat 624
Glu Lys Pro Ala Gln Lys Leu Thr Glu Tyr Arg His Leu Trp Val Asp
195 200 205

ctg agc cac atg atg cag cag ctg gga cga gct tac tcc aac atg tac 672
Leu Ser His Met Met Gln Gln Leu Gly Arg Ala Tyr Ser Asn Met Tyr
210 215 220

```

WO 00/77208

PCT/US00/16211

```

ggc atg tac tgt cta gtc atc ttc ttc aca acg atc atc gcc act tac 720
Gly Met Tyr Cys Leu Val Ile Phe Phe Thr Thr Ile Ile Ala Thr Tyr
225                230                235                240

ggc agc atc agc gaa atc atc gac cac ggg gca acc tac aag gag gtg 768
Gly Ser Ile Ser Glu Ile Ile Asp His Gly Ala Thr Tyr Lys Glu Val
                245                250                255

ggg ctg ttc gtc atc gtg ttc tac tgc atg ggt ctg ctg tac att atc 816
Gly Leu Phe Val Ile Val Phe Tyr Cys Met Gly Leu Leu Tyr Ile Ile
                260                265                270

tgc aac gag gcg cac tac gcc tcc cga aag gtc gga ctt gac ttc cag 864
Cys Asn Glu Ala His Tyr Ala Ser Arg Lys Val Gly Leu Asp Phe Gln
                275                280                285

acg aag ctg ctc aat atc aac ctg act gcc gtg gac gct gcc acc cag 912
Thr Lys Leu Leu Asn Ile Asn Leu Thr Ala Val Asp Ala Ala Thr Gln
                290                295                300

aag gaa gtg gag atg ctg ctt gtg gcc ata aac aag aac cag ccc atc 960
Lys Glu Val Glu Met Leu Leu Val Ala Ile Asn Lys Asn Pro Pro Ile
305                310                315                320

atg aat ctg gac ggg tat gca aac atc aac cgc gag ctg atc acg acc 1008
Met Asn Leu Asp Gly Tyr Ala Asn Ile Asn Arg Glu Leu Ile Thr Thr
                325                330                335

aac atc tgc ttc atg gcc acc tac ctc gtg gtc ctg ctg cag ttt aag 1056
Asn Ile Ser Phe Met Ala Thr Tyr Leu Val Val Leu Leu Gln Phe Lys
                340                345                350

atc acg gag cag aga cgc att ggt cag cag cag gcc 1092
Ile Thr Glu Gln Arg Arg Ile Gly Gln Gln Gln Ala
                355                360

```

<210> 10
 <211> 364
 <212> PRT
 <213> Drosophila melanogaster

```

<400> 10
Met Gly Val Met Pro Ile His Arg Asn Pro Pro Glu Lys Asn Leu Pro
  1                5                10                15

Arg Thr Gly Tyr Ser Trp Gly Ser Lys Gln Val Met Trp Ala Ile Phe
                20                25                30

Ile Tyr Ser Cys Gln Thr Thr Ile Val Val Leu Val Leu Arg Glu Arg
                35                40                45

Val Lys Lys Phe Val Thr Ser Pro Asp Lys Arg Phe Asp Glu Ala Ile
                50                55                60

Tyr Asn Val Ile Phe Ile Ser Leu Leu Phe Thr Asn Phe Leu Leu Pro

```

65	70	75	80
Val Ala Ser Trp Arg His Gly Pro Gln Val Ala Ile Phe Lys Asn Met			
85		90	95
Trp Thr Asn Tyr Gln Tyr Lys Phe Phe Lys Thr Thr Gly Ser Pro Ile			
100	105	110	
Val Phe Pro Asn Leu Tyr Pro Leu Thr Trp Ser Leu Cys Val Phe Ser			
115	120	125	
Trp Leu Leu Ser Ile Ala Ile Asn Leu Ser Gln Tyr Phe Leu Gln Pro			
130	135	140	
Asp Phe Arg Leu Trp Tyr Thr Phe Ala Tyr Tyr Pro Ile Ile Ala Met			
145	150	155	160
Leu Asn Cys Phe Cys Ser Leu Trp Tyr Ile Asn Cys Asn Ala Phe Gly			
165	170	175	
Thr Ala Ser Arg Ala Leu Ser Asp Ala Leu Gln Thr Thr Ile Arg Gly			
180	185	190	
Glu Lys Pro Ala Gln Lys Leu Thr Glu Tyr Arg His Leu Trp Val Asp			
195	200	205	
Leu Ser His Met Met Gln Gln Leu Gly Arg Ala Tyr Ser Asn Met Tyr			
210	215	220	
Gly Met Tyr Cys Leu Val Ile Phe Phe Thr Thr Ile Ile Ala Thr Tyr			
225	230	235	240
Gly Ser Ile Ser Glu Ile Ile Asp His Gly Ala Thr Tyr Lys Glu Val			
245	250	255	
Gly Leu Phe Val Ile Val Phe Tyr Cys Met Gly Leu Leu Tyr Ile Ile			
260	265	270	
Cys Asn Glu Ala His Tyr Ala Ser Arg Lys Val Gly Leu Asp Phe Gln			
275	280	285	
Thr Lys Leu Leu Asn Ile Asn Leu Thr Ala Val Asp Ala Ala Thr Gln			
290	295	300	
Lys Glu Val Glu Met Leu Leu Val Ala Ile Asn Lys Asn Pro Pro Ile			
305	310	315	320
Met Asn Leu Asp Gly Tyr Ala Asn Ile Asn Arg Glu Leu Ile Thr Thr			
325	330	335	
Asn Ile Ser Phe Met Ala Thr Tyr Leu Val Val Leu Leu Gln Phe Lys			
340	345	350	
Ile Thr Glu Gln Arg Arg Ile Gly Gln Gln Gln Ala			
355	360		

<210> 11
 <211> 1134
 <212> DNA
 <213> *Drosophila melanogaster*

<220>
 <221> CDS
 <222> (1)..(1134)
 <223> Coding region GR22B.1

<220>
 <221> misc_feature
 <222> (985)
 <223> Splice site

```

<400> 11
atg aaa atg ttc caa cct cgt cgc ggc ttc agt tgc cat ttg gca tgg 48
Met Lys Met Phe Gln Pro Arg Arg Gly Phe Ser Cys His Leu Ala Trp
  1             5             10             15

ttc atg ctg cag acg acc ctc tac gca tcc tgg cta ctg gga ctg ttc 96
Phe Met Leu Gln Thr Thr Leu Tyr Ala Ser Trp Leu Leu Gly Leu Phe
      20             25             30

ccc ttc act ttc gat tcc cgg aga aag cag ttg aaa cgc tcc cga tgg 144
Pro Phe Thr Phe Asp Ser Arg Arg Lys Gln Leu Lys Arg Ser Arg Trp
      35             40             45

ctt ctc ctc tac ggt ttc gtt ctg cat tct ctg gcc atg tgc cta gcc 192
Leu Leu Leu Tyr Gly Phe Val Leu His Ser Leu Ala Met Cys Leu Ala
      50             55             60

atg agt agc cac ttg gca tcc aag cag cga agg aaa tat aat gca ttt 240
Met Ser Ser His Leu Ala Ser Lys Gln Arg Arg Lys Tyr Asn Ala Phe
      65             70             75             80

gag cgc aat ccg ctg ttg gag aaa ata tat atg caa ttt cag gtt act 288
Glu Arg Asn Pro Leu Leu Glu Lys Ile Tyr Met Gln Phe Gln Val Thr
      85             90             95

aca ttc ttc act ata agc gta cta ctg ctc atg aat gtt tgg aaa agc 336
Thr Phe Phe Thr Ile Ser Val Leu Leu Leu Met Asn Val Trp Lys Ser
      100            105            110

aac acg gtt cgg aaa att gcc aat gaa cta ctc acc ctt gag ggc cag 384
Asn Thr Val Arg Lys Ile Ala Asn Glu Leu Leu Thr Leu Glu Gly Gln
      115            120            125

gta aaa gac ctt ctc act ctg aag aac tgt ccc aac ttt aac tgt ttc 432
Val Lys Asp Leu Leu Thr Leu Lys Asn Cys Pro Asn Phe Asn Cys Phe
      130            135            140

gtg att aaa aaa cat gtg gct gca ata ggt cag ttc gtg att tca att 480
Val Ile Lys Lys His Val Ala Ala Ile Gly Gln Phe Val Ile Ser Ile
      145            150            155            160

```

tac ttc tgc cta tgc caa gaa aat tgc tac ccg aag atc ctt aag atc	528
Tyr Phe Cys Leu Cys Gln Glu Asn Ser Tyr Pro Lys Ile Leu Lys Ile	
165 170 175	
ctt tgc tgc ctg cca tgc gtg ggc ttg cag ctg att ata atg cac ttt	576
Leu Cys Cys Leu Pro Ser Val Gly Leu Gln Leu Ile Ile Met His Phe	
180 185 190	
cat acc gaa atc att ttg gtc tac cgc tat gtg tgg ctg gtc aat gaa	624
His Thr Glu Ile Ile Leu Val Tyr Arg Tyr Val Trp Leu Val Asn Glu	
195 200 205	
acc ctc gaa gat tca cac cat cta agc tcc tca aga att cac gca cta	672
Thr Leu Glu Asp Ser His His Leu Ser Ser Ser Arg Ile His Ala Leu	
210 215 220	
gca tcc ttg tac gat cgc ctg ctg aag ctg agt gaa ttg gtg gta gcg	720
Ala Ser Leu Tyr Asp Arg Leu Leu Lys Leu Ser Glu Leu Val Val Ala	
225 230 235 240	
tgc aac gat ttg caa ctg atc ctt atg cta att att tat tta atc gga	768
Cys Asn Asp Leu Gln Leu Ile Leu Met Leu Ile Ile Tyr Leu Ile Gly	
245 250 255	
aac act gtg caa atc ttt ttc ctc att gtg ctc gga gta agc atg aat	816
Asn Thr Val Gln Ile Phe Phe Leu Ile Val Leu Gly Val Ser Met Asn	
260 265 270	
aag cgg tat att tac tta gtg gct tcc cca caa ttg atc atc aat ttc	864
Lys Arg Tyr Ile Tyr Leu Val Ala Ser Pro Gln Leu Ile Ile Asn Phe	
275 280 285	
tgg gac ttc tgg ctg aat atc gtg gtg tgt gat ctt gcc gga aag tgt	912
Trp Asp Phe Trp Leu Asn Ile Val Val Cys Asp Leu Ala Gly Lys Cys	
290 295 300	
gga gac caa aca tca aag gtc ctg aaa ctg ttc act gat ctg gag cac	960
Gly Asp Gln Thr Ser Lys Val Leu Lys Leu Phe Thr Asp Leu Glu His	
305 310 315 320	
gac gat gaa gag tta gag aga agt ttg aat gaa ttc gca tgg ctg tgc	1008
Asp Asp Glu Glu Leu Glu Arg Ser Leu Asn Glu Phe Ala Trp Leu Cys	
325 330 335	
acc cac cgc aag ttt cgg ttt cag ctg tgt ggt ctc ttt tcc att aac	1056
Thr His Arg Lys Phe Arg Phe Gln Leu Cys Gly Leu Phe Ser Ile Asn	
340 345 350	
cac aac atg ggc ttt caa atg atc atc acc agt ttt cta tac ctg gtt	1104
His Asn Met Gly Phe Gln Met Ile Ile Thr Ser Phe Leu Tyr Leu Val	
355 360 365	
tac ctg ctt cag ttt gac ttc atg aac ttg	1134
Tyr Leu Leu Gln Phe Asp Phe Met Asn Leu	
370 375	

<210> 12

<211> 378

<212> PRT

<213> Drosophila melanogaster

<400> 12

```

Met Lys Met Phe Gln Pro Arg Arg Gly Phe Ser Cys His Leu Ala Trp
  1           5           10           15

Phe Met Leu Gln Thr Thr Leu Tyr Ala Ser Trp Leu Leu Gly Leu Phe
  20           25           30

Pro Phe Thr Phe Asp Ser Arg Arg Lys Gln Leu Lys Arg Ser Arg Trp
  35           40           45

Leu Leu Leu Tyr Gly Phe Val Leu His Ser Leu Ala Met Cys Leu Ala
  50           55           60

Met Ser Ser His Leu Ala Ser Lys Gln Arg Arg Lys Tyr Asn Ala Phe
  65           70           75           80

Glu Arg Asn Pro Leu Leu Glu Lys Ile Tyr Met Gln Phe Gln Val Thr
  85           90           95

Thr Phe Phe Thr Ile Ser Val Leu Leu Met Asn Val Trp Lys Ser
  100          105          110

Asn Thr Val Arg Lys Ile Ala Asn Glu Leu Leu Thr Leu Glu Gly Gln
  115          120          125

Val Lys Asp Leu Leu Thr Leu Lys Asn Cys Pro Asn Phe Asn Cys Phe
  130          135          140

Val Ile Lys Lys His Val Ala Ala Ile Gly Gln Phe Val Ile Ser Ile
  145          150          155          160

Tyr Phe Cys Leu Cys Gln Glu Asn Ser Tyr Pro Lys Ile Leu Lys Ile
  165          170          175

Leu Cys Cys Leu Pro Ser Val Gly Leu Gln Leu Ile Ile Met His Phe
  180          185          190

His Thr Glu Ile Ile Leu Val Tyr Arg Tyr Val Trp Leu Val Asn Glu
  195          200          205

Thr Leu Glu Asp Ser His His Leu Ser Ser Ser Arg Ile His Ala Leu
  210          215          220

Ala Ser Leu Tyr Asp Arg Leu Leu Lys Leu Ser Glu Leu Val Val Ala
  225          230          235          240

Cys Asn Asp Leu Gln Leu Ile Leu Met Leu Ile Ile Tyr Leu Ile Gly
  245          250          255

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

```

Lys Arg Tyr Ile Tyr Leu Val Ala Ser Pro Gln Leu Ile Ile Asn Phe
 275 280 285

Trp Asp Phe Trp Leu Asn Ile Val Val Cys Asp Leu Ala Gly Lys Cys
 290 295 300

Gly Asp Gln Thr Ser Lys Val Leu Lys Leu Phe Thr Asp Leu Glu His
 305 310 315 320

Asp Asp Glu Glu Leu Glu Arg Ser Leu Asn Glu Phe Ala Trp Leu Cys
 325 330 335

Thr His Arg Lys Phe Arg Phe Gln Leu Cys Gly Leu Phe Ser Ile Asn
 340 345 350

His Asn Met Gly Phe Gln Met Ile Ile Thr Ser Phe Leu Tyr Leu Val
 355 360 365

Tyr Leu Leu Gln Phe Asp Phe Met Asn Leu
 370 375

<210> 13
 <211> 1149
 <212> DNA
 <213> Drosophila melanogaster

<220>
 <221> CDS
 <222> (1)..(1149)
 <223> Coding region GR23A.1a

<220>
 <221> misc_feature
 <222> (910)..(1068)
 <223> Splice sites before positions 910 and 1068

<400> 13
 atg gta atg gtt caa tca gtt tca gtt tat aaa gcc aac atg aag act 48
 Met Val Met Val Gln Ser Val Ser Val Tyr Lys Ala Asn Met Lys Thr
 1 5 10 15

ttg gag tgc ctg acc cgc cgt ttc ctg gaa gtt atc ttc tct gta ctg 96
 Leu Glu Cys Leu Thr Arg Arg Phe Leu Glu Val Ile Phe Ser Val Leu
 20 25 30

gcc ttg gtt cca ctt cct cct att tcg cag ttg gga tgg ctg ttt tta 144
 Ala Leu Val Pro Leu Pro Pro Ile Ser Gln Leu Gly Trp Leu Phe Leu
 35 40 45

tca ttg gcc att cgc tgc tgt tgg ata gtt tac ttt ata tat ttg ctg 192
 Ser Leu Ala Ile Arg Cys Cys Trp Ile Val Tyr Phe Ile Tyr Leu Leu
 50 55 60

gac gtt gcc atc agc ttt tcg tgg gtg gcc atc gaa aat gtt gga aat 240

PCT/US00/16211

23

Pro Ile Ala Thr Leu Phe Trp His Gly Asp Ser Gly Ala Glu Asn Gly
 290 295 300

cgg caa atc ggt tgt cta att tcg aag ttg gta aaa ccc cag ggc agc 960
 Arg Gln Ile Gly Cys Leu Ile Ser Lys Leu Val Lys Pro Gln Gly Ser
 305 310 315 320

aaa cta tac aat gat ttg gta agt gaa ttt tca ctt caa aca ctg cac 1008
 Lys Leu Tyr Asn Asp Leu Val Ser Glu Phe Ser Leu Gln Thr Leu His
 325 330 335

cag cga ttt gtg gtg acc gcc aag gat ttc ttc agt ctc aat ctg cat 1056
 Gln Arg Phe Val Val Thr Ala Lys Asp Phe Phe Ser Leu Asn Leu His
 340 345 350

cta ctg agt agt atg ttt gca gct gtg gtc aca tat ctg gtt att ctc 1104
 Leu Leu Ser Ser Met Phe Ala Ala Val Val Thr Tyr Leu Val Ile Leu
 355 360 365

ata cag ttc atg ttc gcc gaa agg agt tca acg cga ggc agc gga 1149
 Ile Gln Phe Met Phe Ala Glu Arg Ser Ser Thr Arg Gly Ser Gly
 370 375 380

<210> 14

<211> 383

<212> PRT

<213> Drosophila melanogaster

<400> 14

Met Val Met Val Gln Ser Val Ser Val Tyr Lys Ala Asn Met Lys Thr
 1 5 10 15

Leu Glu Cys Leu Thr Arg Arg Phe Leu Glu Val Ile Phe Ser Val Leu
 20 25 30

Ala Leu Val Pro Leu Pro Pro Ile Ser Gln Leu Gly Trp Leu Phe Leu
 35 40 45

Ser Leu Ala Ile Arg Cys Cys Trp Ile Val Tyr Phe Ile Tyr Leu Leu
 50 55 60

Asp Val Ala Ile Ser Phe Ser Trp Val Ala Ile Glu Asn Val Gly Asn
 65 70 75 80

Ala Val Gly Thr Met Leu Phe Val Gly Asn Ser Val Leu Gly Phe Ala
 85 90 95

Leu Leu Leu Glu Ser Val Leu Lys Gln Lys Thr His Ser Gln Leu Glu
 100 105 110

Asp Leu Arg Val Gln Thr Glu Leu Gln Leu Gln Arg Leu Gly Met Phe
 115 120 125

Gly Arg Ser Arg His Ala Ala Tyr Leu Leu Pro Leu Ile Gly Val Gln
 130 135 140

WO 00/77208

PCT/US00/16211

```

Phe Thr Cys Asp Leu Val Arg Leu Ala Thr Asn Phe Gly Glu Thr Val
145                      150                      155                      160

Ser Pro Val Phe Cys Ile Ser Leu Pro Leu Met Trp Leu Leu Arg Tyr
                      165                      170                      175

Arg Tyr Val Gln Leu Val Gln His Val Met Asp Leu Asn Gln Arg Ser
                      180                      185                      190

Ile His Leu Arg Arg Ser Leu Leu Ser Met Ala Ser Gly Asn Asp Leu
                      195                      200                      205

Trp Gln Pro Tyr Gly Val Gln Glu Cys Leu Gln Leu Gln Thr Leu Arg
210                      215                      220

Thr Thr Tyr Glu Arg Ile Phe Glu Cys Tyr Glu Thr Phe Ser Asp Cys
225                      230                      235                      240

Tyr Gly Trp Gly Met Leu Gly Leu His Leu Leu Thr Ser Phe Gln Phe
                      245                      250                      255

Val Thr Asn Ala Tyr Trp Met Ile Met Gly Ile Tyr Asp Gly Gly Asn
                      260                      265                      270

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

Pro Ile Ala Thr Leu Phe Trp His Gly Asp Ser Gly Ala Glu Asn Gly
290                      295                      300

Arg Gln Ile Gly Cys Leu Ile Ser Lys Leu Val Lys Pro Gln Gly Ser
305                      310                      315                      320

Lys Leu Tyr Asn Asp Leu Val Ser Glu Phe Ser Leu Gln Thr Leu His
                      325                      330                      335

Gln Arg Phe Val Val Thr Ala Lys Asp Phe Phe Ser Leu Asn Leu His
                      340                      345                      350

Leu Leu Ser Ser Met Phe Ala Ala Val Val Thr Tyr Leu Val Ile Leu
355                      360                      365

Ile Gln Phe Met Phe Ala Glu Arg Ser Ser Thr Arg Gly Ser Gly
370                      375                      380

```

```

<210> 15
<211> 1122
<212> DNA
<213> Drosophila melanogaster

<220>
<221> CDS
<222> (1)..(1122)
<223> Coding region GR23A.1b

```

```

<220>
<221> misc_feature
<222> (901)..(1042)
<223> Splice sites before positions 901 and 1042

<400> 15
atg ttt ccg ccg aca aga gtg caa gct agc tca cga gtg gtc ctt aag 48
Met Phe Pro Pro Thr Arg Val Gln Ala Ser Ser Arg Val Val Leu Lys
  1             5             10             15

att ttc cat ttc att ttg gtg gca ttt tct ttg aga tca agg aga cta 96
Ile Phe His Phe Ile Leu Val Ala Phe Ser Leu Arg Ser Arg Arg Leu
      20             25             30

tct cgt ttg gtc ctg tgg tta cag ttt ctg ggt tgg cta acc tgg ttt 144
Ser Arg Leu Val Leu Trp Leu Gln Phe Leu Gly Trp Leu Thr Trp Phe
      35             40             45

att tcc atg tgg act caa agt gtg ata tat gcc caa acg att gac tgc 192
Ile Ser Met Trp Thr Gln Ser Val Ile Tyr Ala Gln Thr Ile Asp Cys
      50             55             60

aca ttg gac tgc tca ctg cgg cat ata ctc aca ttt ttc cag acg gtc 240
Thr Leu Asp Cys Ser Leu Arg His Ile Leu Thr Phe Phe Gln Thr Val
      65             70             75             80

tca cat gcc ttc att gtg gtt acc tcc ttt ctt gat ggt ttc aga atc 288
Ser His Ala Phe Ile Val Val Thr Ser Phe Leu Asp Gly Phe Arg Ile
      85             90             95

aag cag gac cag ttg gac gaa ccg att gcg ttc gag gac tcc gat ccc 336
Lys Gln Asp Gln Leu Asp Glu Pro Ile Ala Phe Glu Asp Ser Asp Pro
      100            105            110

tgg ttg gcc ttc act gtg ctg gcg atg ctg gtt ccc act ttg gga gta 384
Trp Leu Ala Phe Thr Val Leu Ala Met Leu Val Pro Thr Leu Gly Val
      115            120            125

gag tat ttg gtg tgc tgc aac gcg cca gaa tat gcc ttt cgt atc agg 432
Glu Tyr Leu Val Cys Ser Asn Ala Pro Glu Tyr Ala Phe Arg Ile Arg
      130            135            140

atc tat cac ctg aaa acg ctg ccc agt ttt ctg gct ctg cag gtg cag 480
Ile Tyr His Leu Lys Thr Leu Pro Ser Phe Leu Ala Leu Gln Val Gln
      145            150            155            160

att ata tcc ttt atc ctg gag gtc atg aag gtg aac ata agg gtt cgc 528
Ile Ile Ser Phe Ile Leu Glu Val Met Lys Val Asn Ile Arg Val Arg
      165            170            175

cag acc aag ctg cag ctg ctg att ctg gcc agg gaa tta tcc tgt cgt 576
Gln Thr Lys Leu Gln Leu Leu Ile Leu Ala Arg Glu Leu Ser Cys Arg
      180            185            190

tgg cca cag aga aag caa aag ccg cag ttt tcc gac caa cag gcc cat 624
Trp Pro Gln Arg Lys Gln Lys Pro Gln Phe Ser Asp Gln Gln Ala His

```

195	200	205	
cga gtg aaa gac ctg aaa agg cgc tat aat gac cta cat tat ttg ttt	672		
Arg Val Lys Asp Leu Lys Arg Arg Tyr Asn Asp Leu His Tyr Leu Phe			
210 215 220			
gtt cgc att aat ggc tat ttt gga ggc agc ctt ctg acc atc atc att	720		
Val Arg Ile Asn Gly Tyr Phe Gly Gly Ser Leu Leu Thr Ile Ile Ile			
225 230 235 240			
gtt cac ttt gcg atc ttt gtt tcc aac tcc tat tgg cta ttc gtg gat	768		
Val His Phe Ala Ile Phe Val Ser Asn Ser Tyr Trp Leu Phe Val Asp			
245 250 255			
atc aga aca aga ccc tgg agg ata tat gcc atc ttg ctc aac ttg gga	816		
Ile Arg Thr Arg Pro Trp Arg Ile Tyr Ala Ile Leu Leu Asn Leu Gly			
260 265 270			
ttt att ttc aat gta gcc ctg caa atg gca gct gct tgt tgg cac tgt	864		
Phe Ile Phe Asn Val Ala Leu Gln Met Ala Ala Ala Cys Trp His Cys			
275 280 285			
caa caa agc tat aat cta ggc cgg caa atc ggt tgt cta att tgc aag	912		
Gln Gln Ser Tyr Asn Leu Gly Arg Gln Ile Gly Cys Leu Ile Ser Lys			
290 295 300			
ttg gta aaa ccc cag ggc agc aaa cta tac aat gat ttg gta agt gaa	960		
Leu Val Lys Pro Gln Gly Ser Lys Leu Tyr Asn Asp Leu Val Ser Glu			
305 310 315 320			
ttt tca ctt caa aca ctg cac cag cga ttt gtg gtg acc gcc aag gat	1008		
Phe Ser Leu Gln Thr Leu His Gln Arg Phe Val Val Thr Ala Lys Asp			
325 330 335			
ttc ttc agt ctc aat ctg cat cta ctg agt agt atg ttt gca gct gtg	1056		
Phe Phe Ser Leu Asn Leu His Leu Leu Ser Ser Met Phe Ala Ala Val			
340 345 350			
gtc aca tat ctg gtt att ctc ata cag ttc atg ttc gcc gaa agg agt	1104		
Val Thr Tyr Leu Val Ile Leu Ile Gln Phe Met Phe Ala Glu Arg Ser			
355 360 365			
tca acg cga ggc agc gga	1122		
Ser Thr Arg Gly Ser Gly			
370			
<210> 16			
<211> 374			
<212> PRT			
<213> Drosophila melanogaster			
<400> 16			
Met Phe Pro Pro Thr Arg Val Gln Ala Ser Ser Arg Val Val Leu Lys			
1 5 10 15			

Ile Phe His Phe Ile Leu Val Ala Phe Ser Leu Arg Ser Arg Arg Leu
 20 25 30
 Ser Arg Leu Val Leu Trp Leu Gln Phe Leu Gly Trp Leu Thr Trp Phe
 35 40 45
 Ile Ser Met Trp Thr Gln Ser Val Ile Tyr Ala Gln Thr Ile Asp Cys
 50 55 60
 Thr Leu Asp Cys Ser Leu Arg His Ile Leu Thr Phe Phe Gln Thr Val
 65 70 75 80
 Ser His Ala Phe Ile Val Val Thr Ser Phe Leu Asp Gly Phe Arg Ile
 85 90 95
 Lys Gln Asp Gln Leu Asp Glu Pro Ile Ala Phe Glu Asp Ser Asp Pro
 100 105 110
 Trp Leu Ala Phe Thr Val Leu Ala Met Leu Val Pro Thr Leu Gly Val
 115 120 125
 Glu Tyr Leu Val Cys Ser Asn Ala Pro Glu Tyr Ala Phe Arg Ile Arg
 130 135 140
 Ile Tyr His Leu Lys Thr Leu Pro Ser Phe Leu Ala Leu Gln Val Gln
 145 150 155 160
 Ile Ile Ser Phe Ile Leu Glu Val Met Lys Val Asn Ile Arg Val Arg
 165 170 175
 Gln Thr Lys Leu Gln Leu Leu Ile Leu Ala Arg Glu Leu Ser Cys Arg
 180 185 190
 Trp Pro Gln Arg Lys Gln Lys Pro Gln Phe Ser Asp Gln Gln Ala His
 195 200 205
 Arg Val Lys Asp Leu Lys Arg Arg Tyr Asn Asp Leu His Tyr Leu Phe
 210 215 220
 Val Arg Ile Asn Gly Tyr Phe Gly Gly Ser Leu Leu Thr Ile Ile Ile
 225 230 235 240
 Val His Phe Ala Ile Phe Val Ser Asn Ser Tyr Trp Leu Phe Val Asp
 245 250 255
 Ile Arg Thr Arg Pro Trp Arg Ile Tyr Ala Ile Leu Leu Asn Leu Gly
 260 265 270
 Phe Ile Phe Asn Val Ala Leu Gln Met Ala Ala Ala Cys Trp His Cys
 275 280 285
 Gln Gln Ser Tyr Asn Leu Gly Arg Gln Ile Gly Cys Leu Ile Ser Lys
 290 295 300
 Leu Val Lys Pro Gln Gly Ser Lys Leu Tyr Asn Asp Leu Val Ser Glu
 305 310 315 320

Phe Ser Leu Gln Thr Leu His Gln Arg Phe Val Val Thr Ala Lys Asp
 325 330 335

Phe Phe Ser Leu Asn Leu His Leu Leu Ser Ser Met Phe Ala Ala Val
 340 345 350

Val Thr Tyr Leu Val Ile Leu Ile Gln Phe Met Phe Ala Glu Arg Ser
 355 360 365

Ser Thr Arg Gly Ser Gly
 370

<210> 17
 <211> 1349
 <212> DNA
 <213> Drosophila melanogaster

<220>
 <221> CDS
 <222> (1)..(1347)
 <223> Coding region GR28A.1

<400> 17
 atg gcc ttt aag ttg tgg gag cgc ttt tca cag gcg gac aat gtg ttc 48
 Met Ala Phe Lys Leu Trp Glu Arg Phe Ser Gln Ala Asp Asn Val Phe
 1 5 10 15

cag gca ctt cga ccg cta aca ttc ata tcg cta ttg ggc ctg gct cca 96
 Gln Ala Leu Arg Pro Leu Thr Phe Ile Ser Leu Leu Gly Leu Ala Pro
 20 25 30

ttt cgt ttg aat ttg aat cct cgc aag gag gtg caa aca tcg aag ttc 144
 Phe Arg Leu Asn Leu Asn Pro Arg Lys Glu Val Gln Thr Ser Lys Phe
 35 40 45

tcc ttc ttt gcc ggc ata gtg cac ttc ctg ttc ttc gtc ctg tgt ttt 192
 Ser Phe Phe Ala Gly Ile Val His Phe Leu Phe Phe Val Leu Cys Phe
 50 55 60

ggc atc tcc gta aag gag gga gat tcc ata ata ggg tac ttc ttt cag 240
 Gly Ile Ser Val Lys Glu Gly Asp Ser Ile Ile Gly Tyr Phe Phe Gln
 65 70 75 80

acc aat atc acc aga ttc agc gat gga acc cta cgc ctg act ggc atc 288
 Thr Asn Ile Thr Arg Phe Ser Asp Gly Thr Leu Arg Leu Thr Gly Ile
 85 90 95

ctg gcc atg tcc act att ttt ggc ttt gcc atg ttc aag aga caa cgt 336
 Leu Ala Met Ser Thr Ile Phe Gly Phe Ala Met Phe Lys Arg Gln Arg
 100 105 110

ttg gtc agc ata ata cag aac aac ata gtg gtg gac gag ata ttt gtg 384
 Leu Val Ser Ile Ile Gln Asn Asn Ile Val Val Asp Glu Ile Phe Val
 115 120 125

```

agg ctg ggc atg aag ttg gac tac cgc agg ata ctg ttg tcc agc ttt 432
Arg Leu Gly Met Lys Leu Asp Tyr Arg Arg Ile Leu Leu Ser Ser Phe
130 135 140

ctc ata tcc ttg ggc atg ctg ctg ttc aac gtc att tac ttg tgt gtg 480
Leu Ile Ser Leu Gly Met Leu Leu Phe Asn Val Ile Tyr Leu Cys Val
145 150 155 160

agc tac agc ctg ctg gtc agt gcc acc ata tgc ccc tca ttt gtg act 528
Ser Tyr Ser Leu Leu Val Ser Ala Thr Ile Ser Pro Ser Phe Val Thr
165 170 175

ttc aca acc ttc gcc ctg ccg cac atc aat atc agc ctg atg gtc ttc 576
Phe Thr Thr Phe Ala Leu Pro His Ile Asn Ile Ser Leu Met Val Phe
180 185 190

aag ttt ctt tgc acc acg gac ttg gcc agg agc cgg ttt agt atg tta 624
Lys Phe Leu Cys Thr Thr Asp Leu Ala Arg Ser Arg Phe Ser Met Leu
195 200 205

aac gaa atc ctg cag gat att ttg gat gcc cat ata gag caa tta agc 672
Asn Glu Ile Leu Gln Asp Ile Leu Asp Ala His Ile Glu Gln Leu Ser
210 215 220

gcc ttg gaa ctc tca ccc atg cac tcg gtt gtc aat cac aga cgc tac 720
Ala Leu Glu Leu Ser Pro Met His Ser Val Val Asn His Arg Arg Tyr
225 230 235 240

tca cat cgc cta cga aat ttg att agt acg cca atg aag cgg tac agt 768
Ser His Arg Leu Arg Asn Leu Ile Ser Thr Pro Met Lys Arg Tyr Ser
245 250 255

gtt acc tcc gtc ata cgc ctc aat ccc gaa tac gca atc aaa cag gtg 816
Val Thr Ser Val Ile Arg Leu Asn Pro Glu Tyr Ala Ile Lys Gln Val
260 265 270

tcc aac ata cac aat ctg ctc tgc gac att tgc cag acc atc gag gaa 864
Ser Asn Ile His Asn Leu Leu Cys Asp Ile Cys Gln Thr Ile Glu Glu
275 280 285

tac ttc aca tat ccg ctg ctc gga atc ata gcc ata tcc ttt ctg ttc 912
Tyr Phe Thr Tyr Pro Leu Leu Gly Ile Ile Ala Ile Ser Phe Leu Phe
290 295 300

att ctc ttt gat gac ttt tac att ttg gaa gcc att ctg aat ccc aaa 960
Ile Leu Phe Asp Asp Phe Tyr Ile Leu Glu Ala Ile Leu Asn Pro Lys
305 310 315 320

cga ctg gat gtc ttt gag gcc gat gag ttc ttt gcc ttc ttc ctg atg 1008
Arg Leu Asp Val Phe Glu Ala Asp Glu Phe Phe Ala Phe Phe Leu Met
325 330 335

cag ctc atc tgg tat ata gtg atc ata gtg ctg atc gtg gag ggc agc 1056
Gln Leu Ile Trp Tyr Ile Val Ile Ile Val Leu Ile Val Glu Gly Ser
340 345 350

```

```

agt cgg act att ttg cat agc agc tat acg gca gct ata gtt cac aag 1104
Ser Arg Thr Ile Leu His Ser Ser Tyr Thr Ala Ala Ile Val His Lys
      355              360              365

att ctc aat atc acc gat gat cca gaa ctc aga gat cga ctt ttc cgt 1152
Ile Leu Asn Ile Thr Asp Asp Pro Glu Leu Arg Asp Arg Leu Phe Arg
      370              375              380

ctg tcc ttg cag ctg tcg cat cgg aag gtc ctt ttc aca gcc gca gga 1200
Leu Ser Leu Gln Leu Ser His Arg Lys Val Leu Phe Thr Ala Ala Gly
      385              390              395              400

ctt ttt cgt ctg gat cgc aca ctg att ttt acg att act ggt gct gcc 1248
Leu Phe Arg Leu Asp Arg Thr Leu Ile Phe Thr Ile Thr Gly Ala Ala
      405              410              415

acg tgc tac ctc att atc cta att cag ttt cga ttc aca cat cac atg 1296
Thr Cys Tyr Leu Ile Ile Leu Ile Gln Phe Arg Phe Thr His His Met
      420              425              430

gac gac acc agc tcc aat tca aca aat aat tta cat tcc att cat ctc 1344
Asp Asp Thr Ser Ser Asn Ser Thr Asn Asn Leu His Ser Ile His Leu
      435              440              445

ggc ga 1349
Gly

```

<210> 18

<211> 449

<212> PRT

<213> *Drosophila melanogaster*

<400> 18

```

Met Ala Phe Lys Leu Trp Glu Arg Phe Ser Gln Ala Asp Asn Val Phe
  1              5              10              15

Gln Ala Leu Arg Pro Leu Thr Phe Ile Ser Leu Leu Gly Leu Ala Pro
      20              25              30

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

Ser Phe Phe Ala Gly Ile Val His Phe Leu Phe Phe Val Leu Cys Phe
      50              55              60

Gly Ile Ser Val Lys Glu Gly Asp Ser Ile Ile Gly Tyr Phe Phe Gln
      65              70              75              80

Thr Asn Ile Thr Arg Phe Ser Asp Gly Thr Leu Arg Leu Thr Gly Ile
      85              90              95

Leu Ala Met Ser Thr Ile Phe Gly Phe Ala Met Phe Lys Arg Gln Arg
      100             105             110

Leu Val Ser Ile Ile Gln Asn Asn Ile Val Val Asp Glu Ile Phe Val

```

115	120	125
Arg Leu Gly Met Lys Leu Asp Tyr Arg Arg Ile Leu Leu Ser Ser Phe		
130	135	140
Leu Ile Ser Leu Gly Met Leu Leu Phe Asn Val Ile Tyr Leu Cys Val		
145	150	155 160
Ser Tyr Ser Leu Leu Val Ser Ala Thr Ile Ser Pro Ser Phe Val Thr		
	165	170 175
Phe Thr Thr Phe Ala Leu Pro His Ile Asn Ile Ser Leu Met Val Phe		
	180	185 190
Lys Phe Leu Cys Thr Thr Asp Leu Ala Arg Ser Arg Phe Ser Met Leu		
	195	200 205
Asn Glu Ile Leu Gln Asp Ile Leu Asp Ala His Ile Glu Gln Leu Ser		
	210	215 220
Ala Leu Glu Leu Ser Pro Met His Ser Val Val Asn His Arg Arg Tyr		
	225	230 235 240
Ser His Arg Leu Arg Asn Leu Ile Ser Thr Pro Met Lys Arg Tyr Ser		
	245	250 255
Val Thr Ser Val Ile Arg Leu Asn Pro Glu Tyr Ala Ile Lys Gln Val		
	260	265 270
Ser Asn Ile His Asn Leu Leu Cys Asp Ile Cys Gln Thr Ile Glu Glu		
	275	280 285
Tyr Phe Thr Tyr Pro Leu Leu Gly Ile Ile Ala Ile Ser Phe Leu Phe		
	290	295 300
Ile Leu Phe Asp Asp Phe Tyr Ile Leu Glu Ala Ile Leu Asn Pro Lys		
	305	310 315 320
Arg Leu Asp Val Phe Glu Ala Asp Glu Phe Phe Ala Phe Phe Leu Met		
	325	330 335
Gln Leu Ile Trp Tyr Ile Val Ile Ile Val Leu Ile Val Glu Gly Ser		
	340	345 350
Ser Arg Thr Ile Leu His Ser Ser Tyr Thr Ala Ala Ile Val His Lys		
	355	360 365
Ile Leu Asn Ile Thr Asp Asp Pro Glu Leu Arg Asp Arg Leu Phe Arg		
	370	375 380
Leu Ser Leu Gln Leu Ser His Arg Lys Val Leu Phe Thr Ala Ala Gly		
	385	390 395 400
Leu Phe Arg Leu Asp Arg Thr Leu Ile Phe Thr Ile Thr Gly Ala Ala		
	405	410 415

Thr Cys Tyr Leu Ile Ile Leu Ile Gln Phe Arg Phe Thr His His Met
420 425 430

Asp Asp Thr Ser Ser Asn Ser Thr Asn Asn Leu His Ser Ile His Leu
435 440 445

Gly

<210> 19

<211> 1176

<212> DNA

<213> Drosophila melanogaster

<220>

<221> CDS

<222> (1)..(1176)

<223> Coding region GR32D.1

<220>

<221> misc_feature

<222> (931)..(1084)

<223> Splice sites before positions 931, 994 and 1084.

<400> 19

atg cca ata tac gaa cag gtc tcc gac tat gag gtg ggt cct ccg acc 48
Met Pro Ile Tyr Glu Gln Val Ser Asp Tyr Glu Val Gly Pro Pro Thr
1 5 10 15

aaa acg aat gag ttc tac tcg ttc ttc gta aga ggc gtg gtc cac gcc 96
Lys Thr Asn Glu Phe Tyr Ser Phe Phe Val Arg Gly Val Val His Ala
20 25 30

ctg acc atc ttc aat gtc tac agc tta ttt aca ccg ata tcg gcg caa 144
Leu Thr Ile Phe Asn Val Tyr Ser Leu Phe Thr Pro Ile Ser Ala Gln
35 40 45

tta ttt ttc tcc tac cgc gag acg gac aat gtg aac cag tgg atc gag 192
Leu Phe Phe Ser Tyr Arg Glu Thr Asp Asn Val Asn Gln Trp Ile Glu
50 55 60

ctg ctg ctc tgc atc cta acc tat acc ctt acg gtt ttc gtt tgt gcc 240
Leu Leu Leu Cys Ile Leu Thr Tyr Thr Leu Thr Val Phe Val Cys Ala
65 70 75 80

cac aat acc acg agc atg ttg cgg ata atg aat gaa atc ctt caa etc 288
His Asn Thr Thr Ser Met Leu Arg Ile Met Asn Glu Ile Leu Gln Leu
85 90 95

gac gag gaa gtg cgt cgc cag ttt ggt gcc aat ttg agc caa aat ttt 336
Asp Glu Glu Val Arg Arg Gln Phe Gly Ala Asn Leu Ser Gln Asn Phe
100 105 110

ggg ttc ttg gta aag ttt tta gta gga atc acc gct tgc cag gct tat 384

Gly Phe Leu Val Lys Phe Leu Val Gly Ile Thr Ala Cys Gln Ala Tyr	
115 120 125	
ata atc gtg ctc aaa ata tat gcc gtg caa ggc gag atc aca ccc acc	432
Ile Ile Val Leu Lys Ile Tyr Ala Val Gln Gly Glu Ile Thr Pro Thr	
130 135 140	
tcc tat ata cta ttg gcc ttc tat ggc atc cag aac ggt ctg acc gcc	480
Ser Tyr Ile Leu Leu Ala Phe Tyr Gly Ile Gln Asn Gly Leu Thr Ala	
145 150 155 160	
aca tat ata gta ttc gca tcg gct ttg ctc agg att gtg tac atc cgg	528
Thr Tyr Ile Val Phe Ala Ser Ala Leu Leu Arg Ile Val Tyr Ile Arg	
165 170 175	
ttt cat ttc atc aac cag ctg cta aat ggt tac acc tat ggg cag cag	576
Phe His Phe Ile Asn Gln Leu Leu Asn Gly Tyr Thr Tyr Gly Gln Gln	
180 185 190	
cat agg cgc aaa gaa ggt gga gca cga gcg agg cgg cag cgt ggt gac	624
His Arg Arg Lys Glu Gly Gly Ala Arg Ala Arg Arg Gln Arg Gly Asp	
195 200 205	
gtc aat ccc aat gtc aat cct gct cta atg gaa cat ttc ccg gaa gac	672
Val Asn Pro Asn Val Asn Pro Ala Leu Met Glu His Phe Pro Glu Asp	
210 215 220	
tcg cta ttc ata tac cgc atg cac aac aaa ctg ctg cgt atc tac aag	720
Ser Leu Phe Ile Tyr Arg Met His Asn Lys Leu Leu Arg Ile Tyr Lys	
225 230 235 240	
ggc atc aac gat tgc tgc aac ttg att ctg gtc tcg ttt ctg ggc tac	768
Gly Ile Asn Asp Cys Cys Asn Leu Ile Leu Val Ser Phe Leu Gly Tyr	
245 250 255	
tcc ttt tac acg gtc acc acc aac tgc tac aac ctc ttt gtc cag att	816
Ser Phe Tyr Thr Val Thr Thr Asn Cys Tyr Asn Leu Phe Val Gln Ile	
260 265 270	
acc ggc aag ggc atg gtt tca cca aac ata ttg cag tgg tgc ttc gcc	864
Thr Gly Lys Gly Met Val Ser Pro Asn Ile Leu Gln Trp Cys Phe Ala	
275 280 285	
tgg tta tgt ctc cac gtt tcc ctg ctg gct ttg ctg tca cgc agc tgt	912
Trp Leu Cys Leu His Val Ser Leu Leu Ala Leu Leu Ser Arg Ser Cys	
290 295 300	
ggg ctg acc acc acg gag gcc aat gcc aca tcc caa att ctt gca agg	960
Gly Leu Thr Thr Thr Glu Ala Asn Ala Thr Ser Gln Ile Leu Ala Arg	
305 310 315 320	
gtg tat gcc aag tcg aag gag tat cag aat atc att gat aag ttc ctt	1008
Val Tyr Ala Lys Ser Lys Glu Tyr Gln Asn Ile Ile Asp Lys Phe Leu	
325 330 335	
acg aag agc att aaa cag gag gtg caa ttc acg gca tat gga ttt ttt	1056

Thr Lys Ser Ile Lys Gln Glu Val Gln Phe Thr Ala Tyr Gly Phe Phe
 340 345 350
 gcg ata gat aac tcc aca cta ttc aag ata ttt tcg gcc gtc aca aca 1104
 Ala Ile Asp Asn Ser Thr Leu Phe Lys Ile Phe Ser Ala Val Thr Thr
 355 360 365
 tac ttg gta atc ttg att cag ttc aaa cag ctc gaa gac tcg aaa gta 1152
 Tyr Leu Val Ile Leu Ile Gln Phe Lys Gln Leu Glu Asp Ser Lys Val
 370 375 380
 gag gat cct gta cca gaa caa act
 Glu Asp Pro Val Pro Glu Gln Thr 1176
 385 390

 <210> 20
 <211> 392
 <212> PRT
 <213> *Drosophila melanogaster*

 <400> 20
 Met Pro Ile Tyr Glu Gln Val Ser Asp Tyr Glu Val Gly Pro Pro Thr
 1 5 10 15
 Lys Thr Asn Glu Phe Tyr Ser Phe Phe Val Arg Gly Val Val His Ala
 20 25 30
 Leu Thr Ile Phe Asn Val Tyr Ser Leu Phe Thr Pro Ile Ser Ala Gln
 35 40 45
 Leu Phe Phe Ser Tyr Arg Glu Thr Asp Asn Val Asn Gln Trp Ile Glu
 50 55 60
 Leu Leu Leu Cys Ile Leu Thr Tyr Thr Leu Thr Val Phe Val Cys Ala
 65 70 75 80
 His Asn Thr Thr Ser Met Leu Arg Ile Met Asn Glu Ile Leu Gln Leu
 85 90 95
 Asp Glu Glu Val Arg Arg Gln Phe Gly Ala Asn Leu Ser Gln Asn Phe
 100 105 110
 Gly Phe Leu Val Lys Phe Leu Val Gly Ile Thr Ala Cys Gln Ala Tyr
 115 120 125
 Ile Ile Val Leu Lys Ile Tyr Ala Val Gln Gly Glu Ile Thr Pro Thr
 130 135 140
 Ser Tyr Ile Leu Leu Ala Phe Tyr Gly Ile Gln Asn Gly Leu Thr Ala
 145 150 155 160
 Thr Tyr Ile Val Phe Ala Ser Ala Leu Leu Arg Ile Val Tyr Ile Arg
 165 170 175
 Phe His Phe Ile Asn Gln Leu Leu Asn Gly Tyr Thr Tyr Gly Gln Gln

```

                180                185                190
His Arg Arg Lys Glu Gly Gly Ala Arg Ala Arg Arg Gln Arg Gly Asp
   195                200                205
Val Asn Pro Asn Val Asn Pro Ala Leu Met Glu His Phe Pro Glu Asp
   210                215                220
Ser Leu Phe Ile Tyr Arg Met His Asn Lys Leu Leu Arg Ile Tyr Lys
   225                230                235                240
Gly Ile Asn Asp Cys Cys Asn Leu Ile Leu Val Ser Phe Leu Gly Tyr
   245                250                255
Ser Phe Tyr Thr Val Thr Thr Asn Cys Tyr Asn Leu Phe Val Gln Ile
   260                265                270
Thr Gly Lys Gly Met Val Ser Pro Asn Ile Leu Gln Trp Cys Phe Ala
   275                280                285
Trp Leu Cys Leu His Val Ser Leu Leu Ala Leu Leu Ser Arg Ser Cys
   290                295                300
Gly Leu Thr Thr Thr Glu Ala Asn Ala Thr Ser Gln Ile Leu Ala Arg
   305                310                315                320
Val Tyr Ala Lys Ser Lys Glu Tyr Gln Asn Ile Ile Asp Lys Phe Leu
   325                330                335
Thr Lys Ser Ile Lys Gln Glu Val Gln Phe Thr Ala Tyr Gly Phe Phe
   340                345                350
Ala Ile Asp Asn Ser Thr Leu Phe Lys Ile Phe Ser Ala Val Thr Thr
   355                360                365
Tyr Leu Val Ile Leu Ile Gln Phe Lys Gln Leu Glu Asp Ser Lys Val
   370                375                380
Glu Asp Pro Val Pro Glu Gln Thr
   385                390

```

```

<210> 21
<211> 1419
<212> DNA
<213> Drosophila melanogaster

```

```

<220>
<221> CDS
<222> (1)..(1416)
<223> Coding region GR33C.1

```

```

<400> 21
atg aaa cgg aag gcc gtg gaa gtc ata ggc ctc att cca ctg aat cgc   48
Met Lys Arg Lys Ala Val Glu Val Ile Gly Leu Ile Pro Leu Asn Arg

```

1	5	10	15	
cag caa tcg gaa act aac ttc ata ctc gac tac gcc atg atg tgc att				96
Gln Gln Ser Glu Thr Asn Phe Ile Leu Asp Tyr Ala Met Met Cys Ile				
	20	25	30	
gtg ccc atc ttc tat gtg gct tgc tat ctt ctc ata aat ctt agc cac				144
Val Pro Ile Phe Tyr Val Ala Cys Tyr Leu Leu Ile Asn Leu Ser His				
	35	40	45	
att att ggc ctc tgt tta ctg gac tct tgc aat agt gtt tgc aag ctg				192
Ile Ile Gly Leu Cys Leu Leu Asp Ser Cys Asn Ser Val Cys Lys Leu				
	50	55	60	
agc agc cat ctc ttt atg cat ttg ggc gca ttt cta tat ctg acc atc				240
Ser Ser His Leu Phe Met His Leu Gly Ala Phe Leu Tyr Leu Thr Ile				
	65	70	75	80
acc ctg ctg tca ctt tat cgc cga aag gag ttt ttc cag cag ttt gat				288
Thr Leu Leu Ser Leu Tyr Arg Arg Lys Glu Phe Phe Gln Gln Phe Asp				
	85	90	95	
gcg aga ctt aat gac att gat gca gtt atc cag aaa tgc cag cgg gtg				336
Ala Arg Leu Asn Asp Ile Asp Ala Val Ile Gln Lys Cys Gln Arg Val				
	100	105	110	
gcg gaa atg gac aag gtg aag gtt act gcg gtg aaa cac agt gtg gcc				384
Ala Glu Met Asp Lys Val Lys Val Thr Ala Val Lys His Ser Val Ala				
	115	120	125	
tat cac ttc acc tgg ctt ttc cta ttc tgc gtt ttc acc ttt gcc ctt				432
Tyr His Phe Thr Trp Leu Phe Leu Phe Cys Val Phe Thr Phe Ala Leu				
	130	135	140	
tac tat gac gtc aga tct ttg tac ttg acc ttc ggc aat ctc gcc ttc				480
Tyr Tyr Asp Val Arg Ser Leu Tyr Leu Thr Phe Gly Asn Leu Ala Phe				
	145	150	155	160
att ccg ttc atg gtg tcc agt ttc cca tat ttg gcc ggc agc atc att				528
Ile Pro Phe Met Val Ser Ser Phe Pro Tyr Leu Ala Gly Ser Ile Ile				
	165	170	175	
cag ggt gag ttc atc tat cac gtg tcg gtc atc tcg cag cgc ttc gag				576
Gln Gly Glu Phe Ile Tyr His Val Ser Val Ile Ser Gln Arg Phe Glu				
	180	185	190	
cag att aac atg ctg ctg gag aag att aac cag gag gcg cgc cat cgc				624
Gln Ile Asn Met Leu Leu Glu Lys Ile Asn Gln Glu Ala Arg His Arg				
	195	200	205	
cac gca ccc ctt acc gtg ttc gat atc gag agc gag ggc aaa aag gag				672
His Ala Pro Leu Thr Val Phe Asp Ile Glu Ser Glu Gly Lys Lys Glu				
	210	215	220	
cgg aag acc gtt aca ccg att acg gtc atg gac ggc agg acg acg aca				720
Arg Lys Thr Val Thr Pro Ile Thr Val Met Asp Gly Arg Thr Thr Thr				

225	230	235	240	
gga ttt ggc aat gag aac aag ttc gcc ggc gaa atg aag cgc cag gag				768
Gly Phe Gly Asn Glu Asn Lys Phe Ala Gly Glu Met Lys Arg Gln Glu				
	245	250	255	
ggg caa caa aag aac gac gat gac gat ttg gac acc agc aac gac gag				816
Gly Gln Gln Lys Asn Asp Asp Asp Leu Asp Thr Ser Asn Asp Glu				
	260	265	270	
gac gag gat gac ttt gat tat gac aat gcc acc atc gcg gaa aat act				864
Asp Glu Asp Asp Phe Asp Tyr Asp Asn Ala Thr Ile Ala Glu Asn Thr				
	275	280	285	
gga aac aca tcg gaa gcc aat tta cca gat ctc ttc aag ctg cat gat				912
Gly Asn Thr Ser Glu Ala Asn Leu Pro Asp Leu Phe Lys Leu His Asp				
	290	295	300	
aaa atc ctg gcg ctc agc gtg att aca aac ggc gag ttt gga cca cag				960
Lys Ile Leu Ala Leu Ser Val Ile Thr Asn Gly Glu Phe Gly Pro Gln				
	305	310	315	320
tgt gta ccc tat atg gcg gcc tgc ttt gtg gtg agc atc ttt ggc att				1008
Cys Val Pro Tyr Met Ala Ala Cys Phe Val Val Ser Ile Phe Gly Ile				
	325	330	335	
ttc ctg gag acc aag gtc aac ttc att gtg ggc ggc aag agt cga ctg				1056
Phe Leu Glu Thr Lys Val Asn Phe Ile Val Gly Gly Lys Ser Arg Leu				
	340	345	350	
ctg gac tat atg acc tat ttg tac gtg att tgg agc ttc acc acc atg				1104
Leu Asp Tyr Met Thr Tyr Leu Tyr Val Ile Trp Ser Phe Thr Thr Met				
	355	360	365	
atg gtg gcc tac ata gtg ctt cga cta tgc tgc aat gcc aac aat cat				1152
Met Val Ala Tyr Ile Val Leu Arg Leu Cys Cys Asn Ala Asn Asn His				
	370	375	380	
tcc aaa caa tcg gcg atg att gtc cat gag att atg caa aag aaa ccg				1200
Ser Lys Gln Ser Ala Met Ile Val His Glu Ile Met Gln Lys Lys Pro				
	385	390	395	400
gca ttt atg ttg agc aac gat ctg ttc tac aac aaa atg aag tca ttc				1248
Ala Phe Met Leu Ser Asn Asp Leu Phe Tyr Asn Lys Met Lys Ser Phe				
	405	410	415	
aca ctg cag ttt ctt cac tgg gag ggt ttc ttt caa ttc aac ggc gtg				1296
Thr Leu Gln Phe Leu His Trp Glu Gly Phe Phe Gln Phe Asn Gly Val				
	420	425	430	
gga ttg ttt gct ctg gac tac aca ttc att ttc tcg act gta agt gca				1344
Gly Leu Phe Ala Leu Asp Tyr Thr Phe Ile Phe Ser Thr Val Ser Ala				
	435	440	445	
gcc aca tcc tat tta att gtc ctg ctg cag ttt gac atg act gca att				1392
Ala Thr Ser Tyr Leu Ile Val Leu Leu Gln Phe Asp Met Thr Ala Ile				

450 455 460

ttg cgc aac gag ggg cta atg tca taa 1419
 Leu Arg Asn Glu Gly Leu Met Ser
 465 470

<210> 22
 <211> 472
 <212> PRT
 <213> *Drosophila melanogaster*

<400> 22
 Met Lys Arg Lys Ala Val Glu Val Ile Gly Leu Ile Pro Leu Asn Arg
 1 5 10 15
 Gln Gln Ser Glu Thr Asn Phe Ile Leu Asp Tyr Ala Met Met Cys Ile
 20 25 30
 Val Pro Ile Phe Tyr Val Ala Cys Tyr Leu Leu Ile Asn Leu Ser His
 35 40 45
 Ile Ile Gly Leu Cys Leu Leu Asp Ser Cys Asn Ser Val Cys Lys Leu
 50 55 60
 Ser Ser His Leu Phe Met His Leu Gly Ala Phe Leu Tyr Leu Thr Ile
 65 70 75 80
 Thr Leu Leu Ser Leu Tyr Arg Arg Lys Glu Phe Phe Gln Gln Phe Asp
 85 90 95
 Ala Arg Leu Asn Asp Ile Asp Ala Val Ile Gln Lys Cys Gln Arg Val
 100 105 110
 Ala Glu Met Asp Lys Val Lys Val Thr Ala Val Lys His Ser Val Ala
 115 120 125
 Tyr His Phe Thr Trp Leu Phe Leu Phe Cys Val Phe Thr Phe Ala Leu
 130 135 140
 Tyr Tyr Asp Val Arg Ser Leu Tyr Leu Thr Phe Gly Asn Leu Ala Phe
 145 150 155 160
 Ile Pro Phe Met Val Ser Ser Phe Pro Tyr Leu Ala Gly Ser Ile Ile
 165 170 175
 Gln Gly Glu Phe Ile Tyr His Val Ser Val Ile Ser Gln Arg Phe Glu
 180 185 190
 Gln Ile Asn Met Leu Leu Glu Lys Ile Asn Gln Glu Ala Arg His Arg
 195 200 205
 His Ala Pro Leu Thr Val Phe Asp Ile Glu Ser Glu Gly Lys Lys Glu
 210 215 220
 Arg Lys Thr Val Thr Pro Ile Thr Val Met Asp Gly Arg Thr Thr Thr

WO 00/77208

PCT/US00/16211

225 230 235 240
 Gly Phe Gly Asn Glu Asn Lys Phe Ala Gly Glu Met Lys Arg Gln Glu
 245 250 255
 Gly Gln Gln Lys Asn Asp Asp Asp Asp Leu Asp Thr Ser Asn Asp Glu
 260 265 270
 Asp Glu Asp Asp Phe Asp Tyr Asp Asn Ala Thr Ile Ala Glu Asn Thr
 275 280 285
 Gly Asn Thr Ser Glu Ala Asn Leu Pro Asp Leu Phe Lys Leu His Asp
 290 295 300
 Lys Ile Leu Ala Leu Ser Val Ile Thr Asn Gly Glu Phe Gly Pro Gln
 305 310 315 320
 Cys Val Pro Tyr Met Ala Ala Cys Phe Val Val Ser Ile Phe Gly Ile
 325 330 335
 Phe Leu Glu Thr Lys Val Asn Phe Ile Val Gly Gly Lys Ser Arg Leu
 340 345 350
 Leu Asp Tyr Met Thr Tyr Leu Tyr Val Ile Trp Ser Phe Thr Thr Met
 355 360 365
 Met Val Ala Tyr Ile Val Leu Arg Leu Cys Cys Asn Ala Asn Asn His
 370 375 380
 Ser Lys Gln Ser Ala Met Ile Val His Glu Ile Met Gln Lys Lys Pro
 385 390 395 400
 Ala Phe Met Leu Ser Asn Asp Leu Phe Tyr Asn Lys Met Lys Ser Phe
 405 410 415
 Thr Leu Gln Phe Leu His Trp Glu Gly Phe Phe Gln Phe Asn Gly Val
 420 425 430
 Gly Leu Phe Ala Leu Asp Tyr Thr Phe Ile Phe Ser Thr Val Ser Ala
 435 440 445
 Ala Thr Ser Tyr Leu Ile Val Leu Leu Gln Phe Asp Met Thr Ala Ile
 450 455 460
 Leu Arg Asn Glu Gly Leu Met Ser
 465 470

<210> 23

<211> 1113

<212> DNA

<213> *Drosophila melanogaster*

<220>

<221> CDS

<222> (1)..(1113)

<223> Coding region GR36B.1

<400> 23

```

atg ttt gac tgg gtc ggt ttg ttg tta aag gta ctc tac tac tat ggg 48
Met Phe Asp Trp Val Gly Leu Leu Leu Lys Val Leu Tyr Tyr Tyr Gly
  1             5             10             15

cag atc att gga ctt atc aac ttc gaa att gac tgg caa aga ggt cgt 96
Gln Ile Ile Gly Leu Ile Asn Phe Glu Ile Asp Trp Gln Arg Gly Arg
      20             25             30

gtc gtt gca gcc caa aga ggc att ctt ttc gca atc gca att aac gtc 144
Val Val Ala Ala Gln Arg Gly Ile Leu Phe Ala Ile Ala Ile Asn Val
      35             40             45

tta att tgc atg gtg ctg ctt ttg caa ata tcc aag aaa ttc aat ctc 192
Leu Ile Cys Met Val Leu Leu Leu Gln Ile Ser Lys Lys Phe Asn Leu
      50             55             60

gat gtg tac ttc ggt agg gct aac cag ctg cat caa tat gtg atc atc 240
Asp Val Tyr Phe Gly Arg Ala Asn Gln Leu His Gln Tyr Val Ile Ile
      65             70             75             80

cga gca caa cta atg cgc ctt gtt gaa tgt gta ctt cgg cta ttt ctg 288
Arg Ala Gln Leu Met Arg Leu Val Glu Cys Val Leu Arg Leu Phe Leu
      85             90             95

aaa aaa ccg cat gta aag caa atg tcc cga tgg gca att ctg gta aag 336
Lys Lys Pro His Val Lys Gln Met Ser Arg Trp Ala Ile Leu Val Lys
      100             105             110

ttc tct gta ggt gtc gtc agc aat ttc cta caa atg gcc atc tct atg 384
Phe Ser Val Gly Val Val Ser Asn Phe Leu Gln Met Ala Ile Ser Met
      115             120             125

gaa tca ttg gat cgc ttg ggg ttc aac gaa ttc gtg gga atg gct tcg 432
Glu Ser Leu Asp Arg Leu Gly Phe Asn Glu Phe Val Gly Met Ala Ser
      130             135             140

gat ttc tgg atg tcg gcc att ata aat atg gcc ata tca caa cac tat 480
Asp Phe Trp Met Ser Ala Ile Ile Asn Met Ala Ile Ser Gln His Tyr
      145             150             155             160

ttg gta ata ctt ttc gtt cga gcc tat tac cat ttg ctc aag aca gag 528
Leu Val Ile Leu Phe Val Arg Ala Tyr Tyr His Leu Leu Lys Thr Glu
      165             170             175

gtg cgg cag gcg atc cat gaa agc caa atg tta agt gag att tac cca 576
Val Arg Gln Ala Ile His Glu Ser Gln Met Leu Ser Glu Ile Tyr Pro
      180             185             190

cgg aga gcg gct ttc atg acc aag tgt tgt tac ttg gct gat cga ata 624
Arg Arg Ala Ala Phe Met Thr Lys Cys Cys Tyr Leu Ala Asp Arg Ile
      195             200             205

```

gat aat ata gca aaa ctt cag aat caa ctg caa tcg att gtg acc cag 672
 Asp Asn Ile Ala Lys Leu Gln Asn Gln Leu Ser Ile Val Thr Gln
 210 215 220

 ttg aac caa gtg ttt ggc atc caa ggg ata atg gtt tat ggc gga tac 720
 Leu Asn Gln Val Phe Gly Ile Gln Gly Ile Met Val Tyr Gly Gly Tyr
 225 230 235 240

 tat ata ttc tca gta gct aca act tac ata acg tac agt tta gct ata 768
 Tyr Ile Phe Ser Val Ala Thr Thr Tyr Ile Thr Tyr Ser Leu Ala Ile
 245 250 255

 aat ggt ata gaa gaa ctg cac ttg agt gtc aga gca gcg gca ctg gta 816
 Asn Gly Ile Glu Glu Leu His Leu Ser Val Arg Ala Ala Leu Val
 260 265 270

 ttt agt tgg ttt tta ttc tac tac acg agt gca ata cta aat ctg ttt 864
 Phe Ser Trp Phe Leu Phe Tyr Tyr Thr Ser Ala Ile Leu Asn Leu Phe
 275 280 285

 gtt atg ctc aaa ctc ttc gat gat cac aag gag atg gaa cgg ata cta 912
 Val Met Leu Lys Leu Phe Asp Asp His Lys Glu Met Glu Arg Ile Leu
 290 295 300

 gaa gag aga act ctg ttt act tcc gcc ttg gat gtc cgc ttg gag caa 960
 Glu Glu Arg Thr Leu Phe Thr Ser Ala Leu Asp Val Arg Leu Glu Gln
 305 310 315 320

 tcc ttt gaa agc att caa ttg cag cta att aga aac ccg ttg aaa att 1008
 Ser Phe Glu Ser Ile Gln Leu Gln Leu Ile Arg Asn Pro Leu Lys Ile
 325 330 335

 gaa gta ttg gat ata ttt acc att act cgc agt tca tct gct gcc atg 1056
 Glu Val Leu Asp Ile Phe Thr Ile Thr Arg Ser Ser Ser Ala Ala Met
 340 345 350

 att gga tct ata ata acg aat tcg ata ttt ctt att caa tac gat atg 1104
 Ile Gly Ser Ile Ile Thr Asn Ser Ile Phe Leu Ile Gln Tyr Asp Met
 355 360 365

 gaa tat ttt 1113
 Glu Tyr Phe
 370

<210> 24
 <211> 371
 <212> PRT
 <213> *Drosophila melanogaster*

<400> 24
 Met Phe Asp Trp Val Gly Leu Leu Leu Lys Val Leu Tyr Tyr Tyr Gly
 1 5 10 15
 Gln Ile Ile Gly Leu Ile Asn Phe Glu Ile Asp Trp Gln Arg Gly Arg
 20 25 30

Val Val Ala Ala Gln Arg Gly Ile Leu Phe Ala Ile Ala Ile Asn Val
 35 40 45
 Leu Ile Cys Met Val Leu Leu Gln Ile Ser Lys Lys Phe Asn Leu
 50 55 60
 Asp Val Tyr Phe Gly Arg Ala Asn Gln Leu His Gln Tyr Val Ile Ile
 65 70 75 80
 Arg Ala Gln Leu Met Arg Leu Val Glu Cys Val Leu Arg Leu Phe Leu
 85 90 95
 Lys Lys Pro His Val Lys Gln Met Ser Arg Trp Ala Ile Leu Val Lys
 100 105 110
 Phe Ser Val Gly Val Val Ser Asn Phe Leu Gln Met Ala Ile Ser Met
 115 120 125
 Glu Ser Leu Asp Arg Leu Gly Phe Asn Glu Phe Val Gly Met Ala Ser
 130 135 140
 Asp Phe Trp Met Ser Ala Ile Ile Asn Met Ala Ile Ser Gln His Tyr
 145 150 155 160
 Leu Val Ile Leu Phe Val Arg Ala Tyr Tyr His Leu Leu Lys Thr Glu
 165 170 175
 Val Arg Gln Ala Ile His Glu Ser Gln Met Leu Ser Glu Ile Tyr Pro
 180 185 190
 Arg Arg Ala Ala Phe Met Thr Lys Cys Cys Tyr Leu Ala Asp Arg Ile
 195 200 205
 Asp Asn Ile Ala Lys Leu Gln Asn Gln Leu Gln Ser Ile Val Thr Gln
 210 215 220
 Leu Asn Gln Val Phe Gly Ile Gln Gly Ile Met Val Tyr Gly Gly Tyr
 225 230 235 240
 Tyr Ile Phe Ser Val Ala Thr Thr Tyr Ile Thr Tyr Ser Leu Ala Ile
 245 250 255
 Asn Gly Ile Glu Glu Leu His Leu Ser Val Arg Ala Ala Ala Leu Val
 260 265 270
 Phe Ser Trp Phe Leu Phe Tyr Tyr Thr Ser Ala Ile Leu Asn Leu Phe
 275 280 285
 Val Met Leu Lys Leu Phe Asp Asp His Lys Glu Met Glu Arg Ile Leu
 290 295 300
 Glu Glu Arg Thr Leu Phe Thr Ser Ala Leu Asp Val Arg Leu Glu Gln
 305 310 315 320
 Ser Phe Glu Ser Ile Gln Leu Gln Leu Ile Arg Asn Pro Leu Lys Ile
 325 330 335

Glu Val Leu Asp Ile Phe Thr Ile Thr Arg Ser Ser Ser Ala Ala Met
 340 345 350

Ile Gly Ser Ile Ile Thr Asn Ser Ile Phe Leu Ile Gln Tyr Asp Met
 355 360 365

Glu Tyr Phe
 370

<210> 25
 <211> 1134
 <212> DNA
 <213> *Drosophila melanogaster*

<220>
 <221> CDS
 <222> (1)..(1134)
 <223> Coding region GR36B.2

<400> 25
 atg gtc gat tgg gtc gtg ttg ttg ctg aag gca gtc cac ata tac tgc 48
 Met Val Asp Trp Val Val Leu Leu Leu Lys Ala Val His Ile Tyr Cys
 1 5 10 15

tac ttg ata gga tta agt aac ttt gag ttc gat tgc cga aca gga cgc 96
 Tyr Leu Ile Gly Leu Ser Asn Phe Glu Phe Asp Cys Arg Thr Gly Arg
 20 25 30

gta ttt aaa tca agg cga tgt act atc tac gcc ttc atg gcc aac atc 144
 Val Phe Lys Ser Arg Arg Cys Thr Ile Tyr Ala Phe Met Ala Asn Ile
 35 40 45

ttt att ttg att aca ata att tat aat ttt acc gct cat ggt gat acc 192
 Phe Ile Leu Ile Thr Ile Ile Tyr Asn Phe Thr Ala His Gly Asp Thr
 50 55 60

. aat tta ctt ttt caa agc gca aat aag ctg cat gaa tat gtt atc atc 240
 Asn Leu Leu Phe Gln Ser Ala Asn Lys Leu His Glu Tyr Val Ile Ile
 65 70 75 80

ata atg tcc ggc ctc aag atc cgt ggc caa atg atg cag cta gtt aaa 288
 Ile Met Ser Gly Leu Lys Ile Arg Gly Gln Met Met Gln Leu Val Lys
 85 90 95

gac gtt atc cgt cta tat atg atc aat ccg cag tta aag agt atg att 336
 Asp Val Ile Arg Leu Tyr Met Ile Asn Pro Gln Leu Lys Ser Met Ile
 100 105 110

cgc tgg gga att cta ttg aaa gct ttc att agt ttt gca ata gag ctc 384
 Arg Trp Gly Ile Leu Leu Lys Ala Phe Ile Ser Phe Ala Ile Glu Leu
 115 120 125

ctt caa gtg aca ctc tcc gtg gat gca ttg gac cgc caa gga aca gcg 432
 Leu Gln Val Thr Leu Ser Val Asp Ala Leu Asp Arg Gln Gly Thr Ala

130	135	140	
gaa atg atg ggc ttg ctt gta aaa tta tgc gtt tcg ttc att atg aat			480
Glu Met Met Gly Leu Leu Val Lys Leu Cys Val Ser Phe Ile Met Asn			
145	150	155	160
ttg gcc ata tca cag cat ttt ttg gta ata ctt tta att cgg gca caa			528
Leu Ala Ile Ser Gln His Phe Leu Val Ile Leu Leu Ile Arg Ala Gln			
165	170		175
tat cgg att atg aac gca aag ctg cga atg gtg atc gag gaa agc agg			576
Tyr Arg Ile Met Asn Ala Lys Leu Arg Met Val Ile Glu Glu Ser Arg			
180	185		190
agg ttg agt ttc ctg cag ctc cga aat gga gct ttt atg acg aga tgc			624
Arg Leu Ser Phe Leu Gln Leu Arg Asn Gly Ala Phe Met Thr Arg Cys			
195	200		205
tgc tat cta tct gat cag ttg gaa gat ata ggt gag gtt cag agc caa			672
Cys Tyr Leu Ser Asp Gln Leu Glu Asp Ile Gly Glu Val Gln Ser Gln			
210	215		220
cta caa tcg atg gtg ggt caa ctt gat gag gta ttt ggc atg caa ggg			720
Leu Gln Ser Met Val Gly Gln Leu Asp Glu Val Phe Gly Met Gln Gly			
225	230		235
ctt atg gct tat agt gag tac tac cta tcc att gtg ggt aca tct tac			768
Leu Met Ala Tyr Ser Glu Tyr Tyr Leu Ser Ile Val Gly Thr Ser Tyr			
245	250		255
atg tcg tat agc att tat aaa tat ggc ccg cat aat ctg aaa cta tca			816
Met Ser Tyr Ser Ile Tyr Lys Tyr Gly Pro His Asn Leu Lys Leu Ser			
260	265		270
gct aag acc tcg atc att gtc tgc att ttg ata acc cta ttt tac ctt			864
Ala Lys Thr Ser Ile Ile Val Cys Ile Leu Ile Thr Leu Phe Tyr Leu			
275	280		285
gat gcc ttg gtc aac tgc aac aat atg cta cgt gtg ttg gac cat cac			912
Asp Ala Leu Val Asn Cys Asn Asn Met Leu Arg Val Leu Asp His His			
290	295		300
aag gat ttt ttg ggt cta ttg gag gaa cga act gtg ttt gct tcc agc			960
Lys Asp Phe Leu Gly Leu Leu Glu Glu Arg Thr Val Phe Ala Ser Ser			
305	310		315
ttg gac att cga ctg gag gaa tcc ttt gaa agt cta cag ttg caa cta			1008
Leu Asp Ile Arg Leu Glu Glu Ser Phe Glu Ser Leu Gln Leu Gln Leu			
325	330		335
gct cga aac cca tta aaa att aat gta atg ggt atg ttc cct atc acg			1056
Ala Arg Asn Pro Leu Lys Ile Asn Val Met Gly Met Phe Pro Ile Thr			
340	345		350
cgt ggc tca act gcg gct atg tgt gct tct gtt ata gtg aat tcg ata			1104
Arg Gly Ser Thr Ala Ala Met Cys Ala Ser Val Ile Val Asn Ser Ile			

355 360 365
 ttt cta att caa ttt gac atg gaa ttc ttt 1134
 Phe Leu Ile Gln Phe Asp Met Glu Phe Phe
 370 375

 <210> 26
 <211> 378
 <212> PRT
 <213> *Drosophila melanogaster*

 <400> 26
 Met Val Asp Trp Val Val Leu Leu Leu Lys Ala Val His Ile Tyr Cys
 1 5 10 15
 Tyr Leu Ile Gly Leu Ser Asn Phe Glu Phe Asp Cys Arg Thr Gly Arg
 20 25 30
 Val Phe Lys Ser Arg Arg Cys Thr Ile Tyr Ala Phe Met Ala Asn Ile
 35 40 45
 Phe Ile Leu Ile Thr Ile Ile Tyr Asn Phe Thr Ala His Gly Asp Thr
 50 55 60
 Asn Leu Leu Phe Gln Ser Ala Asn Lys Leu His Glu Tyr Val Ile Ile
 65 70 75 80
 Ile Met Ser Gly Leu Lys Ile Arg Gly Gln Met Met Gln Leu Val Lys
 85 90 95
 Asp Val Ile Arg Leu Tyr Met Ile Asn Pro Gln Leu Lys Ser Met Ile
 100 105 110
 Arg Trp Gly Ile Leu Leu Lys Ala Phe Ile Ser Phe Ala Ile Glu Leu
 115 120 125
 Leu Gln Val Thr Leu Ser Val Asp Ala Leu Asp Arg Gln Gly Thr Ala
 130 135 140
 Glu Met Met Gly Leu Leu Val Lys Leu Cys Val Ser Phe Ile Met Asn
 145 150 155 160
 Leu Ala Ile Ser Gln His Phe Leu Val Ile Leu Leu Ile Arg Ala Gln
 165 170 175
 Tyr Arg Ile Met Asn Ala Lys Leu Arg Met Val Ile Glu Glu Ser Arg
 180 185 190
 Arg Leu Ser Phe Leu Gln Leu Arg Asn Gly Ala Phe Met Thr Arg Cys
 195 200 205
 Cys Tyr Leu Ser Asp Gln Leu Glu Asp Ile Gly Glu Val Gln Ser Gln
 210 215 220
 Leu Gln Ser Met Val Gly Gln Leu Asp Glu Val Phe Gly Met Gln Gly

225 230 235 240
 Leu Met Ala Tyr Ser Glu Tyr Tyr Leu Ser Ile Val Gly Thr Ser Tyr
 245 250 255
 Met Ser Tyr Ser Ile Tyr Lys Tyr Gly Pro His Asn Leu Lys Leu Ser
 260 265 270
 Ala Lys Thr Ser Ile Ile Val Cys Ile Leu Ile Thr Leu Phe Tyr Leu
 275 280 285
 Asp Ala Leu Val Asn Cys Asn Asn Met Leu Arg Val Leu Asp His His
 290 295 300
 Lys Asp Phe Leu Gly Leu Leu Glu Glu Arg Thr Val Phe Ala Ser Ser
 305 310 315 320
 Leu Asp Ile Arg Leu Glu Glu Ser Phe Glu Ser Leu Gln Leu Gln Leu
 325 330 335
 Ala Arg Asn Pro Leu Lys Ile Asn Val Met Gly Met Phe Pro Ile Thr
 340 345 350
 Arg Gly Ser Thr Ala Ala Met Cys Ala Ser Val Ile Val Asn Ser Ile
 355 360 365
 Phe Leu Ile Gln Phe Asp Met Glu Phe Phe
 370 375

<210> 27
 <211> 1131
 <212> DNA
 <213> Drosophila melanogaster

<220>
 <221> CDS
 <222> (1)..(1131)
 <223> Coding region GR36B.3

<400> 27
 atg gac ttg gaa agt ttt ttg ttg gga gcg gtt tac tac tat gga ctc 48
 Met Asp Leu Glu Ser Phe Leu Leu Gly Ala Val Tyr Tyr Tyr Gly Leu
 1 5 10 15

 ttc atc ggt ctc agt aac ttt gag ttc gac tgg aac aca gga cgt gta 96
 Phe Ile Gly Leu Ser Asn Phe Glu Phe Asp Trp Asn Thr Gly Arg Val
 20 25 30

 ttt aca aaa aaa tgg agt act ttg tac gca att gcg ttg gat tct tgt 144
 Phe Thr Lys Lys Trp Ser Thr Leu Tyr Ala Ile Ala Leu Asp Ser Cys
 35 40 45

 ata ttt gca cta tat att tat cac tgg acc gga aac aca aac att gtt 192
 Ile Phe Ala Leu Tyr Ile Tyr His Trp Thr Gly Asn Thr Asn Ile Val

50	55	60	
aat gca att ttc ggc aga gca aac atg ttg cac gaa tat gtt gtc gcc			240
Asn Ala Ile Phe Gly Arg Ala Asn Met Leu His Glu Tyr Val Val Ala			
65	70	75	80
ata ttg acc gga ctt aga att cgt tgc aag atg atg gat ttg gcg tca			288
Ile Leu Thr Gly Leu Arg Ile Arg Cys Lys Met Met Asp Leu Ala Ser			
	85	90	95
aaa gtg gtc cga atg tat gta gct agg ccg cag gta agg aga atg tct			336
Lys Val Val Arg Met Tyr Val Ala Arg Pro Gln Val Arg Arg Met Ser			
	100	105	110
cgc tgg ggc att ctc acc aag ttt att ttt ggt tct atc acc gat ggc			384
Arg Trp Gly Ile Leu Thr Lys Phe Ile Phe Gly Ser Ile Thr Asp Gly			
	115	120	125
ctc caa atg gcc atg gtc ttg agt gca atg ggg agt gtg gac tct caa			432
Leu Gln Met Ala Met Val Leu Ser Ala Met Gly Ser Val Asp Ser Gln			
	130	135	140
ttc tac ctg gga ctg ggc ttg cag tac tgg atg ttt gtc att ctc aac			480
Phe Tyr Leu Gly Leu Gly Leu Gln Tyr Trp Met Phe Val Ile Leu Asn			
	145	150	155
atg gct atg atg cag caa cac atg ata atg ctt ttc gtt cgg aca cag			528
Met Ala Met Met Gln Gln His Met Ile Met Leu Phe Val Arg Thr Gln			
	165	170	175
ttc cag ctc att aat acg gaa ctg cgt cag gtg atc gat gaa gcc aag			576
Phe Gln Leu Ile Asn Thr Glu Leu Arg Gln Val Ile Asp Glu Ala Lys			
	180	185	190
gat ctg ctc tta agt ccc cgg cat cag gga gtt ttc atg acc aag tgt			624
Asp Leu Leu Leu Ser Pro Arg His Gln Gly Val Phe Met Thr Lys Cys			
	195	200	205
tgt tcg ttg gcg gat cag ata gag aat ata gca aga att cag agc caa			672
Cys Ser Leu Ala Asp Gln Ile Glu Asn Ile Ala Arg Ile Gln Ser Gln			
	210	215	220
ctg cag acg att atg aac caa atg gaa gaa gta ttc ggc att cag ggg			720
Leu Gln Thr Ile Met Asn Gln Met Glu Glu Val Phe Gly Ile Gln Gly			
	225	230	235
gcc atg acc tat ggg ggc tac tac ttg tcc tcg gtt ggc acc tgt tac			768
Ala Met Thr Tyr Gly Gly Tyr Tyr Leu Ser Ser Val Gly Thr Cys Tyr			
	245	250	255
ttg gca tat agt atc ctc aag cat ggc tat gag aat ttg agc atg aca			816
Leu Ala Tyr Ser Ile Leu Lys His Gly Tyr Glu Asn Leu Ser Met Thr			
	260	265	270
ctg agt act gta atc ctg gcg tac tct tgg tgt ttt ttt tac tac ctt			864
Leu Ser Thr Val Ile Leu Ala Tyr Ser Trp Cys Phe Phe Tyr Tyr Leu			

```

                275                280                285
gat ggt atg ctc aat tta tct gtc atg ctc cac gta cag gat gac tac 912
Asp Gly Met Leu Asn Leu Ser Val Met Leu His Val Gln Asp Asp Tyr
    290                295                300

tgg gaa atg cta caa ata ctc ggg aag cgg aca ata ttc gtt ggc ttg 960
Trp Glu Met Leu Gln Ile Leu Gly Lys Arg Thr Ile Phe Val Gly Leu
    305                310                315                320

gat gtc cgc ctc gaa gaa gcc ttt gag aac ctt aat ttg cag ttg ata 1008
Asp Val Arg Leu Glu Glu Ala Phe Glu Asn Leu Asn Leu Gln Leu Ile
    325                330                335

cga aat cgg tta aaa ata acg gtt gtt aag tta tat gat gta act cgc 1056
Arg Asn Pro Leu Lys Ile Thr Val Val Lys Leu Tyr Asp Val Thr Arg
    340                345                350

agc aat aca atg gcc atg ttc gga aac ctg att acg cat tca att ttt 1104
Ser Asn Thr Met Ala Met Phe Gly Asn Leu Ile Thr His Ser Ile Phe
    355                360                365

tta att cag tat gat att gaa cat ttt 1131
Leu Ile Gln Tyr Asp Ile Glu His Phe
    370                375

```

<210> 28

<211> 377

<212> PRT

<213> Drosophila melanogaster

<400> 28

```

Met Asp Leu Glu Ser Phe Leu Leu Gly Ala Val Tyr Tyr Tyr Gly Leu
  1                5                10                15

Phe Ile Gly Leu Ser Asn Phe Glu Phe Asp Trp Asn Thr Gly Arg Val
    20                25                30

Phe Thr Lys Lys Trp Ser Thr Leu Tyr Ala Ile Ala Leu Asp Ser Cys
    35                40                45

Ile Phe Ala Leu Tyr Ile Tyr His Trp Thr Gly Asn Thr Asn Ile Val
    50                55                60

Asn Ala Ile Phe Gly Arg Ala Asn Met Leu His Glu Tyr Val Val Ala
    65                70                75                80

Ile Leu Thr Gly Leu Arg Ile Arg Cys Lys Met Met Asp Leu Ala Ser
    85                90                95

Lys Val Val Arg Met Tyr Val Ala Arg Pro Gln Val Arg Arg Met Ser
    100                105                110

Arg Trp Gly Ile Leu Thr Lys Phe Ile Phe Gly Ser Ile Thr Asp Gly
    115                120                125

```

Leu Gln Met Ala Met Val Leu Ser Ala Met Gly Ser Val Asp Ser Gln
 130 135 140
 Phe Tyr Leu Gly Leu Gly Leu Gln Tyr Trp Met Phe Val Ile Leu Asn
 145 150 155 160
 Met Ala Met Met Gln Gln His Met Ile Met Leu Phe Val Arg Thr Gln
 165 170 175
 Phe Gln Leu Ile Asn Thr Glu Leu Arg Gln Val Ile Asp Glu Ala Lys
 180 185 190
 Asp Leu Leu Leu Ser Pro Arg His Gln Gly Val Phe Met Thr Lys Cys
 195 200 205
 Cys Ser Leu Ala Asp Gln Ile Glu Asn Ile Ala Arg Ile Gln Ser Gln
 210 215 220
 Leu Gln Thr Ile Met Asn Gln Met Glu Glu Val Phe Gly Ile Gln Gly
 225 230 235 240
 Ala Met Thr Tyr Gly Gly Tyr Tyr Leu Ser Ser Val Gly Thr Cys Tyr
 245 250 255
 Leu Ala Tyr Ser Ile Leu Lys His Gly Tyr Glu Asn Leu Ser Met Thr
 260 265 270
 Leu Ser Thr Val Ile Leu Ala Tyr Ser Trp Cys Phe Phe Tyr Tyr Leu
 275 280 285
 Asp Gly Met Leu Asn Leu Ser Val Met Leu His Val Gln Asp Asp Tyr
 290 295 300
 Trp Glu Met Leu Gln Ile Leu Gly Lys Arg Thr Ile Phe Val Gly Leu
 305 310 315 320
 Asp Val Arg Leu Glu Glu Ala Phe Glu Asn Leu Asn Leu Gln Leu Ile
 325 330 335
 Arg Asn Pro Leu Lys Ile Thr Val Val Lys Leu Tyr Asp Val Thr Arg
 340 345 350
 Ser Asn Thr Met Ala Met Phe Gly Asn Leu Ile Thr His Ser Ile Phe
 355 360 365
 Leu Ile Gln Tyr Asp Ile Glu His Phe
 370 375

<210> 29
 <211> 1044
 <212> DNA
 <213> Drosophila melanogaster

<220>

```

<221> CDS
<222> (1)..(1044)
<223> Coding region GR39D.1

<220>
<221> misc_feature
<222> (232)..(991)
<223> Splice sites before positions 232, 823 and 991

<400> 29
atg ctc tat tcc ttt cat ccg tac ctc aaa tac ttt gcc ctt ttg ggg 48
Met Leu Tyr Ser Phe His Pro Tyr Leu Lys Tyr Phe Ala Leu Leu Gly
  1             5             10             15

ctt gtg cct tgg tca gag agt tgt gcc caa tct aag ttc gta cag aaa 96
Leu Val Pro Trp Ser Glu Ser Cys Ala Gln Ser Lys Phe Val Gln Lys
  20             25             30

gtg tac tca gcg atc cta ata att tta aat gca gtt cat ttc ggg att 144
Val Tyr Ser Ala Ile Leu Ile Ile Leu Asn Ala Val His Phe Gly Ile
  35             40             45

tca att tat ttt cct caa agt gca gaa cta ttt ctt tct ctt atg ggt 192
Ser Ile Tyr Phe Pro Gln Ser Ala Glu Leu Phe Leu Ser Leu Met Gly
  50             55             60

aac ggg att ggt ttc gta act agg att gcg tgt ggt act tac ttg gga 240
Asn Gly Ile Gly Phe Val Thr Arg Ile Ala Cys Gly Thr Tyr Leu Gly
  65             70             75             80

ctt cga ttg cag tgc gaa ctt aaa ata cat gta ggg cgg tta aag tgg 288
Leu Arg Leu Gln Cys Glu Leu Lys Ile His Val Gly Arg Leu Lys Trp
  85             90             95

cag tcg tat gct aaa att cta gca ctc ggt att gga ttt ttg gtg acg 336
Gln Ser Tyr Ala Lys Ile Leu Ala Leu Gly Ile Gly Phe Leu Val Thr
  100            105            110

gta ttg cca tcc atc tat gtt gct cta agc ggt agc ctt ctc tac ttt 384
Val Leu Pro Ser Ile Tyr Val Ala Leu Ser Gly Ser Leu Leu Tyr Phe
  115            120            125

tgg tcc tcc ctc ctg tca atc ctt att ata agg atg caa ttc gta ttg 432
Trp Ser Ser Leu Leu Ser Ile Leu Ile Ile Arg Met Gln Phe Val Leu
  130            135            140

gtt ctt ttg aac gta gaa ttg ctg ggt cac cac gtg agt ctg tta ggc 480
Val Leu Leu Asn Val Glu Leu Leu Gly His His Val Ser Leu Leu Gly
  145            150            155            160

ata cga ctt caa aac gta cta gag tgt cat ctg atg ggc gcc aac tgc 528
Ile Arg Leu Gln Asn Val Leu Glu Cys His Leu Met Gly Ala Asn Cys
  165            170            175

aca ttg gac ggc aat gcc aat cgg ctt tgc tcc cta gag ttt ctg tta 576
Thr Leu Asp Gly Asn Ala Asn Arg Leu Cys Ser Leu Glu Phe Leu Leu

```

180	185	190	
gca ctt aag cag agc cat atg	caa ctg cat tat ctg ttt acc cac ttc	624	
Ala Leu Lys Gln Ser His Met	Gln Leu His Tyr Leu Phe Thr His Phe		
195	200	205	
aac gat ctt ttc ggc tgg tcc ata ctt ggc acc tac gtg gtt ctg ttt	672		
Asn Asp Leu Phe Gly Trp Ser Ile Leu Gly Thr Tyr Val Val Leu Phe			
210	215	220	
tca gat agt acc gtc aat att tac tgg acc cag cag gtt ctg gta gag	720		
Ser Asp Ser Thr Val Asn Ile Tyr Trp Thr Gln Gln Val Leu Val Glu			
225	230	235	
gtt tat gaa tac aaa tac ctt tac gca aca ttt tcc gtt ttt gta ccc	768		
Val Tyr Glu Tyr Lys Tyr Leu Tyr Ala Thr Phe Ser Val Phe Val Pro			
245	250	255	
tca ttc ttc aac atc tta gtg ttt tgc cgt tgt gga gag ttt tgc caa	816		
Ser Phe Phe Asn Ile Leu Val Phe Cys Arg Cys Gly Glu Phe Cys Gln			
260	265	270	
cga cag agt gtc ttg att gga agt tat tta agg aac ttg tcc tgc cat	864		
Arg Gln Ser Val Leu Ile Gly Ser Tyr Leu Arg Asn Leu Ser Cys His			
275	280	285	
ccc tcg att gga aga gaa act tcg tat aaa gat ctg ctg atg gag ttt	912		
Pro Ser Ile Gly Arg Glu Thr Ser Tyr Lys Asp Leu Leu Met Glu Phe			
290	295	300	
att ttg caa gtg gaa caa aat gta ttg gct atc aat gct gag ggc ttt	960		
Ile Leu Gln Val Glu Gln Asn Val Leu Ala Ile Asn Ala Glu Gly Phe			
305	310	315	
atg agc act gat aat tcg ctt cta atg tcg att ttg gct gcg aaa gtt	1008		
Met Ser Thr Asp Asn Ser Leu Leu Met Ser Ile Leu Ala Ala Lys Val			
325	330	335	
acc tat tta atc gtg ctt atg caa ttt agt tca gtc	1044		
Thr Tyr Leu Ile Val Leu Met Gln Phe Ser Ser Val			
340	345		

<210> 30

<211> 348

<212> PRT

<213> Drosophila melanogaster

<400> 30

Met Leu Tyr Ser Phe His Pro Tyr Leu Lys Tyr Phe Ala Leu Leu Gly

1

5

10

15

Leu Val Pro Trp Ser Glu Ser Cys Ala Gln Ser Lys Phe Val Gln Lys

20

25

30

Val Tyr Ser Ala Ile Leu Ile Ile Leu Asn Ala Val His Phe Gly Ile

35					40					45						
Ser	Ile	Tyr	Phe	Pro	Gln	Ser	Ala	Glu	Leu	Phe	Leu	Ser	Leu	Met	Gly	
50					55					60						
Asn	Gly	Ile	Gly	Phe	Val	Thr	Arg	Ile	Ala	Cys	Gly	Thr	Tyr	Leu	Gly	
65					70					75					80	
Leu	Arg	Leu	Gln	Cys	Glu	Leu	Lys	Ile	His	Val	Gly	Arg	Leu	Lys	Trp	
85					90					95						
Gln	Ser	Tyr	Ala	Lys	Ile	Leu	Ala	Leu	Gly	Ile	Gly	Phe	Leu	Val	Thr	
100					105					110						
Val	Leu	Pro	Ser	Ile	Tyr	Val	Ala	Leu	Ser	Gly	Ser	Leu	Leu	Tyr	Phe	
115					120					125						
Trp	Ser	Ser	Leu	Leu	Ser	Ile	Leu	Ile	Ile	Arg	Met	Gln	Phe	Val	Leu	
130					135					140						
Val	Leu	Leu	Asn	Val	Glu	Leu	Leu	Gly	His	His	Val	Ser	Leu	Leu	Gly	
145					150					155					160	
Ile	Arg	Leu	Gln	Asn	Val	Leu	Glu	Cys	His	Leu	Met	Gly	Ala	Asn	Cys	
165					170					175						
Thr	Leu	Asp	Gly	Asn	Ala	Asn	Arg	Leu	Cys	Ser	Leu	Glu	Phe	Leu	Leu	
180					185					190						
Ala	Leu	Lys	Gln	Ser	His	Met	Gln	Leu	His	Tyr	Leu	Phe	Thr	His	Phe	
195					200					205						
Asn	Asp	Leu	Phe	Gly	Trp	Ser	Ile	Leu	Gly	Thr	Tyr	Val	Val	Leu	Phe	
210					215					220						
Ser	Asp	Ser	Thr	Val	Asn	Ile	Tyr	Trp	Thr	Gln	Gln	Val	Leu	Val	Glu	
225					230					235					240	
Val	Tyr	Glu	Tyr	Lys	Tyr	Leu	Tyr	Ala	Thr	Phe	Ser	Val	Phe	Val	Pro	
245					250					255						
Ser	Phe	Phe	Asn	Ile	Leu	Val	Phe	Cys	Arg	Cys	Gly	Glu	Phe	Cys	Gln	
260					265					270						
Arg	Gln	Ser	Val	Leu	Ile	Gly	Ser	Tyr	Leu	Arg	Asn	Leu	Ser	Cys	His	
275					280					285						
Pro	Ser	Ile	Gly	Arg	Glu	Thr	Ser	Tyr	Lys	Asp	Leu	Leu	Met	Glu	Phe	
290					295					300						
Ile	Leu	Gln	Val	Glu	Gln	Asn	Val	Leu	Ala	Ile	Asn	Ala	Glu	Gly	Phe	
305					310					315					320	
Met	Ser	Thr	Asp	Asn	Ser	Leu	Leu	Met	Ser	Ile	Leu	Ala	Ala	Lys	Val	
325					330					335						

Thr Tyr Leu Ile Val Leu Met Gln Phe Ser Ser Val
340 345

<210> 31
<211> 1116
<212> DNA
<213> Drosophila melanogaster

<220>
<221> CDS
<222> (1)..(1116)
<223> Coding region GR39D.2a

<220>
<221> misc_feature
<222> (880)..(1033)
<223> Splice sites before positions 880, 943 and 1033

<400> 31
atg ggc aca aga aat cga aag ctt ctg ttt ttc ctg cac tat cag cgc 48
Met Gly Thr Arg Asn Arg Lys Leu Leu Phe Phe Leu His Tyr Gln Arg
1 5 10 15
tac ttg ggc ctt aca aac ttg gac ttc tca aaa tcg ctg cat att tac 96
Tyr Leu Gly Leu Thr Asn Leu Asp Phe Ser Lys Ser Leu His Ile Tyr
20 25 30
tgg ctc cat ggt act tgg tct tca act gcg att caa att gtg gtc gtt 144
Trp Leu His Gly Thr Trp Ser Ser Thr Ala Ile Gln Ile Val Val Val
35 40 45
ggg gtt ttt atg gcc gct ctg ttg gga gca ctg gcc gaa tct ctt tac 192
Gly Val Phe Met Ala Ala Leu Leu Gly Ala Leu Ala Glu Ser Leu Tyr
50 55 60
tac atg gag act aaa tcc cag act ggc aac acc ttt gac aat gca gtg 240
Tyr Met Glu Thr Lys Ser Gln Thr Gly Asn Thr Phe Asp Asn Ala Val
65 70 75 80
ata cta acc acc tcg gtc act caa ttg cta gcc aat ctt tgg ctc cgt 288
Ile Leu Thr Thr Ser Val Thr Gln Leu Leu Ala Asn Leu Trp Leu Arg
85 90 95
tcg cag cag aaa tca caa gta aac cta ctt caa cga ctt tcg cag gtg 336
Ser Gln Gln Lys Ser Gln Val Asn Leu Leu Gln Arg Leu Ser Gln Val
100 105 110
gta gaa ctt ttg caa ttt gag cct tat gca gtt cca cag ttc cgt tgg 384
Val Glu Leu Leu Gln Phe Glu Pro Tyr Ala Val Pro Gln Phe Arg Trp
115 120 125
tta tac cgt atc tgg cta tta gtg tgc ctt atc tat ggg gcc atg gtg 432
Leu Tyr Arg Ile Trp Leu Leu Val Cys Leu Ile Tyr Gly Ala Met Val
130 135 140

```

acg cat ttt ggc ata aat tgg ttg aca act atg cag atc agc cgt gtc 480
Thr His Phe Gly Ile Asn Trp Leu Thr Thr Met Gln Ile Ser Arg Val
145          150          155          160

ctg act ttg ata gga ttt gta tat agg tgc gtt ttg gcc aac ttt caa 528
Leu Thr Leu Ile Gly Phe Val Tyr Arg Cys Val Leu Ala Asn Phe Gln
165          170          175

ttc acc tgt tat acc ggg atg gtg gtg atc ttg aaa aag ctg ctt caa 576
Phe Thr Cys Tyr Thr Gly Met Val Val Ile Leu Lys Lys Leu Leu Gln
180          185          190

gtt cag gtt aag caa ctg gag cac ttg gtg tcc acc acc acc atc tca 624
Val Gln Val Lys Gln Leu Glu His Leu Val Ser Thr Thr Thr Ile Ser
195          200          205

atg gct gga gta gcc ggt tgt ttg aga acc cac gat gaa atc cta ctg 672
Met Ala Gly Val Ala Gly Cys Leu Arg Thr His Asp Glu Ile Leu Leu
210          215          220

ttg ggt caa aga gaa ctg att gcc gtt tat ggt gga gtt ata cta ttt 720
Leu Gly Gln Arg Glu Leu Ile Ala Val Tyr Gly Gly Val Ile Leu Phe
225          230          235

ctc ttt att tac caa gtc atg cag tgt ata tta ata ttt tac atc agc 768
Leu Phe Ile Tyr Gln Val Met Gln Cys Ile Leu Ile Phe Tyr Ile Ser
245          250          255

aac cta gag ggg ttt cat tca agc aat gac ctg gtt ctc att ttc tgt 816
Asn Leu Glu Gly Phe His Ser Ser Asn Asp Leu Val Leu Ile Phe Cys
260          265          270

tgg ctg gca ccg atg ctc ttc tat ctc atc cta cct tta gtc gtt aat 864
Trp Leu Ala Pro Met Leu Phe Tyr Leu Ile Leu Pro Leu Val Val Asn
275          280          285

gac ata cat aat cag gca aat aaa aca gca aag atg ctg aca aaa gta 912
Asp Ile His Asn Gln Ala Asn Lys Thr Ala Lys Met Leu Thr Lys Val
290          295          300

ccc cga acc ggg act ggg ttg gat aga atg att gaa aaa ttc tta ctc 960
Pro Arg Thr Gly Thr Gly Leu Asp Arg Met Ile Glu Lys Phe Leu Leu
305          310          315

aag aac ctt cga cag aag ccc att tta acc gct tat gga ttt ttc gct 1008
Lys Asn Leu Arg Gln Lys Pro Ile Leu Thr Ala Tyr Gly Phe Phe Ala
325          330          335

ctg gat aaa agt act ttg ttt aag cta ttt act gca atc ttc acg tat 1056
Leu Asp Lys Ser Thr Leu Phe Lys Leu Phe Thr Ala Ile Phe Thr Tyr
340          345          350

atg gtt att ctg gtc caa ttc aag gag atg gaa aat tcc aca aag tct 1104
Met Val Ile Leu Val Gln Phe Lys Glu Met Glu Asn Ser Thr Lys Ser
355          360          365

```

att aat aaa ttt
ile Asn Lys Phe
370

1116

<210> 32
<211> 372
<212> PRT
<213> Drosophila melanogaster

<400> 32
Met Gly Thr Arg Asn Arg Lys Leu Leu Phe Phe Leu His Tyr Gln Arg
1 5 10 15
Tyr Leu Gly Leu Thr Asn Leu Asp Phe Ser Lys Ser Leu His Ile Tyr
20 25 30
Trp Leu His Gly Thr Trp Ser Ser Thr Ala Ile Gln Ile Val Val Val
35 40 45
Gly Val Phe Met Ala Ala Leu Leu Gly Ala Leu Ala Glu Ser Leu Tyr
50 55 60
Tyr Met Glu Thr Lys Ser Gln Thr Gly Asn Thr Phe Asp Asn Ala Val
65 70 75 80
Ile Leu Thr Thr Ser Val Thr Gln Leu Leu Ala Asn Leu Trp Leu Arg
85 90 95
Ser Gln Gln Lys Ser Gln Val Asn Leu Leu Gln Arg Leu Ser Gln Val
100 105 110
Val Glu Leu Leu Gln Phe Glu Pro Tyr Ala Val Pro Gln Phe Arg Trp
115 120 125
Leu Tyr Arg Ile Trp Leu Leu Val Cys Leu Ile Tyr Gly Ala Met Val
130 135 140
Thr His Phe Gly Ile Asn Trp Leu Thr Thr Met Gln Ile Ser Arg Val
145 150 155 160
Leu Thr Leu Ile Gly Phe Val Tyr Arg Cys Val Leu Ala Asn Phe Gln
165 170 175
Phe Thr Cys Tyr Thr Gly Met Val Val Ile Leu Lys Lys Leu Leu Gln
180 185 190
Val Gln Val Lys Gln Leu Glu His Leu Val Ser Thr Thr Thr Ile Ser
195 200 205
Met Ala Gly Val Ala Gly Cys Leu Arg Thr His Asp Glu Ile Leu Leu
210 215 220
Leu Gly Gln Arg Glu Leu Ile Ala Val Tyr Gly Gly Val Ile Leu Phe
225 230 235 240

Leu Phe Ile Tyr Gln Val Met Gln Cys Ile Leu Ile Phe Tyr Ile Ser
 245 250 255
 Asn Leu Glu Gly Phe His Ser Ser Asn Asp Leu Val Leu Ile Phe Cys
 260 265 270
 Trp Leu Ala Pro Met Leu Phe Tyr Leu Ile Leu Pro Leu Val Val Asn
 275 280 285
 Asp Ile His Asn Gln Ala Asn Lys Thr Ala Lys Met Leu Thr Lys Val
 290 295 300
 Pro Arg Thr Gly Thr Gly Leu Asp Arg Met Ile Glu Lys Phe Leu Leu
 305 310 315 320
 Lys Asn Leu Arg Gln Lys Pro Ile Leu Thr Ala Tyr Gly Phe Phe Ala
 325 330 335
 Leu Asp Lys Ser Thr Leu Phe Lys Leu Phe Thr Ala Ile Phe Thr Tyr
 340 345 350
 Met Val Ile Leu Val Gln Phe Lys Glu Met Glu Asn Ser Thr Lys Ser
 355 360 365
 Ile Asn Lys Phe
 370

<210> 33
 <211> 1143
 <212> DNA
 <213> Drosophila melanogaster

<220>
 <221> CDS
 <222> (1)..(1143)
 <223> Coding region GR39D.2b

<220>
 <221> misc_feature
 <222> (907)..(1060)
 <223> Splice sites before positions 907, 970 and 1060

<400> 33
 atg gac ttc caa cca ggt gaa ctc tgt gct tac tac cgc ctt tgc cga 48
 Met Asp Phe Gln Pro Gly Glu Leu Cys Ala Tyr Tyr Arg Leu Cys Arg
 1 5 10 15
 tat cta ggg ata ttc tgt att gat tat aat ccc act aaa aag aaa ttc 96
 Tyr Leu Gly Ile Phe Cys Ile Asp Tyr Asn Pro Thr Lys Lys Lys Phe
 20 25 30
 cga ctg cgg cgc agt gtt ctc tgt tac ata gtt cat ttt gcc ttg caa 144
 Arg Leu Arg Arg Ser Val Leu Cys Tyr Ile Val His Phe Ala Leu Gln
 35 40 45

WO 00/77208

PCT/US00/16211

```

gcc tac tta gct ggt tgc atc tcc gtc atg gtc aca taa tgg tgg tgg ---
Ala Tyr Leu Val Gly Cys Ile Ser Val Met Val Thr Tyr Trp Arg Arg
    50              55              60

tgc ttc aaa agc gag ctt acc acg act gga aac cac ttc gac cgt ctt 240
Cys Phe Lys Ser Glu Leu Thr Thr Thr Gly Asn His Phe Asp Arg Leu
    65              70              75              80

gta atg gta att gcc ctt ggt att ctg gtt gtc cag aat gcg tgg ctc 288
Val Met Val Ile Ala Leu Gly Ile Leu Val Val Gln Asn Ala Trp Leu
    85              90              95

atc tgg ctg caa gcc cca cac cta cga att gtc agg caa ata gag ttt 336
Ile Trp Leu Gln Ala Pro His Leu Arg Ile Val Arg Gln Ile Glu Phe
    100             105             110

tat cga agg aat cac ttg gct aat gtt cga ctg ctc ctc ccc aaa cgt 384
Tyr Arg Arg Asn His Leu Ala Asn Val Arg Leu Leu Leu Pro Lys Arg
    115             120             125

ctg ctt tgg cta att att gca acc aat gtt gtc tac atg gct aac ttc 432
Leu Leu Trp Leu Ile Ile Ala Thr Asn Val Val Tyr Met Ala Asn Phe
    130             135             140

att aag acg tgc ata ttc gaa tgg ctg acg gat gct tct cga ctt ttt 480
Ile Lys Thr Cys Ile Phe Glu Trp Leu Thr Asp Ala Ser Arg Leu Phe
    145             150             155             160

gtc att acc tcc ttg gga ttt cct tta cga tat ctg gtt act agc ttt 528
Val Ile Thr Ser Leu Gly Phe Pro Leu Arg Tyr Leu Val Thr Ser Phe
    165             170             175

aca atg ggc aca tat ttt tgc atg gtg cat att gta cgc ctg gtg ctc 576
Thr Met Gly Thr Tyr Phe Cys Met Val His Ile Val Arg Leu Val Leu
    180             185             190

gac tgg aat cag tcg cag att aac gcg ata ata gat gaa tcg gca gac 624
Asp Trp Asn Gln Ser Gln Ile Asn Ala Ile Ile Asp Glu Ser Ala Asp
    195             200             205

ctc aaa atg act agc ccc aat cgt ctg cgt tta cgt gta tgc ctg gag 672
Leu Lys Met Thr Ser Pro Asn Arg Leu Arg Leu Arg Val Cys Leu Glu
    210             215             220

atg cac gat cgc ctg atg ctg ctc tgc aat gat gag atc agt ctt gtc 720
Met His Asp Arg Leu Met Leu Leu Cys Asn Asp Glu Ile Ser Leu Val
    225             230             235             240

tac ggg ttt ata gcc tgg ctg tct tgg atg ttt gcc tcg ctt gat gta 768
Tyr Gly Phe Ile Ala Trp Leu Ser Trp Met Phe Ala Ser Leu Asp Val
    245             250             255

act ggc gta att tat ctg act atg gtt att caa act aaa aaa tca atc 816
Thr Gly Val Ile Tyr Leu Thr Met Val Ile Gln Thr Lys Lys Ser Ile
    260             265             270

```

```

gtt cta aaa ttg ata aca aac gta gtg tgg ctt tgg cca act ttt atg 864
Val Leu Lys Leu Ile Thr Asn Val Val Trp Leu Ser Pro Thr Phe Met
      275                      280                      285

acg tgc gcc gct agc ttc atg agt aat cgt gtt act att cag gca aat 912
Thr Cys Ala Ala Ser Phe Met Ser Asn Arg Val Thr Ile Gln Ala Asn
      290                      295                      300

aaa aca gca aag atg ctg aca aaa gta ccc cga acc ggg act ggg ttg 960
Lys Thr Ala Lys Met Leu Thr Lys Val Pro Arg Thr Gly Thr Gly Leu
      305                      310                      315                      320

gat aga atg att gaa aaa ttc tta ctc aag aac ctt cga cag aag ccc 1008
Asp Arg Met Ile Glu Lys Phe Leu Leu Lys Asn Leu Arg Gln Lys Pro
      325                      330                      335

att tta acc gct tat gga ttt ttc gct ctg gat aaa agt act ttg ttt 1056
Ile Leu Thr Ala Tyr Gly Phe Phe Ala Leu Asp Lys Ser Thr Leu Phe
      340                      345                      350

aag cta ttt act gca atc ttc acg tat atg gtt att ctg gtc caa ttc 1104
Lys Leu Phe Thr Ala Ile Phe Thr Tyr Met Val Ile Leu Val Gln Phe
      355                      360                      365

aag gag atg gaa aat tcc aca aag tct att aat aaa ttt 1143
Lys Glu Met Glu Asn Ser Thr Lys Ser Ile Asn Lys Phe
      370                      375                      380

```

<210> 34

<211> 381

<212> PRT

<213> *Drosophila melanogaster*

<400> 34

```

Met Asp Phe Gln Pro Gly Glu Leu Cys Ala Tyr Tyr Arg Leu Cys Arg
  1              5              10              15

, Tyr Leu Gly Ile Phe Cys Ile Asp Tyr Asn Pro Thr Lys Lys Lys Phe
      20              25              30

Arg Leu Arg Arg Ser Val Leu Cys Tyr Ile Val His Phe Ala Leu Gln
      35              40              45

Ala Tyr Leu Val Gly Cys Ile Ser Val Met Val Thr Tyr Trp Arg Arg
      50              55              60

Cys Phe Lys Ser Glu Leu Thr Thr Thr Gly Asn His Phe Asp Arg Leu
      65              70              75              80

Val Met Val Ile Ala Leu Gly Ile Leu Val Val Gln Asn Ala Trp Leu
      85              90              95

Ile Trp Leu Gln Ala Pro His Leu Arg Ile Val Arg Gln Ile Glu Phe
      100             105             110

```

Tyr Arg Arg Asn His Leu Ala Asn Val Arg Leu Leu Leu Pro Lys Arg
 115 120 125
 Leu Leu Trp Leu Ile Ile Ala Thr Asn Val Val Tyr Met Ala Asn Phe
 130 135 140
 Ile Lys Thr Cys Ile Phe Glu Trp Leu Thr Asp Ala Ser Arg Leu Phe
 145 150 155 160
 Val Ile Thr Ser Leu Gly Phe Pro Leu Arg Tyr Leu Val Thr Ser Phe
 165 170 175
 Thr Met Gly Thr Tyr Phe Cys Met Val His Ile Val Arg Leu Val Leu
 180 185 190
 Asp Trp Asn Gln Ser Gln Ile Asn Ala Ile Ile Asp Glu Ser Ala Asp
 195 200 205
 Leu Lys Met Thr Ser Pro Asn Arg Leu Arg Leu Arg Val Cys Leu Glu
 210 215 220
 Met His Asp Arg Leu Met Leu Leu Cys Asn Asp Glu Ile Ser Leu Val
 225 230 235 240
 Tyr Gly Phe Ile Ala Trp Leu Ser Trp Met Phe Ala Ser Leu Asp Val
 245 250 255
 Thr Gly Val Ile Tyr Leu Thr Met Val Ile Gln Thr Lys Lys Ser Ile
 260 265 270
 Val Leu Lys Leu Ile Thr Asn Val Val Trp Leu Ser Pro Thr Phe Met
 275 280 285
 Thr Cys Ala Ala Ser Phe Met Ser Asn Arg Val Thr Ile Gln Ala Asn
 290 295 300
 Lys Thr Ala Lys Met Leu Thr Lys Val Pro Arg Thr Gly Thr Gly Leu
 305 310 315 320
 Asp Arg Met Ile Glu Lys Phe Leu Leu Lys Asn Leu Arg Gln Lys Pro
 325 330 335
 Ile Leu Thr Ala Tyr Gly Phe Phe Ala Leu Asp Lys Ser Thr Leu Phe
 340 345 350
 Lys Leu Phe Thr Ala Ile Phe Thr Tyr Met Val Ile Leu Val Gln Phe
 355 360 365
 Lys Glu Met Glu Asn Ser Thr Lys Ser Ile Asn Lys Phe
 370 375 380

<210> 35
 <211> 1143
 <212> DNA

<213> Drosophila melanogaster

<220>

<221> CDS

<222> (1)..(1143)

<223> Coding region GR39D.2c

<220>

<221> misc_feature

<222> (907)..(1060)

<223> Splice sites before positions 907, 970 and 1060

<400> 35

atg	aaa	cgc	aac	gca	ttt	gaa	gag	ttg	agg	gtt	cag	ttg	cga	acc	tta	48
Met	Lys	Arg	Asn	Ala	Phe	Glu	Glu	Leu	Arg	Val	Gln	Leu	Arg	Thr	Leu	
1			5					10					15			

aaa	tgg	ctt	gga	gtc	ctg	cgg	ttc	act	atc	gac	ttt	aat	aag	tgc	tta	96
Lys	Trp	Leu	Gly	Val	Leu	Arg	Phe	Thr	Ile	Asp	Phe	Asn	Lys	Cys	Leu	
	20					25						30				

gta	cgg	gaa	aat	gcg	tct	gag	gag	cgc	agt	gcc	tgg	ctg	tat	ttg	ata	144
Val	Arg	Glu	Asn	Ala	Ser	Glu	Glu	Arg	Ser	Ala	Trp	Leu	Tyr	Leu	Ile	
	35					40					45					

ggt	gta	gtg	gga	att	acg	tgc	agc	ttg	atc	gtc	tac	agc	act	tac	ttt	192
Gly	Val	Val	Gly	Ile	Thr	Cys	Ser	Leu	Ile	Val	Tyr	Ser	Thr	Tyr	Phe	
	50					55				60						

ccc	agt	cac	ttt	att	atg	gga	aaa	cac	aac	acc	acg	ggc	aat	tgc	tat	240
Pro	Ser	His	Phe	Ile	Met	Gly	Lys	His	Asn	Thr	Thr	Gly	Asn	Cys	Tyr	
65					70					75				80		

gcc	cta	atc	aac	atc	agg	tcc	tgt	tcc	att	gta	aca	atg	ctg	att	tac	288
Ala	Leu	Ile	Asn	Ile	Arg	Ser	Cys	Ser	Ile	Val	Thr	Met	Leu	Ile	Tyr	
			85					90					95			

aca	cag	ctc	tac	ata	caa	cgt	ttt	cga	ttc	gtc	gcc	ctc	tta	cag	tcc	336
Thr	Gln	Leu	Tyr	Ile	Gln	Arg	Phe	Arg	Phe	Val	Ala	Leu	Leu	Gln	Ser	
			100				105						110			

att	cta	cgt	ttt	aac	cag	atc	tct	ggg	agt	cac	aga	gaa	gaa	gga	aga	384
Ile	Leu	Arg	Phe	Asn	Gln	Ile	Ser	Gly	Ser	His	Arg	Glu	Glu	Gly	Arg	
	115					120					125					

ttt	gct	ttt	tat	tat	tat	aca	cat	tta	tcc	cta	cta	atc	att	tgt	atg	432
Phe	Ala	Phe	Tyr	Tyr	Tyr	Thr	His	Leu	Ser	Leu	Leu	Ile	Ile	Cys	Met	
	130					135				140						

ctg	aac	tat	gct	tat	ggt	tac	tgg	aca	gca	ggt	gtt	cgc	ttg	acc	aca	480
Leu	Asn	Tyr	Ala	Tyr	Gly	Tyr	Trp	Thr	Ala	Gly	Val	Arg	Leu	Thr	Thr	
145					150					155				160		

att	cct	att	tat	ctg	ctt	caa	tac	gga	ttt	tca	tac	ctc	ttc	ctc	gga	528
Ile	Pro	Ile	Tyr	Leu	Leu	Gln	Tyr	Gly	Phe	Ser	Tyr	Leu	Phe	Leu	Gly	
				165				170					175			

61

```

cag gtg gtg gtc ctg ttt gcc tgc att caa caa ata tta ctt tcg att 576
Gln Val Val Val Leu Phe Ala Cys Ile Gln Gln Ile Leu Leu Ser Ile
      180                      185                      190

tta aag tac tat aac caa gta gtt ctt aaa aat att aaa tca agc aag 624
Leu Lys Tyr Tyr Asn Gln Val Val Leu Lys Asn Ile Lys Ser Ser Lys
      195                      200                      205

gag agt cgc gaa ttc tat tac aac ttt tgc aaa tac aac caa gta ata 672
Glu Ser Arg Glu Phe Tyr Tyr Asn Phe Cys Lys Tyr Asn Gln Val Ile
      210                      215                      220

tgg cta agc tat acc gag atc aac cat tgt ttc ggt ttg cta cta tta 720
Trp Leu Ser Tyr Thr Glu Ile Asn His Cys Phe Gly Leu Leu Leu Leu
      225                      230                      235                      240

ctc gta acc gga tta att ctg cta atc acc cct tct ggg ccg ttc tat 768
Leu Val Thr Gly Leu Ile Leu Leu Ile Thr Pro Ser Gly Pro Phe Tyr
      245                      250                      255

ttg gta tct acc ata ttt gaa gga cga ttt cgt cag aat tgg cag ttc 816
Leu Val Ser Thr Ile Phe Glu Gly Arg Phe Arg Gln Asn Trp Gln Phe
      260                      265                      270

agc tta atg tcg ttc act gcc ata ctt tgg agc tta cca tgg ata gtt 864
Ser Leu Met Ser Phe Thr Ala Ile Leu Trp Ser Leu Pro Trp Ile Val
      275                      280                      285

ttg ctg gtt ttg gca atg ggc agg aat gat gta cag aag gag gca aat 912
Leu Leu Val Leu Ala Met Gly Arg Asn Asp Val Gln Lys Glu Ala Asn
      290                      295                      300

aaa aca gca aag atg ctg aca aaa gta ccc cga acc ggg act ggg ttg 960
Lys Thr Ala Lys Met Leu Thr Lys Val Pro Arg Thr Gly Thr Gly Leu
      305                      310                      315                      320

gat aga atg att gaa aaa ttc tta ctc aag aac ctt cga cag aag ccc 1008
Asp Arg Met Ile Glu Lys Phe Leu Leu Lys Asn Leu Arg Gln Lys Pro
      325                      330                      335

att tta acc gct tat gga ttt ttc gct ctg gat aaa agt act ttg ttt 1056
Ile Leu Thr Ala Tyr Gly Phe Phe Ala Leu Asp Lys Ser Thr Leu Phe
      340                      345                      350

aag cta ttt act gca atc ttc acg tat atg gtt att ctg gtc caa ttc 1104
Lys Leu Phe Thr Ala Ile Phe Thr Tyr Met Val Ile Leu Val Gln Phe
      355                      360                      365

aag gag atg gaa aat tcc aca aag tct att aat aaa ttt 1143
Lys Glu Met Glu Asn Ser Thr Lys Ser Ile Asn Lys Phe
      370                      375                      380

<210> 36
<211> 381
<212> PRT

```

<213> Drosophila melanogaster

<400> 36

```

Met Lys Arg Asn Ala Phe Glu Glu Leu Arg Val Gln Leu Arg Thr Leu
 1              5              10              15

Lys Trp Leu Gly Val Leu Arg Phe Thr Ile Asp Phe Asn Lys Cys Leu
      20              25              30

Val Arg Glu Asn Ala Ser Glu Glu Arg Ser Ala Trp Leu Tyr Leu Ile
      35              40              45

Gly Val Val Gly Ile Thr Cys Ser Leu Ile Val Tyr Ser Thr Tyr Phe
 50              55              60

Pro Ser His Phe Ile Met Gly Lys His Asn Thr Thr Gly Asn Cys Tyr
 65              70              75              80

Ala Leu Ile Asn Ile Arg Ser Cys Ser Ile Val Thr Met Leu Ile Tyr
      85              90              95

Thr Gln Leu Tyr Ile Gln Arg Phe Arg Phe Val Ala Leu Leu Gln Ser
    100              105              110

Ile Leu Arg Phe Asn Gln Ile Ser Gly Ser His Arg Glu Glu Gly Arg
    115              120              125

Phe Ala Phe Tyr Tyr Tyr Thr His Leu Ser Leu Leu Ile Ile Cys Met
    130              135              140

Leu Asn Tyr Ala Tyr Gly Tyr Trp Thr Ala Gly Val Arg Leu Thr Thr
    145              150              155              160

Ile Pro Ile Tyr Leu Leu Gln Tyr Gly Phe Ser Tyr Leu Phe Leu Gly
    165              170              175

Gln Val Val Val Leu Phe Ala Cys Ile Gln Gln Ile Leu Leu Ser Ile
    180              185              190

Leu Lys Tyr Tyr Asn Gln Val Val Leu Lys Asn Ile Lys Ser Ser Lys
    195              200              205

Glu Ser Arg Glu Phe Tyr Tyr Asn Phe Cys Lys Tyr Asn Gln Val Ile
    210              215              220

Trp Leu Ser Tyr Thr Glu Ile Asn His Cys Phe Gly Leu Leu Leu Leu
    225              230              235              240

Leu Val Thr Gly Leu Ile Leu Leu Ile Thr Pro Ser Gly Pro Phe Tyr
    245              250              255

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

Ser Leu Met Ser Phe Thr Ala Ile Leu Trp Ser Leu Pro Trp Ile Val
    275              280              285

```

Leu Leu Val Leu Ala Met Gly Arg Asn Asp Val Gln Lys Glu Ala Asn
 290 295 300
 Lys Thr Ala Lys Met Leu Thr Lys Val Pro Arg Thr Gly Thr Gly Leu
 305 310 315 320
 Asp Arg Met Ile Glu Lys Phe Leu Leu Lys Asn Leu Arg Gln Lys Pro
 325 330 335
 Ile Leu Thr Ala Tyr Gly Phe Phe Ala Leu Asp Lys Ser Thr Leu Phe
 340 345 350
 Lys Leu Phe Thr Ala Ile Phe Thr Tyr Met Val Ile Leu Val Gln Phe
 355 360 365
 Lys Glu Met Glu Asn Ser Thr Lys Ser Ile Asn Lys Phe
 370 375 380

<210> 37
 <211> 1113
 <212> DNA
 <213> *Drosophila melanogaster*

<220>
 <221> CDS
 <222> (1)..(1113)
 <223> Coding region GR39D.2d
 <220>
 <221> misc_feature
 <222> (877)..(1030)
 <223> Splice sites before positions 877, 940 and 1030

<400> 37
 atg tca aaa gtc tgc cgg gac cta cgt atc tat ctc cgt ctt ctt cac 48
 Met Ser Lys Val Cys Arg Asp Leu Arg Ile Tyr Leu Arg Leu Leu His
 1 5 10 15
 atc atg ggc atg atg tgc tgg cat ttc gat tcc gac cac tgt caa cta 96
 Ile Met Gly Met Met Cys Trp His Phe Asp Ser Asp His Cys Gln Leu
 20 25 30
 gtg gcc aca tct gga agc gag cgc tat gcc gtc gtt tat gct ggc tgt 144
 Val Ala Thr Ser Gly Ser Glu Arg Tyr Ala Val Val Tyr Ala Gly Cys
 35 40 45
 att tta gtt tct act acg gct ggc ttt atc ttt gcg ctt tta cat ccg 192
 Ile Leu Val Ser Thr Thr Ala Gly Phe Ile Phe Ala Leu Leu His Pro
 50 55 60
 agt cga ttc cac ata gcc att tat aac cag acg gga aat ttt tac gag 240
 Ser Arg Phe His Ile Ala Ile Tyr Asn Gln Thr Gly Asn Phe Tyr Glu
 65 70 75 80

```

gcc gtc ata ttt cgc agc acg tgt gtg gtg ctt ttc ttg gtt tat gtg 288
Ala Val Ile Phe Arg Ser Thr Cys Val Val Leu Phe Leu Val Tyr Val
      85                      90                      95

ata ttg tat gca tgg cgg cat cgg tac agg gat ttg gtt caa cat ata 336
Ile Leu Tyr Ala Trp Arg His Arg Tyr Arg Asp Leu Val Gln His Ile
      100                    105                    110

ttg cgt ctt aac aga cgc tgt gct agc agt tgc aca aac caa cag ttt 384
Leu Arg Leu Asn Arg Arg Cys Ala Ser Ser Cys Thr Asn Gln Gln Phe
      115                    120                    125

ctg cac aac ata ata ctt tac ggt atg ctt acg atc ctc tgc ttc ggc 432
Leu His Asn Ile Ile Leu Tyr Gly Met Leu Thr Ile Leu Cys Phe Gly
      130                    135                    140

aac tat ctc cac ggc tat act cgc gct gga ctg gcc aca ctt ccg ctg 480
Asn Tyr Leu His Gly Tyr Thr Arg Ala Gly Leu Ala Thr Leu Pro Leu
      145                    150                    155                    160

gcc ctt tgt atg ttg gtc tat atc ttc gcc ttt ttg gtt tta tgc ctt 528
Ala Leu Cys Met Leu Val Tyr Ile Phe Ala Phe Leu Val Leu Cys Leu
      165                    170                    175

ctt ctg atg ttc ttt gtt agt cta aag caa gtc atg acc gcg gga ttg 576
Leu Leu Met Phe Phe Val Ser Leu Lys Gln Val Met Thr Ala Gly Leu
      180                    185                    190

att cac tac aac caa cag ctt tgt cag gcc gat ctg atc tct ggt ctt 624
Ile His Tyr Asn Gln Gln Leu Cys Gln Gly Asp Leu Ile Ser Gly Leu
      195                    200                    205

cga ggc agg cag cag att ctg aaa tta tgc gcc gcc gaa cta aac gag 672
Arg Gly Arg Gln Gln Ile Leu Lys Leu Cys Gly Gly Glu Leu Asn Glu
      210                    215                    220

tgc ttc gcc tta ctt atg ctc ccc att gta gcc ctg gtg ctc ctg atg 720
Cys Phe Gly Leu Leu Met Leu Pro Ile Val Ala Leu Val Leu Leu Met
      225                    230                    235                    240

gca cct tcg ggg ccg ttc ttc tta atc agt acc gtt tta gag gga aaa 768
Ala Pro Ser Gly Pro Phe Phe Leu Ile Ser Thr Val Leu Glu Gly Lys
      245                    250                    255

ttt cga ccc gac gaa tgc ttg atc atg ctt ttg acc tcc tct act tgg 816
Phe Arg Pro Asp Glu Cys Leu Ile Met Leu Leu Thr Ser Ser Thr Trp
      260                    265                    270

gat act ccg tgg atg att atg ttg gtt ctc atg tta cgc act aat ggc 864
Asp Thr Pro Trp Met Ile Met Leu Val Leu Met Leu Arg Thr Asn Gly
      275                    280                    285

att tca gag gaa gca aat aaa aca gca aag atg ctg aca aaa gta ccc 912
Ile Ser Glu Glu Ala Asn Lys Thr Ala Lys Met Leu Thr Lys Val Pro
      290                    295                    300

```

```

cga acc ggg act ggg ttg gat aga atg att gaa aaa ttc tta ctc aag 960
Arg Thr Gly Thr Gly Leu Asp Arg Met Ile Glu Lys Phe Leu Leu Lys
305          310          315          320

aac ctt cga cag aag ccc att tta acc gct tat gga ttt ttc gct ctg 1008
Asn Leu Arg Gln Lys Pro Ile Leu Thr Ala Tyr Gly Phe Phe Ala Leu
          325          330          335

gat aaa agt act ttg ttt aag cta ttt act gca atc ttc acg tat atg 1056
Asp Lys Ser Thr Leu Phe Lys Leu Phe Thr Ala Ile Phe Thr Tyr Met
          340          345          350

gtt att ctg gtc caa ttc aag gag atg gaa aat tcc aca aag tct att 1104
Val Ile Leu Val Gln Phe Lys Glu Met Glu Asn Ser Thr Lys Ser Ile
          355          360          365

aat aaa ttt 1113
Asn Lys Phe
          370

```

```

<210> 38
<211> 371
<212> PRT
<213> Drosophila melanogaster

```

```

<400> 38
Met Ser Lys Val Cys Arg Asp Leu Arg Ile Tyr Leu Arg Leu Leu His
  1          5          10          15

Ile Met Gly Met Met Cys Trp His Phe Asp Ser Asp His Cys Gln Leu
          20          25          30

Val Ala Thr Ser Gly Ser Glu Arg Tyr Ala Val Val Tyr Ala Gly Cys
          35          40          45

Ile Leu Val Ser Thr Thr Ala Gly Phe Ile Phe Ala Leu Leu His Pro
          50          55          60

Ser Arg Phe His Ile Ala Ile Tyr Asn Gln Thr Gly Asn Phe Tyr Glu
          65          70          75          80

Ala Val Ile Phe Arg Ser Thr Cys Val Val Leu Phe Leu Val Tyr Val
          85          90          95

Ile Leu Tyr Ala Trp Arg His Arg Tyr Arg Asp Leu Val Gln His Ile
          100          105          110

Leu Arg Leu Asn Arg Arg Cys Ala Ser Ser Cys Thr Asn Gln Gln Phe
          115          120          125

Leu His Asn Ile Ile Leu Tyr Gly Met Leu Thr Ile Leu Cys Phe Gly
          130          135          140

Asn Tyr Leu His Gly Tyr Thr Arg Ala Gly Leu Ala Thr Leu Pro Leu
          145          150          155          160

```

Ala Leu Cys Met Leu Val Tyr Ile Phe Ala Phe Leu Val Leu Cys Leu
165 170 175

Leu Leu Met Phe Phe Val Ser Leu Lys Gln Val Met Thr Ala Gly Leu
180 185 190

Ile His Tyr Asn Gln Gln Leu Cys Gln Gly Asp Leu Ile Ser Gly Leu
195 200 205

Arg Gly Arg Gln Gln Ile Leu Lys Leu Cys Gly Gly Glu Leu Asn Glu
210 215 220

Cys Phe Gly Leu Leu Met Leu Pro Ile Val Ala Leu Val Leu Leu Met
225 230 235 240

Ala Pro Ser Gly Pro Phe Phe Leu Ile Ser Thr Val Leu Glu Gly Lys
245 250 255

Phe Arg Pro Asp Glu Cys Leu Ile Met Leu Leu Thr Ser Ser Thr Trp
260 265 270

Asp Thr Pro Trp Met Ile Met Leu Val Leu Met Leu Arg Thr Asn Gly
275 280 285

Ile Ser Glu Glu Ala Asn Lys Thr Ala Lys Met Leu Thr Lys Val Pro
290 295 300

Arg Thr Gly Thr Gly Leu Asp Arg Met Ile Glu Lys Phe Leu Leu Lys
305 310 315 320

Asn Leu Arg Gln Lys Pro Ile Leu Thr Ala Tyr Gly Phe Phe Ala Leu
325 330 335

Asp Lys Ser Thr Leu Phe Lys Leu Phe Thr Ala Ile Phe Thr Tyr Met
340 345 350

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

Asn Lys Phe
370

<210> 39
<211> 1137
<212> DNA
<213> *Drosophila melanogaster*

<220>
<221> CDS
<222> (1)..(1137)
<223> Coding region GR43C.1

<220>
<221> misc_feature

<222> (133)..(1078)

<223> Splice sites before positions 133, 301, 442, 727
and 1078

<400> 39

atg	aag	tcg	gct	acg	tcg	aag	gtg	gtg	acg	gcc	ctg	gac	gtg	tct	gtg	48
Met	Lys	Ser	Ala	Thr	Ser	Lys	Val	Val	Thr	Ala	Leu	Asp	Val	Ser	Val	
1				5					10				15			

gtg	gtc	atg	gcc	att	gta	tcc	ggt	gtt	tac	tgt	ggc	ctt	ttc	agc	ctg	96
Val	Val	Met	Ala	Ile	Val	Ser	Gly	Val	Tyr	Cys	Gly	Leu	Phe	Ser	Leu	
		20					25					30				

aac	gac	act	ctc	gag	ctc	aac	gat	cga	cta	aac	aag	atc	gat	aat	acc	144
Asn	Asp	Thr	Leu	Glu	Leu	Asn	Asp	Arg	Leu	Asn	Lys	Ile	Asp	Asn	Thr	
		35				40					45					

cta	aat	gcc	tac	aat	aac	ttt	cga	aga	gat	cga	tgg	cga	gcc	ttg	ggc	192
Leu	Asn	Ala	Tyr	Asn	Asn	Phe	Arg	Arg	Asp	Arg	Trp	Arg	Ala	Leu	Gly	
	50					55					60					

atg	gct	gcc	gtc	tct	ctg	ctg	gcc	att	tcg	att	ctg	gtc	ggc	ctg	gat	240
Met	Ala	Ala	Val	Ser	Leu	Leu	Ala	Ile	Ser	Ile	Leu	Val	Gly	Leu	Asp	
	65				70				75					80		

gta	ggg	acg	tgg	atg	cgc	att	gcc	cag	gat	atg	aac	ata	gct	cag	tcg	288
Val	Gly	Thr	Trp	Met	Arg	Ile	Ala	Gln	Asp	Met	Asn	Ile	Ala	Gln	Ser	
		85						90					95			

gat	acg	gag	ctc	aat	gtg	cac	tgg	tac	att	cca	ttc	tat	agt	ctt	tac	336
Asp	Thr	Glu	Leu	Asn	Val	His	Trp	Tyr	Ile	Pro	Phe	Tyr	Ser	Leu	Tyr	
		100				105						110				

ttc	atc	ctc	acg	ggc	ttg	cag	gtt	aac	att	gcc	aac	acg	gcc	tat	gga	384
Phe	Ile	Leu	Thr	Gly	Leu	Gln	Val	Asn	Ile	Ala	Asn	Thr	Ala	Tyr	Gly	
		115				120						125				

cta	gga	agg	cgc	ttt	ggc	cgt	ctg	aac	cga	atg	ctt	tcg	agc	agc	ttt	432
Leu	Gly	Arg	Arg	Phe	Gly	Arg	Leu	Asn	Arg	Met	Leu	Ser	Ser	Ser	Phe	
	130				135					140						

ctt	gca	gag	aac	aat	gca	acc	agt	gcc	atc	aag	cca	caa	aag	gtc	agc	480
Leu	Ala	Glu	Asn	Asn	Ala	Thr	Ser	Ala	Ile	Lys	Pro	Gln	Lys	Val	Ser	
	145			150				155					160			

acc	gtg	aag	aac	gtt	agc	gta	aat	cgg	cca	gcg	atg	cca	tcg	gca	ctc	528
Thr	Val	Lys	Asn	Val	Ser	Val	Asn	Arg	Pro	Ala	Met	Pro	Ser	Ala	Leu	
			165					170					175			

cac	gcg	agc	ctg	acg	aag	ctg	aac	ggc	gaa	acc	ttg	ccc	agt	gaa	gcc	576
His	Ala	Ser	Leu	Thr	Lys	Leu	Asn	Gly	Glu	Thr	Leu	Pro	Ser	Glu	Ala	
		180				185					190					

gca	ggt	gac	aag	gca	gcc	gca	cgc	agt	cta	atc	ctc	aac	gtg	gag	tta	624
Ala	Gly	Asp	Lys	Ala	Ala	Ala	Arg	Ser	Leu	Ile	Leu	Asn	Val	Glu	Leu	
	195					200						205				

```

ctc aaa ttg ggt tat ttt cca gcc aag aac aag ggg ttg ctc ctc aaa 672
Leu Lys Leu Gly Tyr Phe Pro Ala Lys Asn Lys Gly Leu Leu Leu Lys
210 215 220

tcc ctg gcg gac agt cac gag tcg ctg gga aaa tgt gtc cac ctc ctg 720
Ser Leu Ala Asp Ser His Glu Ser Leu Gly Lys Cys Val His Leu Leu
225 230 235 240

tcc aac tcc ttc ggc att gcg gtt ctc ttc atc ctg gtg tcc tgt ctg 768
Ser Asn Ser Phe Gly Ile Ala Val Leu Phe Ile Leu Val Ser Cys Leu
245 250 255

ctg cac ctc gtg gct acg gcc tac ttc ctc ttt ctg gaa ctg ctc agc 816
Leu His Leu Val Ala Thr Ala Tyr Phe Leu Phe Leu Glu Leu Leu Ser
260 265 270

aag cgg gat aac ggc tac ctg tgg gtg cag atg ctc tgg att tgc ttt 864
Lys Arg Asp Asn Gly Tyr Leu Trp Val Gln Met Leu Trp Ile Cys Phe
275 280 285

cac ttc ctg cga ctg ctc atg gtg gtg gag cgg tgc cac ttg gct gcc 912
His Phe Leu Arg Leu Leu Met Val Val Glu Pro Cys His Leu Ala Ala
290 295 300

cga gag tcc cgt aaa acg atc cag att gtt tgc gaa atc gag cgg aaa 960
Arg Glu Ser Arg Lys Thr Ile Gln Ile Val Cys Glu Ile Glu Arg Lys
305 310 315 320

gtg cac gag ccc att ctc gcg gag gca gtt aag aag ttt tgg cag cag 1008
Val His Glu Pro Ile Leu Ala Glu Ala Val Lys Lys Phe Trp Gln Gln
325 330 335

tta ctc gtc gta gat gct gac ttc tcc gcc tgc gga ttg tgc cgc gtg 1056
Leu Leu Val Val Asp Ala Asp Phe Ser Ala Cys Gly Leu Cys Arg Val
340 345 350

aac cgc acc att ctg aca tcg ttt gca tcc gcc ata gcc acc tat ctc 1104
Asn Arg Thr Ile Leu Thr Ser Phe Ala Ser Ala Ile Ala Thr Tyr Leu
355 360 365

gtg att ctc att cag ttt caa cgg acc aat ggc 1137
Val Ile Leu Ile Gln Phe Gln Arg Thr Asn Gly
370 375

```

```

<210> 40
<211> 379
<212> PRT
<213> Drosophila melanogaster

```

```

<400> 40
Met Lys Ser Ala Thr Ser Lys Val Val Thr Ala Leu Asp Val Ser Val
1 5 10 15
Val Val Met Ala Ile Val Ser Gly Val Tyr Cys Gly Leu Phe Ser Leu
20 25 30

```

Asn Asp Thr Leu Glu Leu Asn Asp Arg Leu Asn Lys Ile Asp Asn Thr
 35 40 45
 Leu Asn Ala Tyr Asn Asn Phe Arg Arg Asp Arg Trp Arg Ala Leu Gly
 50 55 60
 Met Ala Ala Val Ser Leu Leu Ala Ile Ser Ile Leu Val Gly Leu Asp
 65 70 75 80
 Val Gly Thr Trp Met Arg Ile Ala Gln Asp Met Asn Ile Ala Gln Ser
 85 90 95
 Asp Thr Glu Leu Asn Val His Trp Tyr Ile Pro Phe Tyr Ser Leu Tyr
 100 105 110
 Phe Ile Leu Thr Gly Leu Gln Val Asn Ile Ala Asn Thr Ala Tyr Gly
 115 120 125
 Leu Gly Arg Arg Phe Gly Arg Leu Asn Arg Met Leu Ser Ser Ser Phe
 130 135 140
 Leu Ala Glu Asn Asn Ala Thr Ser Ala Ile Lys Pro Gln Lys Val Ser
 145 150 155 160
 Thr Val Lys Asn Val Ser Val Asn Arg Pro Ala Met Pro Ser Ala Leu
 165 170 175
 His Ala Ser Leu Thr Lys Leu Asn Gly Glu Thr Leu Pro Ser Glu Ala
 180 185 190
 Ala Gly Asp Lys Ala Ala Ala Arg Ser Leu Ile Leu Asn Val Glu Leu
 195 200 205
 Leu Lys Leu Gly Tyr Phe Pro Ala Lys Asn Lys Gly Leu Leu Lys
 210 215 220
 Ser Leu Ala Asp Ser His Glu Ser Leu Gly Lys Cys Val His Leu Leu
 225 230 235 240
 Ser Asn Ser Phe Gly Ile Ala Val Leu Phe Ile Leu Val Ser Cys Leu
 245 250 255
 Leu His Leu Val Ala Thr Ala Tyr Phe Leu Phe Leu Glu Leu Leu Ser
 260 265 270
 Lys Arg Asp Asn Gly Tyr Leu Trp Val Gln Met Leu Trp Ile Cys Phe
 275 280 285
 His Phe Leu Arg Leu Leu Met Val Val Glu Pro Cys His Leu Ala Ala
 290 295 300
 Arg Glu Ser Arg Lys Thr Ile Gln Ile Val Cys Glu Ile Glu Arg Lys
 305 310 315 320
 Val His Glu Pro Ile Leu Ala Glu Ala Val Lys Lys Phe Trp Gln Gln
 325 330 335

Leu Leu Val Val Asp Ala Asp Phe Ser Ala Cys Gly Leu Cys Arg Val
 340 345 350

Asn Arg Thr Ile Leu Thr Ser Phe Ala Ser Ala Ile Ala Thr Tyr Leu
 355 360 365

Val Ile Leu Ile Gln Phe Gln Arg Thr Asn Gly
 370 375

<210> 41
 <211> 1083
 <212> DNA
 <213> *Drosophila melanogaster*

<220>
 <221> CDS
 <222> (1)..(1083)
 <223> Coding region GR47A.1

<220>
 <221> misc_feature
 <222> (928)
 <223> Splice site

<400> 41
 atg gcc ttt acc agc tcg cag tta tgc agt ttg ctt acc aag ttt acg 48
 Met Ala Phe Thr Ser Ser Gln Leu Cys Ser Leu Leu Thr Lys Phe Thr
 1 5 10 15

gcg ctc aat gga cta aat acc tac tat ttc gat aca aag acg aat gcg 96
 Ala Leu Asn Gly Leu Asn Thr Tyr Tyr Phe Asp Thr Lys Thr Asn Ala
 20 25 30

ttc cga gtc tcc tcg aag ctc aag ata tac tgc gct atc cat cat gcc 144
 Phe Arg Val Ser Ser Lys Leu Lys Ile Tyr Cys Ala Ile His His Ala
 35 40 45

ctc tgt gtg ctg gca ctg gca cac atg tcc tat agc act gct agt aat 192
 Leu Cys Val Leu Ala Leu Ala His Met Ser Tyr Ser Thr Ala Ser Asn
 50 55 60

ttg cgt gtg agt gtg act gtt ctg acg att ggt gga acc atg gcc tgc 240
 Leu Arg Val Ser Val Thr Val Leu Thr Ile Gly Gly Thr Met Ala Cys
 65 70 75 80

tgt gtg aag tcg tgc tgg gaa aag gcg cag ggc att cga aac ttg gcc 288
 Cys Val Lys Ser Cys Trp Glu Lys Ala Gln Gly Ile Arg Asn Leu Ala
 85 90 95

cgc gga ctt gtc aca atg gag cag aag tat ttc gct ggc aga ccc agc 336
 Arg Gly Leu Val Thr Met Glu Gln Lys Tyr Phe Ala Gly Arg Pro Ser
 100 105 110

ggg ctg cta ttg aag tgc cga tac tac att aaa att act ttt ggt tcg 384

Gly	Leu	Leu	Leu	Lys	Cys	Arg	Tyr	Tyr	Ile	Lys	Ile	Thr	Phe	Gly	Ser		
	115						120					125					
ata	acc	ctg	ctg	cgt	atc	cat	ttg	atc	caa	ccg	atc	tac	atg	agg	aga	432	
Ile	Thr	Leu	Leu	Arg	Ile	His	Leu	Ile	Gln	Pro	Ile	Tyr	Met	Arg	Arg		
	130				135				140								
ctg	cta	cca	tcg	cag	ttt	tat	cta	aat	gtg	gga	gcg	tat	tgg	cta	ctc	480	
Leu	Leu	Pro	Ser	Gln	Phe	Tyr	Leu	Asn	Val	Gly	Ala	Tyr	Trp	Leu	Leu		
	145			150				155						160			
tac	aac	atg	ctt	ttg	gca	gct	gtg	ctg	ggc	ttt	tac	ttt	ctg	ctc	tgg	528	
Tyr	Asn	Met	Leu	Leu	Ala	Ala	Val	Leu	Gly	Phe	Tyr	Phe	Leu	Leu	Trp		
			165					170					175				
gaa	atg	tgt	cgc	ata	caa	aag	ctc	ata	aat	gat	caa	atg	aca	ctc	att	576	
Glu	Met	Cys	Arg	Ile	Gln	Lys	Leu	Ile	Asn	Asp	Gln	Met	Thr	Leu	Ile		
			180				185					190					
ttg	gcc	agg	tcg	ggc	cag	aga	aat	cgg	ctg	aaa	aag	atg	cag	cac	tgc	624	
Leu	Ala	Arg	Ser	Gly	Gln	Arg	Asn	Arg	Leu	Lys	Lys	Met	Gln	His	Cys		
		195				200						205					
ctt	aga	ctc	tac	tcc	aag	ctt	cta	ctg	cta	tgt	gac	cag	ttc	aat	agt	672	
Leu	Arg	Leu	Tyr	Ser	Lys	Leu	Leu	Leu	Leu	Cys	Asp	Gln	Phe	Asn	Ser		
	210					215					220						
cag	ctt	gga	cat	gtt	gcg	att	tgg	gtt	ttg	gct	tgc	aaa	agc	tgg	tgc	720	
Gln	Leu	Gly	His	Val	Ala	Ile	Trp	Val	Leu	Ala	Cys	Lys	Ser	Trp	Cys		
	225			230				235						240			
cag	atc	acc	ttt	ggc	tac	gaa	ata	ttc	cag	atg	gtg	gct	gcc	cca	aag	768	
Gln	Ile	Thr	Phe	Gly	Tyr	Glu	Ile	Phe	Gln	Met	Val	Ala	Ala	Pro	Lys		
			245					250					255				
tca	atc	gat	tta	act	atg	tcc	atg	aga	gtg	ttt	gta	att	ttc	acc	tac	816	
Ser	Ile	Asp	Leu	Thr	Met	Ser	Met	Arg	Val	Phe	Val	Ile	Phe	Thr	Tyr		
			260					265					270				
att	ttt	gat	gcc	atg	aac	cta	ttt	ctt	gga	aca	gac	att	tcc	gag	ttg	864	
Ile	Phe	Asp	Ala	Met	Asn	Leu	Phe	Leu	Gly	Thr	Asp	Ile	Ser	Glu	Leu		
		275				280						285					
ttc	agt	act	ttt	aga	gcg	gac	tct	caa	cga	att	ttg	agg	gag	acc	agt	912	
Phe	Ser	Thr	Phe	Arg	Ala	Asp	Ser	Gln	Arg	Ile	Leu	Arg	Glu	Thr	Ser		
		290				295					300						
cgc	ttg	gat	cgt	ttg	ctc	tcc	atg	ttc	gcc	ttg	aag	ttg	gcc	ctc	cat	960	
Arg	Leu	Asp	Arg	Leu	Leu	Ser	Met	Phe	Ala	Leu	Lys	Leu	Ala	Leu	His		
	305				310				315					320			
ccc	aag	cgg	gtc	gtc	tta	ctc	aac	gtg	ttt	acc	ttt	gac	cga	aag	ctg	1008	
Pro	Lys	Arg	Val	Val	Leu	Leu	Asn	Val	Phe	Thr	Phe	Asp	Arg	Lys	Leu		
			325					330					335				
act	cta	acg	ctt	ctg	gcc	aag	tcg	acg	ttg	tac	aca	ata	tgc	tgt	ctg	1056	

Thr Leu Thr Leu Leu Ala Lys Ser Thr Leu Tyr Thr Ile Cys Cys Leu
 340 345 350

caa aac gac tac aac aaa ctc aag gcg 1083
 Gln Asn Asp Tyr Asn Lys Leu Lys Ala
 355 360

<210> 42
 <211> 361
 <212> PRT
 <213> Drosophila melanogaster

<400> 42
 Met Ala Phe Thr Ser Ser Gln Leu Cys Ser Leu Leu Thr Lys Phe Thr
 1 5 10 15
 Ala Leu Asn Gly Leu Asn Thr Tyr Tyr Phe Asp Thr Lys Thr Asn Ala
 20 25 30
 Phe Arg Val Ser Ser Lys Leu Lys Ile Tyr Cys Ala Ile His His Ala
 35 40 45
 Leu Cys Val Leu Ala Leu Ala His Met Ser Tyr Ser Thr Ala Ser Asn
 50 55 60
 Leu Arg Val Ser Val Thr Val Leu Thr Ile Gly Gly Thr Met Ala Cys
 65 70 75 80
 Cys Val Lys Ser Cys Trp Glu Lys Ala Gln Gly Ile Arg Asn Leu Ala
 85 90 95
 Arg Gly Leu Val Thr Met Glu Gln Lys Tyr Phe Ala Gly Arg Pro Ser
 100 105 110
 Gly Leu Leu Leu Lys Cys Arg Tyr Tyr Ile Lys Ile Thr Phe Gly Ser
 115 120 125
 Ile Thr Leu Leu Arg Ile His Leu Ile Gln Pro Ile Tyr Met Arg Arg
 130 135 140
 Leu Leu Pro Ser Gln Phe Tyr Leu Asn Val Gly Ala Tyr Trp Leu Leu
 145 150 155 160
 Tyr Asn Met Leu Leu Ala Ala Val Leu Gly Phe Tyr Phe Leu Leu Trp
 165 170 175
 Glu Met Cys Arg Ile Gln Lys Leu Ile Asn Asp Gln Met Thr Leu Ile
 180 185 190
 Leu Ala Arg Ser Gly Gln Arg Asn Arg Leu Lys Lys Met Gln His Cys
 195 200 205
 Leu Arg Leu Tyr Ser Lys Leu Leu Leu Leu Cys Asp Gln Phe Asn Ser
 210 215 220

Gln Leu Gly His Val Ala Ile Trp Val Leu Ala Cys Lys Ser Trp Cys
 225 230 235 240

Gln Ile Thr Phe Gly Tyr Glu Ile Phe Gln Met Val Ala Ala Pro Lys
 245 250 255

Ser Ile Asp Leu Thr Met Ser Met Arg Val Phe Val Ile Phe Thr Tyr
 260 265 270

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

Phe Ser Thr Phe Arg Ala Asp Ser Gln Arg Ile Leu Arg Glu Thr Ser
 290 295 300

Arg Leu Asp Arg Leu Leu Ser Met Phe Ala Leu Lys Leu Ala Leu His
 305 310 315 320

Pro Lys Arg Val Val Leu Leu Asn Val Phe Thr Phe Asp Arg Lys Leu
 325 330 335

Thr Leu Thr Leu Leu Ala Lys Ser Thr Leu Tyr Thr Ile Cys Cys Leu
 340 345 350

Gln Asn Asp Tyr Asn Lys Leu Lys Ala
 355 360

<210> 43
 <211> 1206
 <212> DNA
 <213> *Drosophila melanogaster*

<220>
 <221> CDS
 <222> (1)..(1206)
 <223> Coding region GR47F.1

<400> 43
 atg caa cgc gac gat ggc ttt gtc tac tgt tac gga aac ttg tac agc 48
 Met Gln Arg Asp Asp Gly Phe Val Tyr Cys Tyr Gly Asn Leu Tyr Ser
 1 5 10 15

ctc ctg ctc tac tgg gga ttg gtc aca atc cgg gtg aga agt cct gat 96
 Leu Leu Leu Tyr Trp Gly Leu Val Thr Ile Arg Val Arg Ser Pro Asp
 20 25 30

cgt gga ggt gcg ttt tca aac agg tgg act gtg tgc tat gcc ctt ttc 144
 Arg Gly Gly Ala Phe Ser Asn Arg Trp Thr Val Cys Tyr Ala Leu Phe
 35 40 45

acg cgc tcc ttt atg gtt atc tgc ttt atg gct acc gtg atg act aag 192
 Thr Arg Ser Phe Met Val Ile Cys Phe Met Ala Thr Val Met Thr Lys
 50 55 60

```

ctg agg gat ccc gag atg tcc gct gcc atg ttt gga cac ctt agc ccg 240
Leu Arg Asp Pro Glu Met Ser Ala Ala Met Phe Gly His Leu Ser Pro
65 70 75 80

ttg gtc aag gca ata ttc acc tgg gag tgc ctc agt tgc tcg gtc acc 288
Leu Val Lys Ala Ile Phe Thr Trp Glu Cys Leu Ser Cys Ser Val Thr
85 90 95

tac att gag tac tgc ctt tcg ctg gac ttg cag aag gac aga cac ctg 336
Tyr Ile Glu Tyr Cys Leu Ser Leu Asp Leu Gln Lys Asp Arg His Leu
100 105 110

aaa ctc gtg gct agg atg caa gag ttc gat cgt tcc gtt ttg atg gtg 384
Lys Leu Val Ala Arg Met Gln Glu Phe Asp Arg Ser Val Leu Met Val
115 120 125

ttc cct cat gtt cag tgg aac tat cga cgc gct cgc ttg aag tat tgg 432
Phe Pro His Val Gln Trp Asn Tyr Arg Arg Ala Arg Leu Lys Tyr Trp
130 135 140

tat ggc act gtg att gtc ggc ttt tgc ttt ttc tcc ttt tcc atc tca 480
Tyr Gly Thr Val Ile Val Gly Phe Cys Phe Phe Ser Phe Ser Ile Ser
145 150 155 160

ctg atc ttc gat acg acg cgt tgc aca tgc ggc ata ccg tcc acc ttg 528
Leu Ile Phe Asp Thr Thr Arg Cys Thr Cys Gly Ile Pro Ser Thr Leu
165 170 175

ctg atg gcc ttc acc tac acc ttg ctg acc agt tcc gta gga ttg ctg 576
Leu Met Ala Phe Thr Tyr Thr Leu Leu Thr Ser Ser Val Gly Leu Leu
180 185 190

ggc ttc gtg cac ata ggt att atg gac ttt ata agg gtt cgt ctg cgt 624
Gly Phe Val His Ile Gly Ile Met Asp Phe Ile Arg Val Arg Leu Arg
195 200 205

ctg gtg cag cag ctg ctt cac caa ctt tac caa gcg gat gac agt tca 672
Leu Val Gln Gln Leu Leu His Gln Leu Tyr Gln Ala Asp Asp Ser Ser
210 215 220

gag gtg cac gaa agg att gct tat ctt ttc gaa atg tcc aag cgc tgt 720
Glu Val His Glu Arg Ile Ala Tyr Leu Phe Glu Met Ser Lys Arg Cys
225 230 235 240

tcg ttt ctc ctg gcg gag tta aac gga gtt ttt gga ttc gcc gca gct 768
Ser Phe Leu Leu Ala Glu Leu Asn Gly Val Phe Gly Phe Ala Ala Ala
245 250 255

gct ggt atc ttt tat gat ttc act atc atg acc tgc ttc gtg tac gtt 816
Ala Gly Ile Phe Tyr Asp Phe Thr Ile Met Thr Cys Phe Val Tyr Val
260 265 270

atc tgc caa aag ctg ctg gag aga gag ccc tgg gat ccc gag tac gtg 864
Ile Cys Gln Lys Leu Leu Glu Arg Glu Pro Trp Asp Pro Glu Tyr Val
275 280 285

```

WO 00/77208

PCT/US00/16211

```

tac atg ttg ctg cac gtg gcc ata cac acc tac aaa gtg gtc att act 912
Tyr Met Leu Leu His Val Ala Ile His Thr Tyr Lys Val Val Ile Thr
290 295 300

agc acc tat ggg tat tta ctt ctg cga gag gtg ggt gac ctt att tcc 960
Ser Thr Tyr Gly Tyr Leu Leu Leu Arg Glu Val Gly Asp Leu Ile Ser
305 310 315 320

gat gtg gct cgc cgg aaa acg gag gac ttt cag cat tgg cgc atg cac 1008
Asp Val Ala Arg Arg Lys Thr Glu Asp Phe Gln His Trp Arg Met His
325 330 335

aat cgg caa gcg gca atg gtg ggc agc aca acc ttg cta agt gtt tct 1056
Asn Arg Gln Ala Ala Met Val Gly Ser Thr Thr Leu Leu Ser Val Ser
340 345 350

acc att tac ctg gtg tac aac gga atg gcc aac tat gtg att atc ctg 1104
Thr Ile Tyr Leu Val Tyr Asn Gly Met Ala Asn Tyr Val Ile Ile Leu
355 360 365

gtg cag ctg cta ttt caa caa cag caa atc aaa gat cac cag ttg aca 1152
Val Gln Leu Leu Phe Gln Gln Gln Ile Lys Asp His Gln Leu Thr
370 375 380

tcg ggc aaa gat gtg gac att gtg ggc cca atg gga ccc atc act cac 1200
Ser Gly Lys Asp Val Asp Ile Val Gly Pro Met Gly Pro Ile Thr His
385 390 395 400

atg gat 1206
Met Asp

```

<210> 44
 <211> 402
 <212> PRT
 <213> *Drosophila melanogaster*

```

<400> 44
Met Gln Arg Asp Asp Gly Phe Val Tyr Cys Tyr Gly Asn Leu Tyr Ser
1 5 10 15

Leu Leu Leu Tyr Trp Gly Leu Val Thr Ile Arg Val Arg Ser Pro Asp
20 25 30

Arg Gly Gly Ala Phe Ser Asn Arg Trp Thr Val Cys Tyr Ala Leu Phe
35 40 45

Thr Arg Ser Phe Met Val Ile Cys Phe Met Ala Thr Val Met Thr Lys
50 55 60

Leu Arg Asp Pro Glu Met Ser Ala Ala Met Phe Gly His Leu Ser Pro
65 70 75 80

Leu Val Lys Ala Ile Phe Thr Trp Glu Cys Leu Ser Cys Ser Val Thr
85 90 95

```

76

Tyr Ile Glu Tyr Cys Leu Ser Leu Asp Leu Gln Lys Asp Arg His Leu
 100 105 110
 Lys Leu Val Ala Arg Met Gln Glu Phe Asp Arg Ser Val Leu Met Val
 115 120 125
 Phe Pro His Val Gln Trp Asn Tyr Arg Arg Ala Arg Leu Lys Tyr Trp
 130 135 140
 Tyr Gly Thr Val Ile Val Gly Phe Cys Phe Phe Ser Phe Ser Ile Ser
 145 150 155 160
 Leu Ile Phe Asp Thr Thr Arg Cys Thr Cys Gly Ile Pro Ser Thr Leu
 165 170 175
 Leu Met Ala Phe Thr Tyr Thr Leu Leu Thr Ser Ser Val Gly Leu Leu
 180 185 190
 Gly Phe Val His Ile Gly Ile Met Asp Phe Ile Arg Val Arg Leu Arg
 195 200 205
 Leu Val Gln Gln Leu Leu His Gln Leu Tyr Gln Ala Asp Asp Ser Ser
 210 215 220
 Glu Val His Glu Arg Ile Ala Tyr Leu Phe Glu Met Ser Lys Arg Cys
 225 230 235 240
 Ser Phe Leu Leu Ala Glu Leu Asn Gly Val Phe Gly Phe Ala Ala Ala
 245 250 255
 Ala Gly Ile Phe Tyr Asp Phe Thr Ile Met Thr Cys Phe Val Tyr Val
 260 265 270
 Ile Cys Gln Lys Leu Leu Glu Arg Glu Pro Trp Asp Pro Glu Tyr Val
 275 280 285
 Tyr Met Leu Leu His Val Ala Ile His Thr Tyr Lys Val Val Ile Thr
 290 295 300
 Ser Thr Tyr Gly Tyr Leu Leu Leu Arg Glu Val Gly Asp Leu Ile Ser
 305 310 315 320
 Asp Val Ala Arg Arg Lys Thr Glu Asp Phe Gln His Trp Arg Met His
 325 330 335
 Asn Arg Gln Ala Ala Met Val Gly Ser Thr Thr Leu Leu Ser Val Ser
 340 345 350
 Thr Ile Tyr Leu Val Tyr Asn Gly Met Ala Asn Tyr Val Ile Ile Leu
 355 360 365
 Val Gln Leu Leu Phe Gln Gln Gln Gln Ile Lys Asp His Gln Leu Thr
 370 375 380
 Ser Gly Lys Asp Val Asp Ile Val Gly Pro Met Gly Pro Ile Thr His
 385 390 395 400

Met Asp

<210> 45
 <211> 1251
 <212> DNA
 <213> *Drosophila melanogaster*

<220>
 <221> CDS
 <222> (1)..(1248)
 <223> Coding region 57B.1

<400> 45
 atg gcc gtg ttg tac ttc ttc cgc gaa ccg gaa acg gtt ttc gat tgt 48
 Met Ala Val Leu Tyr Phe Phe Arg Glu Pro Glu Thr Val Phe Asp Cys
 1 5 10 15

gcc gca ttc att tgc atc ctg cag ttc ctt atg ggc tgc aat ggg ttc 96
 Ala Ala Phe Ile Cys Ile Leu Gln Phe Leu Met Gly Cys Asn Gly Phe
 20 25 30

ggc att cgt cga tgc acc ttt agg atc agt tgg gca agc aga atc tat 144
 Gly Ile Arg Arg Ser Thr Phe Arg Ile Ser Trp Ala Ser Arg Ile Tyr
 35 40 45

tcc atg tgc gtg gcc att gcg gca ttt tgc tgt ctc ttt ggt tcc ctg 192
 Ser Met Ser Val Ala Ile Ala Ala Phe Cys Cys Leu Phe Gly Ser Leu
 50 55 60

agt gtc ctt ttg gca gag gaa gat ata aga gaa cgc ctg gca aag gcc 240
 Ser Val Leu Leu Ala Glu Glu Asp Ile Arg Glu Arg Leu Ala Lys Ala
 65 70 75 80

gat aat cta gtg ctg agc ata tca gcc ttg gaa ctg ctg atg tcc aca 288
 Asp Asn Leu Val Leu Ser Ile Ser Ala Leu Glu Leu Leu Met Ser Thr
 85 90 95

ttg gtc ttt ggg gtg act gtg ata tcc ttg caa gta ttt gca cgc egg 336
 Leu Val Phe Gly Val Thr Val Ile Ser Leu Gln Val Phe Ala Arg Arg
 100 105 110

cac ttg gga atc tat caa aga ttg gct gcc ttg gat gcc agg ctg atg 384
 His Leu Gly Ile Tyr Gln Arg Leu Ala Ala Leu Asp Ala Arg Leu Met
 115 120 125

agc gac ttc gga gct aat ctg aac tat aga aaa atg ctg aga aag aac 432
 Ser Asp Phe Gly Ala Asn Leu Asn Tyr Arg Lys Met Leu Arg Lys Asn
 130 135 140

ata gca gtg ctg gga ata gtg acc act att tac ctg atg gcc atc aac 480
 Ile Ala Val Leu Gly Ile Val Thr Thr Ile Tyr Leu Met Ala Ile Asn
 145 150 155 160

```

agt gcc gcc gtc caa gtg gcc agc ggt cat cga gct ctg ttc ctt cta 528
Ser Ala Ala Val Gln Val Ala Ser Gly His Arg Ala Leu Phe Leu Leu
165 170 175

ttc gct ctg tgc tat aca att gtc acc gga gga ccc cac ttt acg ggc 576
Phe Ala Leu Cys Tyr Thr Ile Val Thr Gly Gly Pro His Phe Thr Gly
180 185 190

tat gtg cac atg acc ctg gcc gag atg ttg ggc ata cgt ttc cgt ttg 624
Tyr Val His Met Thr Leu Ala Glu Met Leu Gly Ile Arg Phe Arg Leu
195 200 205

ctg cag caa ctg cta cag ccg gaa ttt ctt aac tgg cga ttt cct cag 672
Leu Gln Gln Leu Leu Gln Pro Glu Phe Leu Asn Trp Arg Phe Pro Gln
210 215 220

ctg cac gta cag gag tta cgc ata cgc cag gtg gtt tcc atg atc cag 720
Leu His Val Gln Glu Leu Arg Ile Arg Gln Val Val Ser Met Ile Gln
225 230 235 240

gag ctg cac tat ctc atc cag gag att aac cga gtt tac gcc ctc agt 768
Glu Leu His Tyr Leu Ile Gln Glu Ile Asn Arg Val Tyr Ala Leu Ser
245 250 255

ctg tgg gca gcc atg gcc cat gat ctg gcc atg agc acg agt gag ttg 816
Leu Trp Ala Ala Met Ala His Asp Leu Ala Met Ser Thr Ser Glu Leu
260 265 270

tac atc ctg ttt ggc cag tcc gtg ggc att ggc caa cag aat gag gag 864
Tyr Ile Leu Phe Gly Gln Ser Val Gly Ile Gly Gln Gln Asn Glu Glu
275 280 285

gag aac ggc agt tgc tat cga atg ctc ggc tat ttg gcc cta gtc atg 912
Glu Asn Gly Ser Cys Tyr Arg Met Leu Gly Tyr Leu Ala Leu Val Met
290 295 300

atc cca ccg ctc tac aag ctt ctg ata gct cca ttc tat tgc gat cgc 960
Ile Pro Pro Leu Tyr Lys Leu Leu Ile Ala Pro Phe Tyr Cys Asp Arg
305 310 315 320

acc ata tac gag gca agg aga tgc ctt cgc ctc gtc gaa aag ttg gac 1008
Thr Ile Tyr Glu Ala Arg Arg Cys Leu Arg Leu Val Glu Lys Leu Asp
325 330 335

gat tgg ttc ccc caa aag tct tct ctg cga ccc ctg gtc gaa tcc cta 1056
Asp Trp Phe Pro Gln Lys Ser Ser Leu Arg Pro Leu Val Glu Ser Leu
340 345 350

atg tcg tgg cgc att cag gcc aag att cag ttc acc agt ggc cta gat 1104
Met Ser Trp Arg Ile Gln Ala Lys Ile Gln Phe Thr Ser Gly Leu Asp
355 360 365

gtt gtg ctt agc cgc aaa gtt atc ggt ttg ttt aca tcg att ctg gtc 1152
Val Val Leu Ser Arg Lys Val Ile Gly Leu Phe Thr Ser Ile Leu Val
370 375 380

```

WO 00/77208

PCT/US00/16211

```

aat tac ctg ctc ata ctc atc cag ttc gcc atg acc cag aag atg ggc 1200
Asn Tyr Leu Leu Ile Leu Ile Gln Phe Ala Met Thr Gln Lys Met Gly
385          390          395          400

gag caa atc gag cag cag aag att gca ctg cag gaa tgg att gga ttc 1248
Glu Gln Ile Glu Gln Gln Lys Ile Ala Leu Gln Glu Trp Ile Gly Phe
          405          410          415

taa 1251

```

<210> 46
 <211> 416
 <212> PRT
 <213> *Drosophila melanogaster*

```

<400> 46
Met Ala Val Leu Tyr Phe Phe Arg Glu Pro Glu Thr Val Phe Asp Cys
 1          5          10          15

Ala Ala Phe Ile Cys Ile Leu Gln Phe Leu Met Gly Cys Asn Gly Phe
 20          25          30

Gly Ile Arg Arg Ser Thr Phe Arg Ile Ser Trp Ala Ser Arg Ile Tyr
 35          40          45

Ser Met Ser Val Ala Ile Ala Ala Phe Cys Cys Leu Phe Gly Ser Leu
 50          55          60

Ser Val Leu Leu Ala Glu Glu Asp Ile Arg Glu Arg Leu Ala Lys Ala
 65          70          75          80

Asp Asn Leu Val Leu Ser Ile Ser Ala Leu Glu Leu Leu Met Ser Thr
 85          90          95

Leu Val Phe Gly Val Thr Val Ile Ser Leu Gln Val Phe Ala Arg Arg
100          105          110

His Leu Gly Ile Tyr Gln Arg Leu Ala Ala Leu Asp Ala Arg Leu Met
115          120          125

Ser Asp Phe Gly Ala Asn Leu Asn Tyr Arg Lys Met Leu Arg Lys Asn
130          135          140

Ile Ala Val Leu Gly Ile Val Thr Thr Ile Tyr Leu Met Ala Ile Asn
145          150          155          160

Ser Ala Ala Val Gln Val Ala Ser Gly His Arg Ala Leu Phe Leu Leu
165          170          175

Phe Ala Leu Cys Tyr Thr Ile Val Thr Gly Gly Pro His Phe Thr Gly
180          185          190

Tyr Val His Met Thr Leu Ala Glu Met Leu Gly Ile Arg Phe Arg Leu
195          200          205

```

```

Leu Gln Gln Leu Leu Gln Pro Glu Phe Leu Asn Trp Arg Phe Pro Gln
210                215                220

Leu His Val Gln Glu Leu Arg Ile Arg Gln Val Val Ser Met Ile Gln
225                230                235                240

Glu Leu His Tyr Leu Ile Gln Glu Ile Asn Arg Val Tyr Ala Leu Ser
245                250                255

Leu Trp Ala Ala Met Ala His Asp Leu Ala Met Ser Thr Ser Glu Leu
260                265                270

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

Glu Asn Gly Ser Cys Tyr Arg Met Leu Gly Tyr Leu Ala Leu Val Met
290                295                300

Ile Pro Pro Leu Tyr Lys Leu Leu Ile Ala Pro Phe Tyr Cys Asp Arg
305                310                315                320

Thr Ile Tyr Glu Ala Arg Arg Cys Leu Arg Leu Val Glu Lys Leu Asp
325                330                335

Asp Trp Phe Pro Gln Lys Ser Ser Leu Arg Pro Leu Val Glu Ser Leu
340                345                350

Met Ser Trp Arg Ile Gln Ala Lys Ile Gln Phe Thr Ser Gly Leu Asp
355                360                365

Val Val Leu Ser Arg Lys Val Ile Gly Leu Phe Thr Ser Ile Leu Val
370                375                380

Asn Tyr Leu Leu Ile Leu Ile Gln Phe Ala Met Thr Gln Lys Met Gly
385                390                395                400

Glu Gln Ile Glu Gln Gln Lys Ile Ala Leu Gln Glu Trp Ile Gly Phe
405                410                415

```

```

<210> 47
<211> 1185
<212> DNA
<213> Drosophila melanogaster

```

```

<220>
<221> CDS
<222> (1)..(1185)
<223> Coding region GR58A.1

```

```

<220>
<221> misc_feature
<222> (1021)
<223> Splice site

```

```

<400> 47
atg ctg ttg aaa ttc atg tac ata tat gga ata ggc tgt ggt ctg atg   48
Met Leu Leu Lys Phe Met Tyr Ile Tyr Gly Ile Gly Cys Gly Leu Met
  1             5             10             15

cct gct cct ctg aag aaa gga cag ttt ctt ttg gga tat aag cag aga   96
Pro Ala Pro Leu Lys Lys Gly Gln Phe Leu Leu Gly Tyr Lys Gln Arg
      20             25             30

tgg tac ctc atc tat acg gcc tgc ctg cac ggc ggt cta ctg acc gtc  144
Trp Tyr Leu Ile Tyr Thr Ala Cys Leu His Gly Gly Leu Leu Thr Val
      35             40             45

ctg ccc ttc aca ttt ccc cac tac atg tac gac gac agc tac atg agc  192
Leu Pro Phe Thr Phe Pro His Tyr Met Tyr Asp Asp Ser Tyr Met Ser
      50             55             60

agc aat cca gtc ctg aaa tgg aca ttt aat cta acc aac atc acc cgt  240
Ser Asn Pro Val Leu Lys Trp Thr Phe Asn Leu Thr Asn Ile Thr Arg
      65             70             75             80

att atg gcc atg ttt tgc ggc gtt ctc ctg atg tgg ttc aga aga aaa  288
Ile Met Ala Met Phe Ser Gly Val Leu Leu Met Trp Phe Arg Arg Lys
      85             90             95

agg atc ttg aat ttg ggt gaa aac tta ata ctt cac tgt ctt aag tgc  336
Arg Ile Leu Asn Leu Gly Glu Asn Leu Ile Leu His Cys Leu Lys Cys
      100            105            110

aaa aca ctg gat aat cgt tct aag aag tat tct aaa ttg cgg aaa aga  384
Lys Thr Leu Asp Asn Arg Ser Lys Lys Tyr Ser Lys Leu Arg Lys Arg
      115            120            125

gtt cga aat gta ctt ttc cag atg tta ctc gtc gca aat ctc agc att  432
Val Arg Asn Val Leu Phe Gln Met Leu Leu Val Ala Asn Leu Ser Ile
      130            135            140

ctg ctg ggg gct tta att ttg ttt aga att cat tca gta caa aga att  480
Leu Leu Gly Ala Leu Ile Leu Phe Arg Ile His Ser Val Gln Arg Ile
      145            150            155            160

agt aaa act gca atg att gta gcc cat att aca cag ttt ata tat gtg  528
Ser Lys Thr Ala Met Ile Val Ala His Ile Thr Gln Phe Ile Tyr Val
      165            170            175

gtc ttt atg atg acg gga ata tgt gtg atc cta tta gtt ctc cac tgg  576
Val Phe Met Met Thr Gly Ile Cys Val Ile Leu Leu Val Leu His Trp
      180            185            190

caa agt gag aga ctt caa ata gca ctc aag gac ttg tgc tcc ttt ctt  624
Gln Ser Glu Arg Leu Gln Ile Ala Leu Lys Asp Leu Cys Ser Phe Leu
      195            200            205

aat cat gaa gaa cgc aac tcc ttg act tta tca gaa aac aag gcc aat  672
Asn His Glu Glu Arg Asn Ser Leu Thr Leu Ser Glu Asn Lys Ala Asn
      210            215            220

```

```

aga tcc ctc gga aaa ttg gct aag ctc ttt aaa ctt ttt gcg gaa aat 720
Arg Ser Leu Gly Lys Leu Ala Lys Leu Phe Lys Leu Phe Ala Glu Asn
225                230                235                240

cag cga tta gtc cga gaa gtt ttt cga acc ttt gac ttg ccc atc gcc 768
Gln Arg Leu Val Arg Glu Val Phe Arg Thr Phe Asp Leu Pro Ile Ala
245                250                255

ctt ctt ttg ctg aaa atg ttt gtt aca aac gtg aat ctg gtt tat cat 816
Leu Leu Leu Leu Lys Met Phe Val Thr Asn Val Asn Leu Val Tyr His
260                265                270

gga gtg caa ttt ggc aac gat acc ata gaa acc tca agt tac aca aga 864
Gly Val Gln Phe Gly Asn Asp Thr Ile Glu Thr Ser Ser Tyr Thr Arg
275                280                285

atc gtg gga cag tgg gtg gtt att tct cat tac tgg agt gct gtt ttg 912
Ile Val Gly Gln Trp Val Val Ile Ser His Tyr Trp Ser Ala Val Leu
290                295                300

ctc atg aat gtt gtg gat gac gtg acg cgg aga agt gac ctg aaa atg 960
Leu Met Asn Val Val Asp Asp Val Thr Arg Arg Ser Asp Leu Lys Met
305                310                315                320

gga gac ctc ctg cga gaa ttc agt cat ctg gag ctg gtc aaa agg gac 1008
Gly Asp Leu Leu Arg Glu Phe Ser His Leu Glu Leu Val Lys Arg Asp
325                330                335

ttt cat ttg cag ctg gaa ctt ttc tgc gat cac ctt cgt tgt cat ccg 1056
Phe His Leu Gln Leu Glu Leu Phe Ser Asp His Leu Arg Cys His Pro
340                345                350

tca acg tat aag gtc tgt gga cta ttt att ttc aat aag caa acg agt 1104
Ser Thr Tyr Lys Val Cys Gly Leu Phe Ile Phe Asn Lys Gln Thr Ser
355                360                365

ttg gct tat ttt ttc tat gtg ctg gtt caa gtt ttg gtg ctt gta caa 1152
Leu Ala Tyr Phe Phe Tyr Val Leu Val Gln Val Leu Val Leu Val Gln
370                375                380

ttt gat tta aaa aat aaa gtt gaa aaa aga aat 1185
Phe Asp Leu Lys Asn Lys Val Glu Lys Arg Asn
385                390                395

```

<210> 48
 <211> 395
 <212> PRT
 <213> *Drosophila melanogaster*

<400> 48
 Met Leu Leu Lys Phe Met Tyr Ile Tyr Gly Ile Gly Cys Gly Leu Met
 1 5 10 15
 Pro Ala Pro Leu Lys Lys Gly Gln Phe Leu Leu Gly Tyr Lys Gln Arg
 20 25 30

Trp Tyr Leu Ile Tyr Thr Ala Cys Leu His Gly Gly Leu Leu Thr Val
 35 40 45
 Leu Pro Phe Thr Phe Pro His Tyr Met Tyr Asp Asp Ser Tyr Met Ser
 50 55 60
 Ser Asn Pro Val Leu Lys Trp Thr Phe Asn Leu Thr Asn Ile Thr Arg
 65 70 75 80
 Ile Met Ala Met Phe Ser Gly Val Leu Leu Met Trp Phe Arg Arg Lys
 85 90 95
 Arg Ile Leu Asn Leu Gly Glu Asn Leu Ile Leu His Cys Leu Lys Cys
 100 105 110
 Lys Thr Leu Asp Asn Arg Ser Lys Lys Tyr Ser Lys Leu Arg Lys Arg
 115 120 125
 Val Arg Asn Val Leu Phe Gln Met Leu Leu Val Ala Asn Leu Ser Ile
 130 135 140
 Leu Leu Gly Ala Leu Ile Leu Phe Arg Ile His Ser Val Gln Arg Ile
 145 150 155 160
 Ser Lys Thr Ala Met Ile Val Ala His Ile Thr Gln Phe Ile Tyr Val
 165 170 175
 Val Phe Met Met Thr Gly Ile Cys Val Ile Leu Leu Val Leu His Trp
 180 185 190
 Gln Ser Glu Arg Leu Gln Ile Ala Leu Lys Asp Leu Cys Ser Phe Leu
 195 200 205
 Asn His Glu Glu Arg Asn Ser Leu Thr Leu Ser Glu Asn Lys Ala Asn
 210 215 220
 Arg Ser Leu Gly Lys Leu Ala Lys Leu Phe Lys Leu Phe Ala Glu Asn
 225 230 235 240
 Gln Arg Leu Val Arg Glu Val Phe Arg Thr Phe Asp Leu Pro Ile Ala
 245 250 255
 Leu Leu Leu Leu Lys Met Phe Val Thr Asn Val Asn Leu Val Tyr His
 260 265 270
 Gly Val Gln Phe Gly Asn Asp Thr Ile Glu Thr Ser Ser Tyr Thr Arg
 275 280 285
 Ile Val Gly Gln Trp Val Val Ile Ser His Tyr Trp Ser Ala Val Leu
 290 295 300
 Leu Met Asn Val Val Asp Asp Val Thr Arg Arg Ser Asp Leu Lys Met
 305 310 315 320
 Gly Asp Leu Leu Arg Glu Phe Ser His Leu Glu Leu Val Lys Arg Asp
 325 330 335

Phe His Leu Gln Leu Glu Leu Phe Ser Asp His Leu Arg Cys His Pro
 340 345 350
 Ser Thr Tyr Lys Val Cys Gly Leu Phe Ile Phe Asn Lys Gln Thr Ser
 355 360 365
 Leu Ala Tyr Phe Phe Tyr Val Leu Val Gln Val Leu Val Leu Val Gln
 370 375 380
 Phe Asp Leu Lys Asn Lys Val Glu Lys Arg Asn
 385 390 395

<210> 49
 <211> 1224
 <212> DNA
 <213> *Drosophila melanogaster*

<220>
 <221> CDS
 <222> (1)..(1224)
 <223> Coding region GR58A.2

<220>
 <221> misc_feature
 <222> (1015)
 <223> Splice site

<400> 49
 atg ttg cat ccg aaa ctt ggt cgc gtg atg aac gtg gtc tac tac cat 48
 Met Leu His Pro Lys Leu Gly Arg Val Met Asn Val Val Tyr Tyr His
 1 5 10 15
 tcc gtt gtc ttc gcc cta atg agc act acc cta agg ata cgc tcc tgc 96
 Ser Val Val Phe Ala Leu Met Ser Thr Thr Leu Arg Ile Arg Ser Cys
 20 25 30
 agg aaa tgc cta cgc ttg gaa aag gtc tcg cga act tac acc atc tac 144
 Arg Lys Cys Leu Arg Leu Glu Lys Val Ser Arg Thr Tyr Thr Ile Tyr
 35 40 45
 agt ttc ttt gtg ggc ata ttc ctg ttt cta aac ttg tac ttc atg gtg 192
 Ser Phe Phe Val Gly Ile Phe Leu Phe Leu Asn Leu Tyr Phe Met Val
 50 55 60
 cct cgg atc atg gaa gat ggc tac atg aag tac aac ata gtc ctg cag 240
 Pro Arg Ile Met Glu Asp Gly Tyr Met Lys Tyr Asn Ile Val Leu Gln
 65 70 75 80
 tgg aac ttc ttt gtg atg ctc ttc ctt aga gcc atc gct gtg gtg agc 288
 Trp Asn Phe Phe Val Met Leu Phe Leu Arg Ala Ile Ala Val Val Ser
 85 90 95
 tgc tat gga acc ctg tgg cta aag cgc cac aaa atc att cag ctc tat 336
 Cys Tyr Gly Thr Leu Trp Leu Lys Arg His Lys Ile Ile Gln Leu Tyr

	100	105	110	
aaa tac tcg ctg ata tat tgg aag agg ttc gga cac ata acg agg gcc				384
Lys Tyr Ser Leu Ile Tyr Trp Lys Arg Phe Gly His Ile Thr Arg Ala				
	115	120	125	
ata gtg gac aag aaa gaa ctc ctt gac cta caa gag tcc ttg gcc aga				432
Ile Val Asp Lys Lys Glu Leu Leu Asp Leu Gln Glu Ser Leu Ala Arg				
	130	135	140	
ata atg atc cgg aaa ata ata tta ctc tat agt gct ttt ctg tgc tcc				480
Ile Met Ile Arg Lys Ile Ile Leu Leu Tyr Ser Ala Phe Leu Cys Ser				
	145	150	155	160
act gtg ttg cag tat caa ttg ctc agt gtg att aac cca cag att ttc				528
Thr Val Leu Gln Tyr Gln Leu Leu Ser Val Ile Asn Pro Gln Ile Phe				
	165	170	175	
ctg gct ttc tgt gcc cgg ctc acg cac ttt ctg cac ttc ttg tgc gtg				576
Leu Ala Phe Cys Ala Arg Leu Thr His Phe Leu His Phe Leu Cys Val				
	180	185	190	
aaa atg ggt ttc ttc ggg gta ctg gtt cta ctg aat cat caa ttt ttg				624
Lys Met Gly Phe Phe Gly Val Leu Val Leu Leu Asn His Gln Phe Leu				
	195	200	205	
gta atc cat ctg gcc ata aac gct ctt cat ggt agg aaa gcc cgg aag				672
Val Ile His Leu Ala Ile Asn Ala Leu His Gly Arg Lys Ala Arg Lys				
	210	215	220	
aaa tgg aaa gct ctg cgt tcc gta gcc gcc atg cat ctg aaa acc ctt				720
Lys Trp Lys Ala Leu Arg Ser Val Ala Ala Met His Leu Lys Thr Leu				
	225	230	235	240
cga tta gcc aga agg atc ttt gac atg ttt gac atc gct aat gcc acg				768
Arg Leu Ala Arg Arg Ile Phe Asp Met Phe Asp Ile Ala Asn Ala Thr				
	245	250	255	
gtg ttt atc aac atg ttt atg acc gct att aat atc ctt tat cat gcc				816
Val Phe Ile Asn Met Phe Met Thr Ala Ile Asn Ile Leu Tyr His Ala				
	260	265	270	
gtc cag tac agc aat agc agc atc aag tca aat ggt tgg ggc atc tta				864
Val Gln Tyr Ser Asn Ser Ser Ile Lys Ser Asn Gly Trp Gly Ile Leu				
	275	280	285	
ttt ggc aac gga ttg atc gtt ttc aac ttc tgg ggc acc atg gcg ctg				912
Phe Gly Asn Gly Leu Ile Val Phe Asn Phe Trp Gly Thr Met Ala Leu				
	290	295	300	
atg gaa atg ctg gat agt gtg gtg acc tcc tgc aac aac act ggc cag				960
Met Glu Met Leu Asp Ser Val Val Thr Ser Cys Asn Asn Thr Gly Gln				
	305	310	315	320
caa cta agg caa ctt agc gat cta cca aaa gtg ggt cca aag atg caa				1008
Gln Leu Arg Gln Leu Ser Asp Leu Pro Lys Val Gly Pro Lys Met Gln				

```

325          330          335
agg gag ttg gat gtt ttt acc atg cag ctg agg cag aat cgg ttg gtc 1056
Arg Glu Leu Asp Val Phe Thr Met Gln Leu Arg Gln Asn Arg Leu Val
340          345          350

tac aaa ata tgc gga att gtg gag ctg gac aaa ccc gct tgc ctt agt 1104
Tyr Lys Ile Cys Gly Ile Val Glu Leu Asp Lys Pro Ala Cys Leu Ser
355          360          365

tat att ggc tcc att ctg agc aac gtc att atc ctc atg cag ttc gat 1152
Tyr Ile Gly Ser Ile Leu Ser Asn Val Ile Ile Leu Met Gln Phe Asp
370          375          380

ttg aga cgg caa aga caa ccc atc aat gat cgt caa tat ctc atc cat 1200
Leu Arg Arg Gln Arg Gln Pro Ile Asn Asp Arg Gln Tyr Leu Ile His
385          390          395          400

ttg atg aag aac aaa aca aaa gtg 1224
Leu Met Lys Asn Lys Thr Lys Val
405

<210> 50
<211> 408
<212> PRT
<213> Drosophila melanogaster

<400> 50
Met Leu His Pro Lys Leu Gly Arg Val Met Asn Val Val Tyr Tyr His
1          5          10          15

Ser Val Val Phe Ala Leu Met Ser Thr Thr Leu Arg Ile Arg Ser Cys
20          25          30

Arg Lys Cys Leu Arg Leu Glu Lys Val Ser Arg Thr Tyr Thr Ile Tyr
35          40          45

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

Pro Arg Ile Met Glu Asp Gly Tyr Met Lys Tyr Asn Ile Val Leu Gln
65          70          75          80

Trp Asn Phe Phe Val Met Leu Phe Leu Arg Ala Ile Ala Val Val Ser
85          90          95

Cys Tyr Gly Thr Leu Trp Leu Lys Arg His Lys Ile Ile Gln Leu Tyr
100          105          110

Lys Tyr Ser Leu Ile Tyr Trp Lys Arg Phe Gly His Ile Thr Arg Ala
115          120          125

Ile Val Asp Lys Lys Glu Leu Leu Asp Leu Gln Glu Ser Leu Ala Arg
130          135          140

```

Ile Met Ile Arg Lys Ile Ile Leu Leu Tyr Ser Ala Phe Leu Cys Ser
 145 150 155 160
 Thr Val Leu Gln Tyr Gln Leu Leu Ser Val Ile Asn Pro Gln Ile Phe
 165 170 175
 Leu Ala Phe Cys Ala Arg Leu Thr His Phe Leu His Phe Leu Cys Val
 180 185 190
 Lys Met Gly Phe Phe Gly Val Leu Val Leu Leu Asn His Gln Phe Leu
 195 200 205
 Val Ile His Leu Ala Ile Asn Ala Leu His Gly Arg Lys Ala Arg Lys
 210 215 220
 Lys Trp Lys Ala Leu Arg Ser Val Ala Ala Met His Leu Lys Thr Leu
 225 230 235 240
 Arg Leu Ala Arg Arg Ile Phe Asp Met Phe Asp Ile Ala Asn Ala Thr
 245 250 255
 Val Phe Ile Asn Met Phe Met Thr Ala Ile Asn Ile Leu Tyr His Ala
 260 265 270
 Val Gln Tyr Ser Asn Ser Ser Ile Lys Ser Asn Gly Trp Gly Ile Leu
 275 280 285
 Phe Gly Asn Gly Leu Ile Val Phe Asn Phe Trp Gly Thr Met Ala Leu
 290 295 300
 Met Glu Met Leu Asp Ser Val Val Thr Ser Cys Asn Asn Thr Gly Gln
 305 310 315 320
 Gln Leu Arg Gln Leu Ser Asp Leu Pro Lys Val Gly Pro Lys Met Gln
 325 330 335
 Arg Glu Leu Asp Val Phe Thr Met Gln Leu Arg Gln Asn Arg Leu Val
 340 345 350
 Tyr Lys Ile Cys Gly Ile Val Glu Leu Asp Lys Pro Ala Cys Leu Ser
 355 360 365
 Tyr Ile Gly Ser Ile Leu Ser Asn Val Ile Ile Leu Met Gln Phe Asp
 370 375 380
 Leu Arg Arg Gln Arg Gln Pro Ile Asn Asp Arg Gln Tyr Leu Ile His
 385 390 395 400
 Leu Met Lys Asn Lys Thr Lys Val
 405

<210> 51
 <211> 1236
 <212> DNA

<213> Drosophila melanogaster

<220>

<221> CDS

<222> (1)..(1236)

<223> Coding region GR58A.3

<220>

<221> misc_feature

<222> (1066)

<223> Splice site

<400> 51

```

atg aat caa tac ttt ttg ctg cat act tac ttt caa gtg agc cgc ttg   48
Met Asn Gln Tyr Phe Leu Leu His Thr Tyr Phe Gln Val Ser Arg Leu
   1             5             10             15

ata gga ctc tgc aat ctg cac tac gat tcc tcg aat cat cgg ttc atc   96
Ile Gly Leu Cys Asn Leu His Tyr Asp Ser Ser Asn His Arg Phe Ile
      20             25             30

ctc aat cat gtg ccc aca gtg gtt tac tgt gtt atc ctg aat gtg gtc   144
Leu Asn His Val Pro Thr Val Val Tyr Cys Val Ile Leu Asn Val Val
      35             40             45

tat ctc ttg gtc ctt cca ttt gcc ctg ttc gtg cta act ggc aac att   192
Tyr Leu Leu Val Leu Pro Phe Ala Leu Phe Val Leu Thr Gly Asn Ile
      50             55             60

tac cat tgt cca gat gcg ggc atg ttt gga gtg gtc tac aac gtg gtg   240
Tyr His Cys Pro Asp Ala Gly Met Phe Gly Val Val Tyr Asn Val Val
      65             70             75             80

gct ctg acc aaa ctc ctg acc atg ctc ttc cta atg agc agt gtt tgg   288
Ala Leu Thr Lys Leu Leu Thr Met Leu Phe Leu Met Ser Ser Val Trp
      85             90             95

ata cag aga cgt cga ctg tat aag ttg gga aac gac tta atg aag atg   336
Ile Gln Arg Arg Arg Leu Tyr Lys Leu Gly Asn Asp Leu Met Lys Met
      100            105            110

ttg cac aaa ttc cga ttt aac ctt gga aac gat tgc agg aat aga tgc   384
Leu His Lys Phe Arg Phe Asn Leu Gly Asn Asp Cys Arg Asn Arg Cys
      115            120            125

ctg tgc aag ggt ctg ctg acc agc tct cga ttt gtg ctc cta acc cag   432
Leu Cys Lys Gly Leu Leu Thr Ser Ser Arg Phe Val Leu Leu Thr Gln
      130            135            140

cag ctg ttg acc cgg gat tcg gtg gtt aat tgc gag agt aac tcc agt   480
Gln Leu Leu Thr Arg Asp Ser Val Val Asn Cys Glu Ser Asn Ser Ser
      145            150            155            160

ttg aga caa gct atg gtt ccg tac caa agt gct gcc ata gtg tat gcc   528
Leu Arg Gln Ala Met Val Pro Tyr Gln Ser Ala Ala Ile Val Tyr Ala
      165            170            175

```

ttt atc atg att cta tta atg agc tat gtg gat atg aca gtc tat atg	576
Leu Ile Met Ile Leu Leu Met Ser Tyr Val Asp Met Thr Val Tyr Met	
180 185 190	
gtc gaa gtg gct ggc aat tgg ctg ctg gta aat atg acc cag ggg gtt	624
Val Glu Val Ala Gly Asn Trp Leu Leu Val Asn Met Thr Gln Gly Val	
195 200 205	
cgg gaa atg gtc caa gat cta gag gtc ttg ccc gaa cgg aat ggc att	672
Arg Glu Met Val Gln Asp Leu Glu Val Leu Pro Glu Arg Asn Gly Ile	
210 215 220	
ccg cgg gag atg gga ctg atg caa atc ctt gcc gcc tgg cga aaa ctt	720
Pro Arg Glu Met Gly Leu Met Gln Ile Leu Ala Ala Trp Arg Lys Leu	
225 230 235 240	
tgg aga cgc tgt cgc cgt ttg gat gcg ttg ctc aag cag ttc gtt gac	768
Trp Arg Arg Cys Arg Arg Leu Asp Ala Leu Leu Lys Gln Phe Val Asp	
245 250 255	
atc ttc cag tgg cag gtg ctc ttc aac ctg cta acc act tat ata ttc	816
Ile Phe Gln Trp Gln Val Leu Phe Asn Leu Leu Thr Thr Tyr Ile Phe	
260 265 270	
agc att gct gtt ttg ttt cga ttg tgg att tat ttg gag ttc gat aaa	864
Ser Ile Ala Val Leu Phe Arg Leu Trp Ile Tyr Leu Glu Phe Asp Lys	
275 280 285	
aac ttt cat tta tgg aag ggc ata ttg tat gcg att att ttt ctg acc	912
Asn Phe His Leu Trp Lys Gly Ile Leu Tyr Ala Ile Ile Phe Leu Thr	
290 295 300	
cat cac gtc gaa atc gta atg caa ttt tcc att ttc gag atc aac cgc	960
His His Val Glu Ile Val Met Gln Phe Ser Ile Phe Glu Ile Asn Arg	
305 310 315 320	
tgt aag tgg ttg gga ctt tta gaa gat gtc gga aat ctg tgg gac atc	1008
Cys Lys Trp Leu Gly Leu Leu Glu Asp Val Gly Asn Leu Trp Asp Ile	
325 330 335	
aat tat tcg gga agg caa tgc att aaa agc agt gga acg att ctg tca	1056
Asn Tyr Ser Gly Arg Gln Cys Ile Lys Ser Ser Gly Thr Ile Leu Ser	
340 345 350	
aga aag tta gag ttt tcc ctg ctc tac atg aat cgc aaa ctt caa ctg	1104
Arg Lys Leu Glu Phe Ser Leu Leu Tyr Met Asn Arg Lys Leu Gln Leu	
355 360 365	
aat cca aaa cgt gtg aga cgt ctg cac atc gtt gga ttg ttc gat ttg	1152
Asn Pro Lys Arg Val Arg Arg Leu His Ile Val Gly Leu Phe Asp Leu	
370 375 380	
agt aat tta acc gtt cac aac atg acc aga agt ata ata acc aat gtg	1200
Ser Asn Leu Thr Val His Asn Met Thr Arg Ser Ile Ile Thr Asn Val	
385 390 395 400	

tta gtt ctg tgt cag att gcc tat aaa aaa tat ggt
 Leu Val Leu Cys Gln Ile Ala Tyr Lys Lys Tyr Gly
 405 410

1236

<210> 52
 <211> 412
 <212> PRT
 <213> Drosophila melanogaster

<400> 52
 Met Asn Gln Tyr Phe Leu Leu His Thr Tyr Phe Gln Val Ser Arg Leu
 1 5 10 15
 Ile Gly Leu Cys Asn Leu His Tyr Asp Ser Ser Asn His Arg Phe Ile
 20 25 30
 Leu Asn His Val Pro Thr Val Val Tyr Cys Val Ile Leu Asn Val Val
 35 40 45
 Tyr Leu Leu Val Leu Pro Phe Ala Leu Phe Val Leu Thr Gly Asn Ile
 50 55 60
 Tyr His Cys Pro Asp Ala Gly Met Phe Gly Val Val Tyr Asn Val Val
 65 70 75 80
 Ala Leu Thr Lys Leu Leu Thr Met Leu Phe Leu Met Ser Ser Val Trp
 85 90 95
 Ile Gln Arg Arg Arg Leu Tyr Lys Leu Gly Asn Asp Leu Met Lys Met
 100 105 110
 Leu His Lys Phe Arg Phe Asn Leu Gly Asn Asp Cys Arg Asn Arg Cys
 115 120 125
 Leu Cys Lys Gly Leu Leu Thr Ser Ser Arg Phe Val Leu Leu Thr Gln
 130 135 140
 Gln Leu Leu Thr Arg Asp Ser Val Val Asn Cys Glu Ser Asn Ser Ser
 145 150 155 160
 Leu Arg Gln Ala Met Val Pro Tyr Gln Ser Ala Ala Ile Val Tyr Ala
 165 170 175
 Leu Ile Met Ile Leu Leu Met Ser Tyr Val Asp Met Thr Val Tyr Met
 180 185 190
 Val Glu Val Ala Gly Asn Trp Leu Leu Val Asn Met Thr Gln Gly Val
 195 200 205
 Arg Glu Met Val Gln Asp Leu Glu Val Leu Pro Glu Arg Asn Gly Ile
 210 215 220
 Pro Arg Glu Met Gly Leu Met Gln Ile Leu Ala Ala Trp Arg Lys Leu
 225 230 235 240

<400> 53																		
atg	aag	cgg	ata	ggg	caa	gcg	tat	aac	gta	tac	gcc	gtg	ttc	atc	ggc		48	
Met	Lys	Arg	Ile	Gly	Gln	Ala	Tyr	Asn	Val	Tyr	Ala	Val	Phe	Ile	Gly			
1				5					10					15				
atg	acg	tca	tac	gaa	aca	atg	ggg	gga	aaa	ttc	agg	cag	tcg	cgg	att		96	
Met	Thr	Ser	Tyr	Glu	Thr	Met	Gly	Gly	Lys	Phe	Arg	Gln	Ser	Arg	Ile			
			20				25					30						
act	cgg	ata	tat	tgc	ttg	ctt	ata	aat	gcc	atc	ttc	ttg	act	ttg	cta		144	
Thr	Arg	Ile	Tyr	Cys	Leu	Leu	Ile	Asn	Ala	Ile	Phe	Leu	Thr	Leu	Leu			

	35	40	45	
ccg agt gct ttc tgg aaa tcg gcc aag ctt ttg agt acg gcg gac tgg	192			
Pro Ser Ala Phe Trp Lys Ser Ala Lys Leu Leu Ser Thr Ala Asp Trp				
50 55 60				
atg cct tcc tac atg agg gtc act ccg tac atc atg tgt aca atc aac	240			
Met Pro Ser Tyr Met Arg Val Thr Pro Tyr Ile Met Cys Thr Ile Asn				
65 70 75 80				
tat gct gca ata gcc tac acc ttg atc tca aga tgc tac agg gac gcc	288			
Tyr Ala Ala Ile Ala Tyr Thr Leu Ile Ser Arg Cys Tyr Arg Asp Ala				
85 90 95				
atg ttg atg gac ttg caa cgc att gtg ctg gaa gtg aat cgg gaa atg	336			
Met Leu Met Asp Leu Gln Arg Ile Val Leu Glu Val Asn Arg Glu Met				
100 105 110				
ttg cgc acg gga aag aaa atg aac tca cta ctt cga cga atg ttt ttc	384			
Leu Arg Thr Gly Lys Lys Met Asn Ser Leu Leu Arg Arg Met Phe Phe				
115 120 125				
ttg aac att tac att gac ctc tcg tgc ttg tcg tac att ctg gcg gtt	432			
Leu Asn Ile Tyr Ile Asp Leu Ser Cys Leu Ser Tyr Ile Leu Ala Val				
130 135 140				
ttc att tac cag tgg aaa gct caa aac tgg tcg aat ctt tgc aat gga	480			
Phe Ile Tyr Gln Trp Lys Ala Gln Asn Trp Ser Asn Leu Cys Asn Gly				
145 150 155 160				
ctc ttg gtt aac att tcc cta act atc ctg ttt gtc aac aca ttc ttc	528			
Leu Leu Val Asn Ile Ser Leu Thr Ile Leu Phe Val Asn Thr Phe Phe				
165 170 175				
tac ttc acc tct ttg tgg cac ata gcc agg gga tat gac ttt gta aac	576			
Tyr Phe Thr Ser Leu Trp His Ile Ala Arg Gly Tyr Asp Phe Val Asn				
180 185 190				
cag caa ctg aat gag att gtc gcc tgc caa tcg atg gac ttg gag aga	624			
Gln Gln Leu Asn Glu Ile Val Ala Cys Gln Ser Met Asp Leu Glu Arg				
195 200 205				
aag tca aaa gaa ctc cgc ggc ctg tgg gct ctg cac aga aat ctt agt	672			
Lys Ser Lys Glu Leu Arg Gly Leu Trp Ala Leu His Arg Asn Leu Ser				
210 215 220				
tat aca gca cgc agg ata aac aag cac tac ggt ccc caa atg ctt gcc	720			
Tyr Thr Ala Arg Arg Ile Asn Lys His Tyr Gly Pro Gln Met Leu Ala				
225 230 235 240				
atg cga ttt gat tac ttc ata ttc tcc atc atc aac gct tgc ata ggc	768			
Met Arg Phe Asp Tyr Phe Ile Phe Ser Ile Ile Asn Ala Cys Ile Gly				
245 250 255				
acg att tac tcg act acc gac cag gag ccc tcg ctt gaa aag att ttc	816			
Thr Ile Tyr Ser Thr Thr Asp Gln Glu Pro Ser Leu Glu Lys Ile Phe				

260	265	270	
gga tct cta att tac tgg gtg cga agt ttc gat ttc ttt ctg aac gac			864
Gly Ser Leu Ile Tyr Trp Val Arg Ser Phe Asp Phe Phe Leu Asn Asp			
275	280	285	
tat atc tgc gat ctg gtc agc gag tac caa atg cag cca aag ttc ttc			912
Tyr Ile Cys Asp Leu Val Ser Glu Tyr Gln Met Gln Pro Lys Phe Phe			
290	295	300	
gcc ccc gaa agc agc atg tcc aat gag ttg agt tct tac ttg atc tac			960
Ala Pro Glu Ser Ser Met Ser Asn Glu Leu Ser Ser Tyr Leu Ile Tyr			
305	310	315	320
gag agc agc acg cga ttg gat ctg ttg gtc tgc gga ctc tat cgt gtc			1008
Glu Ser Ser Thr Arg Leu Asp Leu Leu Val Cys Gly Leu Tyr Arg Val			
325	330	335	
aat aag cgg aaa tgg ttg caa atg gtt ggc tcc atc gta gtc cac tca			1056
Asn Lys Arg Lys Trp Leu Gln Met Val Gly Ser Ile Val Val His Ser			
340	345	350	
agt atg ctg ttc cag ttc cat ctc gtc atg cga ggt ggt ctc			1098
Ser Met Leu Phe Gln Phe His Leu Val Met Arg Gly Gly Leu			
355	360	365	

<210> 54

<211> 366

<212> PRT

<213> Drosophila melanogaster

<400> 54

Met Lys Arg Ile Gly Gln Ala Tyr Asn Val Tyr Ala Val Phe Ile Gly		
1	5	10

Met Thr Ser Tyr Glu Thr Met Gly Gly Lys Phe Arg Gln Ser Arg Ile		
20	25	30

Thr Arg Ile Tyr Cys Leu Leu Ile Asn Ala Ile Phe Leu Thr Leu Leu		
35	40	45

Pro Ser Ala Phe Trp Lys Ser Ala Lys Leu Leu Ser Thr Ala Asp Trp		
50	55	60

Met Pro Ser Tyr Met Arg Val Thr Pro Tyr Ile Met Cys Thr Ile Asn		
65	70	75

Tyr Ala Ala Ile Ala Tyr Thr Leu Ile Ser Arg Cys Tyr Arg Asp Ala		
85	90	95

Met Leu Met Asp Leu Gln Arg Ile Val Leu Glu Val Asn Arg Glu Met		
100	105	110

Leu Arg Thr Gly Lys Lys Met Asn Ser Leu Leu Arg Arg Met Phe Phe		
115	120	125

Leu Asn Ile Tyr Ile Asp Leu Ser Cys Leu Ser Tyr Ile Leu Ala Val
 130 135 140
 Phe Ile Tyr Gln Trp Lys Ala Gln Asn Trp Ser Asn Leu Cys Asn Gly
 145 150 155 160
 Leu Leu Val Asn Ile Ser Leu Thr Ile Leu Phe Val Asn Thr Phe Phe
 165 170 175
 Tyr Phe Thr Ser Leu Trp His Ile Ala Arg Gly Tyr Asp Phe Val Asn
 180 185 190
 Gln Gln Leu Asn Glu Ile Val Ala Cys Gln Ser Met Asp Leu Glu Arg
 195 200 205
 Lys Ser Lys Glu Leu Arg Gly Leu Trp Ala Leu His Arg Asn Leu Ser
 210 215 220
 Tyr Thr Ala Arg Arg Ile Asn Lys His Tyr Gly Pro Gln Met Leu Ala
 225 230 235 240
 Met Arg Phe Asp Tyr Phe Ile Phe Ser Ile Ile Asn Ala Cys Ile Gly
 245 250 255
 Thr Ile Tyr Ser Thr Thr Asp Gln Glu Pro Ser Leu Glu Lys Ile Phe
 260 265 270
 Gly Ser Leu Ile Tyr Trp Val Arg Ser Phe Asp Phe Phe Leu Asn Asp
 275 280 285
 Tyr Ile Cys Asp Leu Val Ser Glu Tyr Gln Met Gln Pro Lys Phe Phe
 290 295 300
 Ala Pro Glu Ser Ser Met Ser Asn Glu Leu Ser Ser Tyr Leu Ile Tyr
 305 310 315 320
 Glu Ser Ser Thr Arg Leu Asp Leu Leu Val Cys Gly Leu Tyr Arg Val
 325 330 335
 Asn Lys Arg Lys Trp Leu Gln Met Val Gly Ser Ile Val Val His Ser
 340 345 350
 Ser Met Leu Phe Gln Phe His Leu Val Met Arg Gly Gly Leu
 355 360 365

<210> 55
 <211> 1161
 <212> DNA
 <213> Drosophila melanogaster

<220>
 <221> CDS
 <222> (1) .. (1161)
 <223> Coding region GR59C.2

```

<400> 55
atg gtt tat tgg atg att aaa ttg tat ttc cgc tac tcg cta gca att 48
Met Val Tyr Trp Met Ile Lys Leu Tyr Phe Arg Tyr Ser Leu Ala Ile
  1             5             10            15

gga att aca tca cag caa ttt tcg aat cga aag ttt ttc agt acc cta 96
Gly Ile Thr Ser Gln Gln Phe Ser Asn Arg Lys Phe Phe Ser Thr Leu
      20             25            30

ttt tct cgg acg tat gct cta att gcc aac atc gtg acg ctc atc atg 144
Phe Ser Arg Thr Tyr Ala Leu Ile Ala Asn Ile Val Thr Leu Ile Met
      35             40            45

ttg ccc atc gtg atg tgg cag gtt caa ttg gtt ttc cag cag aag aag 192
Leu Pro Ile Val Met Trp Gln Val Gln Leu Val Phe Gln Gln Lys Lys
      50             55            60

acc ttt cca aag ctc att ttg atc acc aat aac gtg agg gaa gcg gtg 240
Thr Phe Pro Lys Leu Ile Leu Ile Thr Asn Asn Val Arg Glu Ala Val
      65             70            75            80

tcc ttc ctg gtc ata ttg tac aca gtg ctg tcg cga gga ttt cgc gat 288
Ser Phe Leu Val Ile Leu Tyr Thr Val Leu Ser Arg Gly Phe Arg Asp
      85             90            95

aca gcg ttt aag gag atg caa cca ttg ctg ctg acg cta ttt cga gaa 336
Thr Ala Phe Lys Glu Met Gln Pro Leu Leu Leu Thr Leu Phe Arg Glu
      100            105            110

gag aag cga tgt ggc ttt aag ggc ata ggc ggt gta cga agg tcc ttg 384
Glu Lys Arg Cys Gly Phe Lys Gly Ile Gly Gly Val Arg Arg Ser Leu
      115            120            125

cgg att ctt ctg ttt gtg aag ttc ttc acg ttg tct tgg ctg tgc gtc 432
Arg Ile Leu Leu Phe Val Lys Phe Phe Thr Leu Ser Trp Leu Cys Val
      130            135            140

aca gac gtt ttg ttt cta ctc tac tcg acc gat gcg tta atc tgg gtt 480
Thr Asp Val Leu Phe Leu Leu Tyr Ser Thr Asp Ala Leu Ile Trp Val
      145            150            155            160

aac gtg ctc agg ttt ttt ttc aag tgt aat acc aac aat att ttg gaa 528
Asn Val Leu Arg Phe Phe Phe Lys Cys Asn Thr Asn Asn Ile Leu Glu
      165            170            175

atg gtg ccc atg ggc tac ttt ctc gcc ctg tgg cac att gct cgt ggc 576
Met Val Pro Met Gly Tyr Phe Leu Ala Leu Trp His Ile Ala Arg Gly
      180            185            190

ttc gat tgc gtc aac agg cgt ctg gac caa att gtg aaa tcg aaa tcg 624
Phe Asp Cys Val Asn Arg Arg Leu Asp Gln Ile Val Lys Ser Lys Ser
      195            200            205

act aga aaa cac agg gag ctg caa cat ctc tgg cta ctt cat gcc tgc 672
Thr Arg Lys His Arg Glu Leu Gln His Leu Trp Leu Leu His Ala Cys
      210            215            220

```

```

ctc acc aag aca gcg ctc aac ata aac aag atc tac gcc ccc cag atg 720
Leu Thr Lys Thr Ala Leu Asn Ile Asn Lys Ile Tyr Ala Pro Gln Met
225                230                235                240

ttg gcc tcc cgg ttc gat aac ttt gta aac ggt gta atc cag gcc tat 768
Leu Ala Ser Arg Phe Asp Asn Phe Val Asn Gly Val Ile Gln Ala Tyr
245                250                255

tgg ggt gct gtg ttc acc ttt gac ctc tcc acg ccc ttc ttt tgg gtc 816
Trp Gly Ala Val Phe Thr Phe Asp Leu Ser Thr Pro Phe Phe Trp Val
260                265                270

gtt tat gga agt gtc caa tac cac gtg cgc tgc ttg gac tac tat ctc 864
Val Tyr Gly Ser Val Gln Tyr His Val Arg Cys Leu Asp Tyr Tyr Leu
275                280                285

atc gat aat atg tgc gat gtg gct gtg gag tac cac gat tca gcc aag 912
Ile Asp Asn Met Cys Asp Val Ala Val Glu Tyr His Asp Ser Ala Lys
290                295                300

cac tcg tgg agt gaa gtt cgt tgg aca aaa gaa gta tct gca ttc ggt 960
His Ser Trp Ser Glu Val Arg Trp Thr Lys Glu Val Ser Ala Phe Gly
305                310                315                320

tct att cta ttg tac ata tgt atg ctc atg caa ctt ctt tca ttt cag 1008
Ser Ile Leu Leu Tyr Ile Cys Met Leu Met Gln Leu Leu Ser Phe Gln
325                330                335

ata agc tcc tat gtg atc tat gcg aat agt acg aag ttg caa ctc tgg 1056
Ile Ser Ser Tyr Val Ile Tyr Ala Asn Ser Thr Lys Leu Gln Leu Trp
340                345                350

agt tgt ggg ctt ttt caa gcc aat cgc agc atg tgg ttt gcc atg ata 1104
Ser Cys Gly Leu Phe Gln Ala Asn Arg Ser Met Trp Phe Ala Met Ile
355                360                365

agc agc gtg ttg tat tac ata cta gtg cta ctg caa ttt cac cta gtt 1152
Ser Ser Val Leu Tyr Tyr Ile Leu Val Leu Leu Gln Phe His Leu Val
370                375                380

atg agg aag 1161
Met Arg Lys
385

```

<210> 56
 <211> 387
 <212> PRT
 <213> *Drosophila melanogaster*

<400> 56
 Met Val Tyr Trp Met Ile Lys Leu Tyr Phe Arg Tyr Ser Leu Ala Ile
 1 5 10 15
 Gly Ile Thr Ser Gln Gln Phe Ser Asn Arg Lys Phe Phe Ser Thr Leu
 20 25 30

Phe Ser Arg Thr Tyr Ala Leu Ile Ala Asn Ile Val Thr Leu Ile Met
 35 40 45
 Leu Pro Ile Val Met Trp Gln Val Gln Leu Val Phe Gln Gln Lys Lys
 50 55 60
 Thr Phe Pro Lys Leu Ile Leu Ile Thr Asn Asn Val Arg Glu Ala Val
 65 70 75 80
 Ser Phe Leu Val Ile Leu Tyr Thr Val Leu Ser Arg Gly Phe Arg Asp
 85 90 95
 Thr Ala Phe Lys Glu Met Gln Pro Leu Leu Leu Thr Leu Phe Arg Glu
 100 105 110
 Glu Lys Arg Cys Gly Phe Lys Gly Ile Gly Gly Val Arg Arg Ser Leu
 115 120 125
 Arg Ile Leu Leu Phe Val Lys Phe Phe Thr Leu Ser Trp Leu Cys Val
 130 135 140
 Thr Asp Val Leu Phe Leu Leu Tyr Ser Thr Asp Ala Leu Ile Trp Val
 145 150 155 160
 Asn Val Leu Arg Phe Phe Phe Lys Cys Asn Thr Asn Asn Ile Leu Glu
 165 170 175
 Met Val Pro Met Gly Tyr Phe Leu Ala Leu Trp His Ile Ala Arg Gly
 180 185 190
 Phe Asp Cys Val Asn Arg Arg Leu Asp Gln Ile Val Lys Ser Lys Ser
 195 200 205
 Thr Arg Lys His Arg Glu Leu Gln His Leu Trp Leu Leu His Ala Cys
 210 215 220
 Leu Thr Lys Thr Ala Leu Asn Ile Asn Lys Ile Tyr Ala Pro Gln Met
 225 230 235 240
 Leu Ala Ser Arg Phe Asp Asn Phe Val Asn Gly Val Ile Gln Ala Tyr
 245 250 255
 Trp Gly Ala Val Phe Thr Phe Asp Leu Ser Thr Pro Phe Phe Trp Val
 260 265 270
 Val Tyr Gly Ser Val Gln Tyr His Val Arg Cys Leu Asp Tyr Tyr Leu
 275 280 285
 Ile Asp Asn Met Cys Asp Val Ala Val Glu Tyr His Asp Ser Ala Lys
 290 295 300
 His Ser Trp Ser Glu Val Arg Trp Thr Lys Glu Val Ser Ala Phe Gly
 305 310 315 320
 Ser Ile Leu Leu Tyr Ile Cys Met Leu Met Gln Leu Leu Ser Phe Gln
 325 330 335

Ile Ser Ser Tyr Val Ile Tyr Ala Asn Ser Thr Lys Leu Gln Leu Trp
 340 345 350

Ser Cys Gly Leu Phe Gln Ala Asn Arg Ser Met Trp Phe Ala Met Ile
 355 360 365

Ser Ser Val Leu Tyr Tyr Ile Leu Val Leu Leu Gln Phe His Leu Val
 370 375 380

Met Arg Lys
 385

<210> 57
 <211> 1170
 <212> DNA
 <213> Drosophila melanogaster

<220>
 <221> CDS
 <222> (1)..(1170)
 <223> Coding region GR59D.1

<220>
 <221> misc_feature
 <222> (1021)
 <223> Splice site

<400> 57
 atg gct gac ctg ctc aaa ttg tgt ttg aga atc gca tat gcc tac gga 48
 Met Ala Asp Leu Leu Lys Leu Cys Leu Arg Ile Ala Tyr Ala Tyr Gly
 1 5 10 15

cgg ttg acc ggc gta atc aac ttt aag att gat ttg aaa acg ggt caa 96
 Arg Leu Thr Gly Val Ile Asn Phe Lys Ile Asp Leu Lys Thr Gly Gln
 20 25 30

gcg cta gtt acc aga gga gct acg ctt att tca gtg agc aca cac ttg 144
 Ala Leu Val Thr Arg Gly Ala Thr Leu Ile Ser Val Ser Thr His Leu
 35 40 45

ctg atc ttt gcc cta ctc ctc tac caa aca atg cga aaa agt gtg gtg 192
 Leu Ile Phe Ala Leu Leu Leu Tyr Gln Thr Met Arg Lys Ser Val Val
 50 55 60

aac gtc atg tgg aag tat gcc aat tcc ctt cac gaa tac gtg ttc ctg 240
 Asn Val Met Trp Lys Tyr Ala Asn Ser Leu His Glu Tyr Val Phe Leu
 65 70 75 80

gtc ata gcc ggc ttt cgc gta gtc tgt gtc ttt ctg gag ctg gtc agt 288
 Val Ile Ala Gly Phe Arg Val Val Cys Val Phe Leu Glu Leu Val Ser
 85 90 95

cga tgg tcg cag cgt cgc acc ttt gtg agg ctc ttc aac tca ttc cgg 336
 Arg Trp Ser Gln Arg Arg Thr Phe Val Arg Leu Phe Asn Ser Phe Arg

	100	105	110	
aga ctg tat cag aga aat ccg gat ata atc cag tac tgc cga agg agc				384
Arg Leu Tyr Gln Arg Asn Pro Asp Ile Ile Gln Tyr Cys Arg Arg Ser	115	120	125	
atc gtt agc aaa ttc ttt tgc gtc aca atg aca gag aca ctg cac atc				432
Ile Val Ser Lys Phe Phe Cys Val Thr Met Thr Glu Thr Leu His Ile	130	135	140	
ata gtc acc ttg gcc atg atg agg aat cgg ctg agc att gct ctg gct				480
Ile Val Thr Leu Ala Met Met Arg Asn Arg Leu Ser Ile Ala Leu Ala	145	150	155	160
ctg cgc att tgg gcc gta ttg agt ctg acg gcc ata ata aat gta atc				528
Leu Arg Ile Trp Ala Val Leu Ser Leu Thr Ala Ile Ile Asn Val Ile	165	170	175	
atc acg cag tac tat gtg gcc acc gcg tgt gtg cga gga cga tat gcg				576
Ile Thr Gln Tyr Tyr Val Ala Thr Ala Cys Val Arg Gly Arg Tyr Ala	180	185	190	
ctc ctg aac aag gat ctt cag gcg att gtg acc gaa tcc cag tcg ctg				624
Leu Leu Asn Lys Asp Leu Gln Ala Ile Val Thr Glu Ser Gln Ser Leu	195	200	205	
gtt ccc aac gga ggt ggc gtc ttt gtg acc aag tgt tgc tac cta gcg				672
Val Pro Asn Gly Gly Gly Val Phe Val Thr Lys Cys Cys Tyr Leu Ala	210	215	220	
gat cgc ttg gag cga ata gcc aag tcc cag tcg gac cta cag gag ctc				720
Asp Arg Leu Glu Arg Ile Ala Lys Ser Gln Ser Asp Leu Gln Glu Leu	225	230	235	240
gtc gaa aac ttg tcc acg gca tac gaa gga gaa gtg gtc tgc ctg gtc				768
Val Glu Asn Leu Ser Thr Ala Tyr Glu Gly Glu Val Val Cys Leu Val	245	250	255	
atc aca tac tat ctg aat atg ctg ggc acc tcg tat ctg ctg ttc agc				816
Ile Thr Tyr Tyr Leu Asn Met Leu Gly Thr Ser Tyr Leu Leu Phe Ser	260	265	270	
att agc aag tat ggc aat ttt ggg aat aac ctg ctc gtg atc atc act				864
Ile Ser Lys Tyr Gly Asn Phe Gly Asn Asn Leu Leu Val Ile Ile Thr	275	280	285	
ctt tgt ggc att gtc tac ttc gta ttt tac gtc gtc gat tgc tgg atc				912
Leu Cys Gly Ile Val Tyr Phe Val Phe Tyr Val Val Asp Cys Trp Ile	290	295	300	
aac gcg ttt aat gtg ttt tac ctt ttg gat gcc cat gat aag atg gtt				960
Asn Ala Phe Asn Val Phe Tyr Leu Leu Asp Ala His Asp Lys Met Val	305	310	315	320
aag ttg ctg aat aag cga act ttg ttt cag cca ggt ctg gat cat cga				1008
Lys Leu Leu Asn Lys Arg Thr Leu Phe Gln Pro Gly Leu Asp His Arg				

325	330	335	
ttg gaa atg gtt ttt gaa aac ttt gct ctg aac ttg gtg cgg aat cca			1056
Leu Glu Met Val Phe Glu Asn Phe Ala Leu Asn Leu Val Arg Asn Pro			
340	345	350	
ttg aag ctc cat atg tac ggc ctt ttc gag ttt ggt cga gga aca tcc			1104
Leu Lys Leu His Met Tyr Gly Leu Phe Glu Phe Gly Arg Gly Thr Ser			
355	360	365	
ttt gcc gtg ttt aac tcc ctg tta aca cac tcc ctt ctc ctc att caa			1152
Phe Ala Val Phe Asn Ser Leu Leu Thr His Ser Leu Leu Ile Gln			
370	375	380	
tac gac gtg caa aac ttc			1170
Tyr Asp Val Gln Asn Phe			
385	390		

<210> 58

<211> 390

<212> PRT

<213> *Drosophila melanogaster*

<400> 58

Met Ala Asp Leu Leu Lys Leu Cys Leu Arg Ile Ala Tyr Ala Tyr Gly			
1	5	10	15
Arg Leu Thr Gly Val Ile Asn Phe Lys Ile Asp Leu Lys Thr Gly Gln			
20	25	30	
Ala Leu Val Thr Arg Gly Ala Thr Leu Ile Ser Val Ser Thr His Leu			
35	40	45	
Leu Ile Phe Ala Leu Leu Leu Tyr Gln Thr Met Arg Lys Ser Val Val			
50	55	60	
Asn Val Met Trp Lys Tyr Ala Asn Ser Leu His Glu Tyr Val Phe Leu			
65	70	75	80
Val Ile Ala Gly Phe Arg Val Val Cys Val Phe Leu Glu Leu Val Ser			
85	90	95	
Arg Trp Ser Gln Arg Arg Thr Phe Val Arg Leu Phe Asn Ser Phe Arg			
100	105	110	
Arg Leu Tyr Gln Arg Asn Pro Asp Ile Ile Gln Tyr Cys Arg Arg Ser			
115	120	125	
Ile Val Ser Lys Phe Phe Cys Val Thr Met Thr Glu Thr Leu His Ile			
130	135	140	
Ile Val Thr Leu Ala Met Met Arg Asn Arg Leu Ser Ile Ala Leu Ala			
145	150	155	160
Leu Arg Ile Trp Ala Val Leu Ser Leu Thr Ala Ile Ile Asn Val Ile			

```

                165                170                175
Ile Thr Gln Tyr Tyr Val Ala Thr Ala Cys Val Arg Gly Arg Tyr Ala
      180                185                190
Leu Leu Asn Lys Asp Leu Gln Ala Ile Val Thr Glu Ser Gln Ser Leu
      195                200                205
Val Pro Asn Gly Gly Gly Val Phe Val Thr Lys Cys Cys Tyr Leu Ala
      210                215                220
Asp Arg Leu Glu Arg Ile Ala Lys Ser Gln Ser Asp Leu Gln Glu Leu
      225                230                235                240
Val Glu Asn Leu Ser Thr Ala Tyr Glu Gly Glu Val Val Cys Leu Val
      245                250                255
Ile Thr Tyr Tyr Leu Asn Met Leu Gly Thr Ser Tyr Leu Leu Phe Ser
      260                265                270
Ile Ser Lys Tyr Gly Asn Phe Gly Asn Asn Leu Leu Val Ile Ile Thr
      275                280                285
Leu Cys Gly Ile Val Tyr Phe Val Phe Tyr Val Val Asp Cys Trp Ile
      290                295                300
Asn Ala Phe Asn Val Phe Tyr Leu Leu Asp Ala His Asp Lys Met Val
      305                310                315                320
Lys Leu Leu Asn Lys Arg Thr Leu Phe Gln Pro Gly Leu Asp His Arg
      325                330                335
Leu Glu Met Val Phe Glu Asn Phe Ala Leu Asn Leu Val Arg Asn Pro
      340                345                350
Leu Lys Leu His Met Tyr Gly Leu Phe Glu Phe Gly Arg Gly Thr Ser
      355                360                365
Phe Ala Val Phe Asn Ser Leu Leu Thr His Ser Leu Leu Leu Ile Gln
      370                375                380
Tyr Asp Val Gln Asn Phe
      385                390

```

```

<210> 59
<211> 1191
<212> DNA
<213> Drosophila melanogaster

```

```

<220>
<221> CDS
<222> (1)..(1191)
<223> Coding region GR59D.2

```

<220>

<221> misc feature

<222> (1036)

<223> Splice site

<400> 59

atg gtt gac ttg gtg aag acg att ttg ctc att gcc tac tgg tat ggc 48
 Met Val Asp Leu Val Lys Thr Ile Leu Leu Ile Ala Tyr Trp Tyr Gly
 1 5 10 15

ctt gcc gtg gga gtg tcc aac ttc gag gtg gac tgg cta act gga gaa 96
 Leu Ala Val Gly Val Ser Asn Phe Glu Val Asp Trp Leu Thr Gly Glu
 20 25 30

gcc att gcc acc cga agg act acg atc tat gca gca gtg cat aat gcc 144
 Ala Ile Ala Thr Arg Arg Thr Thr Ile Tyr Ala Ala Val His Asn Ala
 35 40 45

agc ctt atc act ctg ctg att ctt ttc aat ctt ggc aat aac tca ctg 192
 Ser Leu Ile Thr Leu Leu Ile Leu Phe Asn Leu Gly Asn Asn Ser Leu
 50 55 60

aaa tcc gag ttc ata agt gct cga tat ctg cat gag tac ttc ttt atg 240
 Lys Ser Glu Phe Ile Ser Ala Arg Tyr Leu His Glu Tyr Phe Phe Met
 65 70 75 80

ctc atg act gcg gtt cga atc tcg gca gtt ctg ctc tca ctg ata acc 288
 Leu Met Thr Ala Val Arg Ile Ser Ala Val Leu Leu Ser Leu Ile Thr
 85 90 95

aga tgg tat cag cgt tcc aga ttc att cga att tgg aat cag ata cta 336
 Arg Trp Tyr Gln Arg Ser Arg Phe Ile Arg Ile Trp Asn Gln Ile Leu
 100 105 110

gcc cta gtt cgc gat aga cct caa gtg gtt cgt ggg cgc tgg tat cgt 384
 Ala Leu Val Arg Asp Arg Pro Gln Val Val Arg Gly Arg Trp Tyr Arg
 115 120 125

cgc agc att atc ctt aaa ttt gtg ttc tgt gtc ctg tca gat tct ctg 432
 Arg Ser Ile Ile Leu Lys Phe Val Phe Cys Val Leu Ser Asp Ser Leu
 130 135 140

cac acc ata tcg gat gtg agt gcg caa cga aag cgg atc act gct gac 480
 His Thr Ile Ser Asp Val Ser Ala Gln Arg Lys Arg Ile Thr Ala Asp
 145 150 155 160

ctg att gtc aaa ctg agc tta ctg gcc aca ctg acc acc att ttt aac 528
 Leu Ile Val Lys Leu Ser Leu Leu Ala Thr Leu Thr Thr Ile Phe Asn
 165 170 175

atg atc gtg tgc cag tac tac ttg gcc atg gtg cag gtg att ggg ctc 576
 Met Ile Val Cys Gln Tyr Tyr Leu Ala Met Val Gln Val Ile Gly Leu
 180 185 190

tac aag att ctg ctc caa gat ctg cga tgc ttg gtg cgc caa gct gaa 624
 Tyr Lys Ile Leu Leu Gln Asp Leu Arg Cys Leu Val Arg Gln Ala Glu

195	200	205	
tgc atc tgc tcc att cgc aat cgg cga ggt gga gtt tac tcc att cag	672		
Cys Ile Cys Ser Ile Arg Asn Arg Arg Gly Gly Val Tyr Ser Ile Gln			
210 215 220			
tgc tgc tgc ttg gca gat cag ctg gat cta att gcc gaa agg cat tac	720		
Cys Cys Ser Leu Ala Asp Gln Leu Asp Leu Ile Ala Glu Arg His Tyr			
225 230 235 240			
ttt ctg aag gac aga ctt gat gag atg tcg gac ctt ttc cag ata cag	768		
Phe Leu Lys Asp Arg Leu Asp Glu Met Ser Asp Leu Phe Gln Ile Gln			
245 250 255			
agc cta agc atg agc ctg gtg tac ttt ttc tcc acc atg ggc tcc atc	816		
Ser Leu Ser Met Ser Leu Val Tyr Phe Phe Ser Thr Met Gly Ser Ile			
260 265 270			
tac ttt agc gtc tgt tgc atc ctg tac agc tcc aca gga ttc ggc tct	864		
Tyr Phe Ser Val Cys Ser Ile Leu Tyr Ser Ser Thr Gly Phe Gly Ser			
275 280 285			
aca tac tgg ggt ctt ctg ctg att gta cta tcc acg gct tcc ttc tac	912		
Thr Tyr Trp Gly Leu Leu Leu Ile Val Leu Ser Thr Ala Ser Phe Tyr			
290 295 300			
atg gac aat tgg ttg tcc gtt aac att ggg ttt cat att cga gat cag	960		
Met Asp Asn Trp Leu Ser Val Asn Ile Gly Phe His Ile Arg Asp Gln			
305 310 315 320			
cag gac gaa cta ttc cga gtg ctg gcg gat cga act ctg ttc tat cgg	1008		
Gln Asp Glu Leu Phe Arg Val Leu Ala Asp Arg Thr Leu Phe Tyr Arg			
325 330 335			
gaa ttg gac aac cga ctg gag gca gcc ttt gag aac ttc caa ctg caa	1056		
Glu Leu Asp Asn Arg Leu Glu Ala Ala Phe Glu Asn Phe Gln Leu Gln			
340 345 350			
ctg gcc agt aac cgg cat gaa ttc tac gtt atg ggt ctc ttt aaa atg	1104		
Leu Ala Ser Asn Arg His Glu Phe Tyr Val Met Gly Leu Phe Lys Met			
355 360 365			
gaa cgt ggt cgt cta atc gct atg cta agc tca gtg atc act cat act	1152		
Glu Arg Gly Arg Leu Ile Ala Met Leu Ser Ser Val Ile Thr His Thr			
370 375 380			
atg gtt ctt gtt cag tgg gaa att caa aac gat gaa tcg	1191		
Met Val Leu Val Gln Trp Glu Ile Gln Asn Asp Glu Ser			
385 390 395			

<210> 60

<211> 397

<212> PRT

<213> Drosophila melanogaster

<400> 60

```

Met Val Asp Leu Val Lys Thr Ile Leu Leu Ile Ala Tyr Trp Tyr Gly
  1           5           10           15

Leu Ala Val Gly Val Ser Asn Phe Glu Val Asp Trp Leu Thr Gly Glu
      20           25           30

Ala Ile Ala Thr Arg Arg Thr Thr Ile Tyr Ala Ala Val His Asn Ala
      35           40           45

Ser Leu Ile Thr Leu Leu Ile Leu Phe Asn Leu Gly Asn Asn Ser Leu
      50           55           60

Lys Ser Glu Phe Ile Ser Ala Arg Tyr Leu His Glu Tyr Phe Phe Met
      65           70           75           80

Leu Met Thr Ala Val Arg Ile Ser Ala Val Leu Leu Ser Leu Ile Thr
      85           90           95

Arg Trp Tyr Gln Arg Ser Arg Phe Ile Arg Ile Trp Asn Gln Ile Leu
      100          105          110

Ala Leu Val Arg Asp Arg Pro Gln Val Val Arg Gly Arg Trp Tyr Arg
      115          120          125

Arg Ser Ile Ile Leu Lys Phe Val Phe Cys Val Leu Ser Asp Ser Leu
      130          135          140

His Thr Ile Ser Asp Val Ser Ala Gln Arg Lys Arg Ile Thr Ala Asp
      145          150          155          160

Leu Ile Val Lys Leu Ser Leu Leu Ala Thr Leu Thr Thr Ile Phe Asn
      165          170          175

Met Ile Val Cys Gln Tyr Tyr Leu Ala Met Val Gln Val Ile Gly Leu
      180          185          190

Tyr Lys Ile Leu Leu Gln Asp Leu Arg Cys Leu Val Arg Gln Ala Glu
      195          200          205

Cys Ile Cys Ser Ile Arg Asn Arg Arg Gly Gly Val Tyr Ser Ile Gln
      210          215          220

Cys Cys Ser Leu Ala Asp Gln Leu Asp Leu Ile Ala Glu Arg His Tyr
      225          230          235          240

Phe Leu Lys Asp Arg Leu Asp Glu Met Ser Asp Leu Phe Gln Ile Gln
      245          250          255

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

Tyr Phe Ser Val Cys Ser Ile Leu Tyr Ser Ser Thr Gly Phe Gly Ser
      275          280          285

Thr Tyr Trp Gly Leu Leu Leu Ile Val Leu Ser Thr Ala Ser Phe Tyr

```

105

```

      290              295              300
Met Asp Asn Trp Leu Ser Val Asn Ile Gly Phe His Ile Arg Asp Gln
305              310              315              320
Gln Asp Glu Leu Phe Arg Val Leu Ala Asp Arg Thr Leu Phe Tyr Arg
325              330              335
Glu Leu Asp Asn Arg Leu Glu Ala Ala Phe Glu Asn Phe Gln Leu Gln
340              345              350
Leu Ala Ser Asn Arg His Glu Phe Tyr Val Met Gly Leu Phe Lys Met
355              360              365
Glu Arg Gly Arg Leu Ile Ala Met Leu Ser Ser Val Ile Thr His Thr
370              375              380
Met Val Leu Val Gln Trp Glu Ile Gln Asn Asp Glu Ser
385              390              395

```

```

<210> 61
<211> 1089
<212> DNA
<213> Drosophila melanogaster

```

```

<220>
<221> CDS
<222> (1)..(1089)
<223> Coding region GR59E.1

```

```

<220>
<221> misc_feature
<222> (52)..(988)
<223> Splice sites before positions 52, 835 and 988

```

```

<400> 61
atg cgg agc tca gca aca aaa ggc gct aaa ctt aag aat tcc ccg agg 48
Met Arg Ser Ser Ala Thr Lys Gly Ala Lys Leu Lys Asn Ser Pro Arg
1 5 10 15
gaa aga ttc tat cga ttg gtc cac ttg agt tgg atg att ctg tgg tac 96
Glu Arg Phe Tyr Arg Leu Val His Leu Ser Trp Met Ile Leu Trp Tyr
20 25 30
gga ttg ttc gtc ctg ggc tcc tac tgg gaa ttc gtg ctc gtc acc aca 144
Gly Leu Phe Val Leu Gly Ser Tyr Trp Glu Phe Val Leu Val Thr Thr
35 40 45
cag agg gta tcc ttg gac cgc tac cta aat gct atc gag tcc gcc atc 192
Gln Arg Val Ser Leu Asp Arg Tyr Leu Asn Ala Ile Glu Ser Ala Ile
50 55 60
tat gta gtt cac ata ttt tcc atc atg ctc ctc acc tgg cag tgc agg 240
Tyr Val Val His Ile Phe Ser Ile Met Leu Leu Thr Trp Gln Cys Arg

```

65	70	75	80	
aac tgg gct ccc aag ctg atg acg aac atc gtt aca tcc gac ctg aat				288
Asn Trp Ala Pro Lys Leu Met Thr Asn Ile Val Thr Ser Asp Leu Asn	85	90	95	
aga gct tat acc ata gac tgc aac aga acc aag aga ttt att agg ctc				336
Arg Ala Tyr Thr Ile Asp Cys Asn Arg Thr Lys Arg Phe Ile Arg Leu	100	105	110	
cag ctg ttt ctg gtg gga att ttt gcc tgt cta gcg atc ttc ttc aac				384
Gln Leu Phe Leu Val Gly Ile Phe Ala Cys Leu Ala Ile Phe Phe Asn	115	120	125	
atc tgg acc cac aaa ttc gtt gta tat aga tgc atc ctg agc atc aac				432
Ile Trp Thr His Lys Phe Val Val Tyr Arg Ser Ile Leu Ser Ile Asn	130	135	140	
agc tac gtg atg ccc aat atc ata agt agc ata tcc ttc gcc cag tac				480
Ser Tyr Val Met Pro Asn Ile Ile Ser Ser Ile Ser Phe Ala Gln Tyr	145	150	155	160
tac ctg ctg ctt caa gga atc gct tgg cgt caa agg agg cta act gaa				528
Tyr Leu Leu Leu Gln Gly Ile Ala Trp Arg Gln Arg Arg Leu Thr Glu	165	170	175	
ggc ctg gag cgg gaa cta act cat ttg cac agt ccc cga ata tca gag				576
Gly Leu Glu Arg Glu Leu Thr His Leu His Ser Pro Arg Ile Ser Glu	180	185	190	
gtc cag aaa atc cga atg cat cac gcg aat cta att gat ttt acc aag				624
Val Gln Lys Ile Arg Met His His Ala Asn Leu Ile Asp Phe Thr Lys	195	200	205	
gcg gtg aac cgg acg ttt cag tac tcc atc cta ctg ctc ttc gtg gga				672
Ala Val Asn Arg Thr Phe Gln Tyr Ser Ile Leu Leu Leu Phe Val Gly	210	215	220	
tgt ttc ctg aac ttc aac ctc gtc ctg ttc ctc gtc tat cag ggc att				720
Cys Phe Leu Asn Phe Asn Leu Val Leu Phe Leu Val Tyr Gln Gly Ile	225	230	235	240
gag aat cct tcc atg gcc gac ttc acc aag tgg gta tgc atg ctt ctc				768
Glu Asn Pro Ser Met Ala Asp Phe Thr Lys Trp Val Cys Met Leu Leu	245	250	255	
tgg ctg gcc atg cac gtg gga aag gtg tgc agc atc ctg cat ttc aac				816
Trp Leu Ala Met His Val Gly Lys Val Cys Ser Ile Leu His Phe Asn	260	265	270	
cag agt atc caa aat gag cat tcc acg tgc ttg acc tta ttg agt aga				864
Gln Ser Ile Gln Asn Glu His Ser Thr Cys Leu Thr Leu Leu Ser Arg	275	280	285	
gtg tgc tat gcc aga aaa gat att cag gat acc ata acc cat ttc atc				912
Val Ser Tyr Ala Arg Lys Asp Ile Gln Asp Thr Ile Thr His Phe Ile				

290 295 300

atc caa atg cgg acg aat gtg cgg cag cat gtc gtg tgt gga gtc ata 960
 Ile Gln Met Arg Thr Asn Val Arg Gln His Val Val Cys Gly Val Ile
 305 310 315 320

aat ctg gac ctc aaa ttc tta acc act ttg ttg gtg gcc tct gct gat 1008
 Asn Leu Asp Leu Lys Phe Leu Thr Thr Leu Leu Val Ala Ser Ala Asp
 325 330 335

ttc ttc atc ttc ctg ctg caa tac gat gtc acc tac gaa gcg ttg tcc 1056
 Phe Phe Ile Phe Leu Leu Gln Tyr Asp Val Thr Tyr Glu Ala Leu Ser
 340 345 350

aag tcg gtt cag ggc aat gtc act aga tat aag 1089
 Lys Ser Val Gln Gly Asn Val Thr Arg Tyr Lys
 355 360

<210> 62
 <211> 363
 <212> PRT
 <213> Drosophila melanogaster

<400> 62
 Met Arg Ser Ser Ala Thr Lys Gly Ala Lys Leu Lys Asn Ser Pro Arg
 1 5 10 15

Glu Arg Phe Tyr Arg Leu Val His Leu Ser Trp Met Ile Leu Trp Tyr
 20 25 30

Gly Leu Phe Val Leu Gly Ser Tyr Trp Glu Phe Val Leu Val Thr Thr
 35 40 45

Gln Arg Val Ser Leu Asp Arg Tyr Leu Asn Ala Ile Glu Ser Ala Ile
 50 55 60

Tyr Val Val His Ile Phe Ser Ile Met Leu Leu Thr Trp Gln Cys Arg
 65 70 75 80

Asn Trp Ala Pro Lys Leu Met Thr Asn Ile Val Thr Ser Asp Leu Asn
 85 90 95

Arg Ala Tyr Thr Ile Asp Cys Asn Arg Thr Lys Arg Phe Ile Arg Leu
 100 105 110

Gln Leu Phe Leu Val Gly Ile Phe Ala Cys Leu Ala Ile Phe Phe Asn
 115 120 125

Ile Trp Thr His Lys Phe Val Val Tyr Arg Ser Ile Leu Ser Ile Asn
 130 135 140

Ser Tyr Val Met Pro Asn Ile Ile Ser Ser Ile Ser Phe Ala Gln Tyr
 145 150 155 160

Tyr Leu Leu Leu Gln Gly Ile Ala Trp Arg Gln Arg Arg Leu Thr Glu

108

```

          165          170          175
Gly Leu Glu Arg Glu Leu Thr His Leu His Ser Pro Arg Ile Ser Glu
          180          185          190
Val Gln Lys Ile Arg Met His His Ala Asn Leu Ile Asp Phe Thr Lys
          195          200          205
Ala Val Asn Arg Thr Phe Gln Tyr Ser Ile Leu Leu Leu Phe Val Gly
          210          215          220
Cys Phe Leu Asn Phe Asn Leu Val Leu Phe Leu Val Tyr Gln Gly Ile
          225          230          235          240
Glu Asn Pro Ser Met Ala Asp Phe Thr Lys Trp Val Cys Met Leu Leu
          245          250          255
Trp Leu Ala Met His Val Gly Lys Val Cys Ser Ile Leu His Phe Asn
          260          265          270
Gln Ser Ile Gln Asn Glu His Ser Thr Cys Leu Thr Leu Leu Ser Arg
          275          280          285
Val Ser Tyr Ala Arg Lys Asp Ile Gln Asp Thr Ile Thr His Phe Ile
          290          295          300
Ile Gln Met Arg Thr Asn Val Arg Gln His Val Val Cys Gly Val Ile
          305          310          315          320
Asn Leu Asp Leu Lys Phe Leu Thr Thr Leu Leu Val Ala Ser Ala Asp
          325          330          335
Phe Phe Ile Phe Leu Leu Gln Tyr Asp Val Thr Tyr Glu Ala Leu Ser
          340          345          350
Lys Ser Val Gln Gly Asn Val Thr Arg Tyr Lys
          355          360

```

```

<210> 63
<211> 1197
<212> DNA
<213> Drosophila melanogaster

<220>
<221> CDS
<222> (1)..(1197)
<223> Coding region GR59E.2

<220>
<221> misc_feature
<222> (937)..(1093)
<223> Splice sites before positions 937 and 1093

<400> 63

```

```

atg gac agt tgg tac tgg gag aat ctg ctg ctg acc atc aat cgg ttc 48
Met Asp Ser Ser Tyr Trp Glu Asn Leu Leu Leu Thr Ile Asn Arg Phe
1 5 10 15

ctg ggc gtg tat ccc agt ggg aga gtg ggc gta ctc cgc tgg ctc cac 96
Leu Gly Val Tyr Pro Ser Gly Arg Val Gly Val Leu Arg Trp Leu His
20 25 30

acg ctc tgg agc ctg ttc ctg ctt atg tac atc tgg act ggc agc att 144
Thr Leu Trp Ser Leu Phe Leu Leu Met Tyr Ile Trp Thr Gly Ser Ile
35 40 45

gtt aag tgc ttg gag ttc aca gtg gag ata ccc act att gaa aaa ctg 192
Val Lys Cys Leu Glu Phe Thr Val Glu Ile Pro Thr Ile Glu Lys Leu
50 55 60

ctc tat ctg atg gag ttc cct gga aat atg gcc acc att gcc atc ctg 240
Leu Tyr Leu Met Glu Phe Pro Gly Asn Met Ala Thr Ile Ala Ile Leu
65 70 75 80

gta tac tat gcc gta ttg aac cgt cca ctt gct cac gga gcg gaa ctc 288
Val Tyr Tyr Ala Val Leu Asn Arg Pro Leu Ala His Gly Ala Glu Leu
85 90 95

cag att gag cgg atc atc aca gga ctc aaa ggc aaa gct aag cga ctg 336
Gln Ile Glu Arg Ile Ile Thr Gly Leu Lys Gly Lys Ala Lys Arg Leu
100 105 110

gtt tat aag aga cat ggt cag agg act ctt cat ctg atg gcg acc act 384
Val Tyr Lys Arg His Gly Gln Arg Thr Leu His Leu Met Ala Thr Thr
115 120 125

tta gtc ttc cat ggc ctg tgt gtc ctg gtt gac gtg gtc aac tat gac 432
Leu Val Phe His Gly Leu Cys Val Leu Val Asp Val Val Asn Tyr Asp
130 135 140

ttc gag ttc tgg acc act tgg agc agt aac agt gtc tac aat ttg cct 480
Phe Glu Phe Trp Thr Thr Trp Ser Ser Asn Ser Val Tyr Asn Leu Pro
145 150 155 160

ggg cta atg atg agt ctg ggg gtg ctc cag tat gcc cag ccc gtc cat 528
Gly Leu Met Met Ser Leu Gly Val Leu Gln Tyr Ala Gln Pro Val His
165 170 175

ttc cta tgg ctg gtg atg gac cag atg agg atg tgt ctc aag gag ctg 576
Phe Leu Trp Leu Val Met Asp Gln Met Arg Met Cys Leu Lys Glu Leu
180 185 190

aag ttg ctg cag aga cca ccg cag ggt tgc acc aag ttg gat gcc tgc 624
Lys Leu Leu Gln Arg Pro Pro Gln Gly Ser Thr Lys Leu Asp Ala Cys
195 200 205

tac gag tcc gcc ttt gct gtc cta gta gat gcg ggt gga gga tct gct 672
Tyr Glu Ser Ala Phe Ala Val Leu Val Asp Ala Gly Gly Gly Ser Ala
210 215 220

```

```

ctg atg att gaa gag atg aga tac acc tgc aat ctt atc gag cag gtc 720
Leu Met Ile Glu Glu Met Arg Tyr Thr Cys Asn Leu Ile Glu Gln Val
225          230          235          240

cac agc cag ttt cta ctg aga ttt ggt ctc tat ctg gtg tta aac ctg 768
His Ser Gln Phe Leu Leu Arg Phe Gly Leu Tyr Leu Val Leu Asn Leu
245          250          255

ctc aat tcc ttg gtc agc atc tgt gtg gag tta tac ttg atc ttc aac 816
Leu Asn Ser Leu Val Ser Ile Cys Val Glu Leu Tyr Leu Ile Phe Asn
260          265          270

ttc ttt gag act cct ctc tgg gag gag tcc gtg ctc ctc gta tac cgc 864
Phe Phe Glu Thr Pro Leu Trp Glu Glu Ser Val Leu Leu Val Tyr Arg
275          280          285

ctg ctc tgg ttg gcc atg cat ggc ggt cgt atc tgg ttc atc ctg tcg 912
Leu Leu Trp Leu Ala Met His Gly Gly Arg Ile Trp Phe Ile Leu Ser
290          295          300

gtc aac gaa cag atc ctg gag cag aaa tgt aac ctt tgc cag ctg ctc 960
Val Asn Glu Gln Ile Leu Glu Gln Lys Cys Asn Leu Cys Gln Leu Leu
305          310          315          320

aac gag ttg gag gtc tgc agc tcc cgc ctg cag agg acc att aat cgc 1008
Asn Glu Leu Glu Val Cys Ser Ser Arg Leu Gln Arg Thr Ile Asn Arg
325          330          335

ttc ctg ctg cag ctc cag cgg agt att gac cag ccc ctg gaa gcc tgc 1056
Phe Leu Leu Gln Leu Gln Arg Ser Ile Asp Gln Pro Leu Glu Ala Cys
340          345          350

gga atc gtt acg ctg gac act cgc tca ttg ggc ggg ttt atc ggc gtc 1104
Gly Ile Val Thr Leu Asp Thr Arg Ser Leu Gly Gly Phe Ile Gly Val
355          360          365

ctg atg gcc att gtc att ttc ctc att caa atc gga ctg ggc aat aag 1152
Leu Met Ala Ile Val Ile Phe Leu Ile Gln Ile Gly Leu Gly Asn Lys
370          375          380

tcg cta atg ggc gtt gcc ctc aac aga tcc aat tgg gtt tac gtt 1197
Ser Leu Met Gly Val Ala Leu Asn Arg Ser Asn Trp Val Tyr Val
385          390          395

```

<210> 64
 <211> 399
 <212> PRT
 <213> Drosophila melanogaster

```

<400> 64
Met Asp Ser Ser Tyr Trp Glu Asn Leu Leu Leu Thr Ile Asn Arg Phe
1          5          10          15

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

```

Thr Leu Trp Ser Leu Phe Leu Leu Met Tyr Ile Trp Thr Gly Ser Ile
 35 40 45
 Val Lys Cys Leu Glu Phe Thr Val Glu Ile Pro Thr Ile Glu Lys Leu
 50 55 60
 Leu Tyr Leu Met Glu Phe Pro Gly Asn Met Ala Thr Ile Ala Ile Leu
 65 70 75 80
 Val Tyr Tyr Ala Val Leu Asn Arg Pro Leu Ala His Gly Ala Glu Leu
 85 90 95
 Gln Ile Glu Arg Ile Ile Thr Gly Leu Lys Gly Lys Ala Lys Arg Leu
 100 105 110
 Val Tyr Lys Arg His Gly Gln Arg Thr Leu His Leu Met Ala Thr Thr
 115 120 125
 Leu Val Phe His Gly Leu Cys Val Leu Val Asp Val Val Asn Tyr Asp
 130 135 140
 Phe Glu Phe Trp Thr Thr Trp Ser Ser Asn Ser Val Tyr Asn Leu Pro
 145 150 155 160
 Gly Leu Met Met Ser Leu Gly Val Leu Gln Tyr Ala Gln Pro Val His
 165 170 175
 Phe Leu Trp Leu Val Met Asp Gln Met Arg Met Cys Leu Lys Glu Leu
 180 185 190
 Lys Leu Leu Gln Arg Pro Pro Gln Gly Ser Thr Lys Leu Asp Ala Cys
 195 200 205
 Tyr Glu Ser Ala Phe Ala Val Leu Val Asp Ala Gly Gly Gly Ser Ala
 210 215 220
 Leu Met Ile Glu Glu Met Arg Tyr Thr Cys Asn Leu Ile Glu Gln Val
 225 230 235 240
 His Ser Gln Phe Leu Leu Arg Phe Gly Leu Tyr Leu Val Leu Asn Leu
 245 250 255
 Leu Asn Ser Leu Val Ser Ile Cys Val Glu Leu Tyr Leu Ile Phe Asn
 260 265 270
 Phe Phe Glu Thr Pro Leu Trp Glu Glu Ser Val Leu Leu Val Tyr Arg
 275 280 285
 Leu Leu Trp Leu Ala Met His Gly Gly Arg Ile Trp Phe Ile Leu Ser
 290 295 300
 Val Asn Glu Gln Ile Leu Glu Gln Lys Cys Asn Leu Cys Gln Leu Leu
 305 310 315 320
 Asn Glu Leu Glu Val Cys Ser Ser Arg Leu Gln Arg Thr Ile Asn Arg
 325 330 335

Phe Leu Leu Gln Leu Gln Arg Ser Ile Asp Gln Pro Leu Glu Ala Cys
 340 345 350

Gly Ile Val Thr Leu Asp Thr Arg Ser Leu Gly Gly Phe Ile Gly Val
 355 360 365

Leu Met Ala Ile Val Ile Phe Leu Ile Gln Ile Gly Leu Gly Asn Lys
 370 375 380

Ser Leu Met Gly Val Ala Leu Asn Arg Ser Asn Trp Val Tyr Val
 385 390 395

<210> 65
 <211> 1317
 <212> DNA
 <213> Drosophila melanogaster

<220>
 <221> CDS
 <222> (1) .. (1317)
 <223> Coding region GR61D.1

<400> 65
 atg tcg agg act tcg gat gac atc cgg aag cac ctg aaa gtg cgg cgc 48
 Met Ser Arg Thr Ser Asp Asp Ile Arg Lys His Leu Lys Val Arg Arg
 1 5 10 15

cag aag cag agg gcc att ttg gcc atg aga tgg cgc tgt gcc cag gga 96
 Gln Lys Gln Arg Ala Ile Leu Ala Met Arg Trp Arg Cys Ala Gln Gly
 20 25 30

gga ttg gag ttc gag cag ctg gac acc ttc tac gga gcc atc agg cca 144
 Gly Leu Glu Phe Glu Gln Leu Asp Thr Phe Tyr Gly Ala Ile Arg Pro
 35 40 45

tgt aag acc act ttt cca aac cac ttt cat atg tgt cta aac ata tta 192
 Cys Lys Thr Thr Phe Pro Asn His Phe His Met Cys Leu Asn Ile Leu
 50 55 60

caa tgg aca ata gat ctg tgt gtg gcg cag ttc ttc ggc att atg cct 240
 Gln Trp Thr Ile Asp Leu Cys Val Ala Gln Phe Phe Gly Ile Met Pro
 65 70 75 80

ttg tcg aat atc cgg agt cgc gat ccc cag gat gtg aag ttc aag gta 288
 Leu Ser Asn Ile Arg Ser Arg Asp Pro Gln Asp Val Lys Phe Lys Val
 85 90 95

agg agc att ggc ctg gca gtc acc ggt ctc ttc ctt ttg ctc ggt ggc 336
 Arg Ser Ile Gly Leu Ala Val Thr Gly Leu Phe Leu Leu Leu Gly Gly
 100 105 110

atg aag acc ttg gtc ggt gcc aat att ctc ttc aca gag ggt ctg aat 384
 Met Lys Thr Leu Val Gly Ala Asn Ile Leu Phe Thr Glu Gly Leu Asn
 115 120 125

```

gcc aag aat atc gtg ggt ttg gtt ttc ctt atc gtg ggc atg gtc aac 432
Ala Lys Asn Ile Val Gly Leu Val Phe Leu Ile Val Gly Met Val Asn
130 135 140

tgg cta aac ttc gtg ggc ttt gct cgc tcc tgg tcg cac ata atg ctg 480
Trp Leu Asn Phe Val Gly Phe Ala Arg Ser Trp Ser His Ile Met Leu
145 150 155 160

ccc tgg agt tcg gtg gac att ctg atg ctc ttt ccg ccc tac aaa cgt 528
Pro Trp Ser Ser Val Asp Ile Leu Met Leu Phe Pro Pro Tyr Lys Arg
165 170 175

ggc aag cga agc ctt cgg tca aag gtc aac gtc ctt gct ctg agt gtg 576
Gly Lys Arg Ser Leu Arg Ser Lys Val Asn Val Leu Ala Leu Ser Val
180 185 190

gtc gtc ctg gca gcg acc aca tgc tgt act acg cct ctg gct act gca 624
Val Val Leu Ala Ala Thr Thr Cys Cys Thr Thr Pro Leu Ala Thr Ala
195 200 205

gct ata gca tgc aca tcc tgc agt gcc aca caa acc act cgc gta tta 672
Ala Ile Ala Cys Thr Ser Cys Ser Ala Thr Gln Thr Thr Arg Val Leu
210 215 220

cct ttg gac ttt att tgg aga agg agt tct ccg aca tca tgt tca tca 720
Pro Leu Asp Phe Ile Trp Arg Arg Ser Ser Pro Thr Ser Cys Ser Ser
225 230 235 240

tgc cat tca ata tat ttt cca ttg gct aaa tgg agc ctt cac ctt tct 768
Cys His Ser Ile Tyr Phe Pro Leu Ala Lys Trp Ser Leu His Leu Ser
245 250 255

gtg gaa ctt cat gga cat ctt tat tgt gat gac cag tat tgg act ggc 816
Val Glu Leu His Gly His Leu Tyr Cys Asp Asp Gln Tyr Trp Thr Gly
260 265 270

gca aag gtt tca aca gtt tgc cgc tcg agt tgg cgc act gga ggg tcg 864
Ala Lys Val Ser Thr Val Cys Arg Ser Ser Trp Arg Thr Gly Gly Ser
275 280 285

tgt aaa ata atc caa act att tca atg tat ttc tct ttt aaa tcc cca 912
Cys Lys Ile Ile Gln Thr Ile Ser Met Tyr Phe Ser Phe Lys Ser Pro
290 295 300

tcc cat tta gca tgt tcc cga agc ctt gtg gta cga cat tcg gag gga 960
Ser His Leu Ala Cys Ser Arg Ser Leu Val Val Arg His Ser Glu Gly
305 310 315 320

tca cat tcg cct ttg cga gct ggc cac aaa ctg cgg cat ccc ata aac 1008
Ser His Ser Pro Leu Arg Ala Gly His Lys Leu Arg His Pro Ile Asn
325 330 335

tac gtt tac ttc tgg tac tcg ctg atc ttc ctg ctg gcc agg acg agt 1056
Tyr Val Tyr Phe Trp Tyr Ser Leu Ile Phe Leu Leu Ala Arg Thr Ser
340 345 350

```

```

ctg gtt ttc atg act gcc tcc aag atc cac gat gcc tcg ctt ctg ccc 1104
Leu Val Phe Met Thr Ala Ser Lys Ile His Asp Ala Ser Leu Leu Pro
    355                360                365

ctg agg tcc ttg tac ttg gtg ccc agc gat ggc tgg acg cag gag gtg 1152
Leu Arg Ser Leu Tyr Leu Val Pro Ser Asp Gly Trp Thr Gln Glu Val
    370                375                380

cag aga ttt gcc gac cag ctg acc agc gag ttt gtc gga ttg tct gga 1200
Gln Arg Phe Ala Asp Gln Leu Thr Ser Glu Phe Val Gly Leu Ser Gly
    385                390                395                400

tat cgt ctc ttc tgc ttg aca aga aag agt ctc ttc gga atg cta gcc 1248
Tyr Arg Leu Phe Cys Leu Thr Arg Lys Ser Leu Phe Gly Met Leu Ala
    405                410                415

acc ctg gtg acc tac gaa ctt atg ctg ctg caa ata gat gct aag agc 1296
Thr Leu Val Thr Tyr Glu Leu Met Leu Leu Gln Ile Asp Ala Lys Ser
    420                425                430

cac aag gga ctg cgg tgc gcg 1317
His Lys Gly Leu Arg Cys Ala
    435

```

```

<210> 66
<211> 439
<212> PRT
<213> Drosophila melanogaster

```

```

<400> 66
Met Ser Arg Thr Ser Asp Asp Ile Arg Lys His Leu Lys Val Arg Arg
  1                5                10                15

Gln Lys Gln Arg Ala Ile Leu Ala Met Arg Trp Arg Cys Ala Gln Gly
    20                25                30

Gly Leu Glu Phe Glu Gln Leu Asp Thr Phe Tyr Gly Ala Ile Arg Pro
  35                40                45

Cys Lys Thr Thr Phe Pro Asn His Phe His Met Cys Leu Asn Ile Leu
  50                55                60

Gln Trp Thr Ile Asp Leu Cys Val Ala Gln Phe Phe Gly Ile Met Pro
  65                70                75                80

Leu Ser Asn Ile Arg Ser Arg Asp Pro Gln Asp Val Lys Phe Lys Val
  85                90                95

Arg Ser Ile Gly Leu Ala Val Thr Gly Leu Phe Leu Leu Leu Gly Gly
 100                105                110

Met Lys Thr Leu Val Gly Ala Asn Ile Leu Phe Thr Glu Gly Leu Asn
 115                120                125

Ala Lys Asn Ile Val Gly Leu Val Phe Leu Ile Val Gly Met Val Asn

```

130					135					140					
Trp	Leu	Asn	Phe	Val	Gly	Phe	Ala	Arg	Ser	Trp	Ser	His	Ile	Met	Leu
145					150					155					160
Pro	Trp	Ser	Ser	Val	Asp	Ile	Leu	Met	Leu	Phe	Pro	Pro	Tyr	Lys	Arg
				165					170					175	
Gly	Lys	Arg	Ser	Leu	Arg	Ser	Lys	Val	Asn	Val	Leu	Ala	Leu	Ser	Val
			180					185					190		
Val	Val	Leu	Ala	Ala	Thr	Thr	Cys	Cys	Thr	Thr	Pro	Leu	Ala	Thr	Ala
		195					200					205			
Ala	Ile	Ala	Cys	Thr	Ser	Cys	Ser	Ala	Thr	Gln	Thr	Thr	Arg	Val	Leu
	210					215					220				
Pro	Leu	Asp	Phe	Ile	Trp	Arg	Arg	Ser	Ser	Pro	Thr	Ser	Cys	Ser	Ser
225				230						235					240
Cys	His	Ser	Ile	Tyr	Phe	Pro	Leu	Ala	Lys	Trp	Ser	Leu	His	Leu	Ser
			245						250				255		
Val	Glu	Leu	His	Gly	His	Leu	Tyr	Cys	Asp	Asp	Gln	Tyr	Trp	Thr	Gly
		260						265					270		
Ala	Lys	Val	Ser	Thr	Val	Cys	Arg	Ser	Ser	Trp	Arg	Thr	Gly	Gly	Ser
		275					280					285			
Cys	Lys	Ile	Ile	Gln	Thr	Ile	Ser	Met	Tyr	Phe	Ser	Phe	Lys	Ser	Pro
	290					295					300				
Ser	His	Leu	Ala	Cys	Ser	Arg	Ser	Leu	Val	Val	Arg	His	Ser	Glu	Gly
305				310						315					320
Ser	His	Ser	Pro	Leu	Arg	Ala	Gly	His	Lys	Leu	Arg	His	Pro	Ile	Asn
			325						330					335	
Tyr	Val	Tyr	Phe	Trp	Tyr	Ser	Leu	Ile	Phe	Leu	Leu	Ala	Arg	Thr	Ser
			340					345					350		
Leu	Val	Phe	Met	Thr	Ala	Ser	Lys	Ile	His	Asp	Ala	Ser	Leu	Leu	Pro
		355					360					365			
Leu	Arg	Ser	Leu	Tyr	Leu	Val	Pro	Ser	Asp	Gly	Trp	Thr	Gln	Glu	Val
					375						380				
Gln	Arg	Phe	Ala	Asp	Gln	Leu	Thr	Ser	Glu	Phe	Val	Gly	Leu	Ser	Gly
385					390					395					400
Tyr	Arg	Leu	Phe	Cys	Leu	Thr	Arg	Lys	Ser	Leu	Phe	Gly	Met	Leu	Ala
			405						410					415	
Thr	Leu	Val	Thr	Tyr	Glu	Leu	Met	Leu	Leu	Gln	Ile	Asp	Ala	Lys	Ser
			420					425					430		

His Lys Gly Leu Arg Cys Ala
435

<210> 67
<211> 1080
<212> DNA
<213> *Drosophila melanogaster*

<220>
<221> CDS
<222> (1)..(1080)
<223> Coding region GR65C.1

```

<400> 67
atg cgc gta cat cag cgg caa agt gcg gtc ata att caa atg ggg cac 48
Met Arg Val His Gln Arg Gln Ser Ala Val Ile Ile Gln Met Gly His
  1             5             10             15

cct cca ttc atg tcc ttg aag ggc ggc aaa tcg ggt ttc gga tca att 96
Pro Pro Phe Met Ser Leu Lys Gly Gly Lys Ser Gly Phe Gly Ser Ile
      20             25             30

gtt tgg cca tcc gcg atg agg gaa gtg aat ctg ctc aac cgc ttt aca 144
Val Trp Pro Ser Ala Met Arg Glu Val Asn Leu Leu Asn Arg Phe Thr
      35             40             45

cgc cag ttc ctg ttt ctc atc gtg ttg gtg acc cag atc tgc ggg gtc 192
Arg Gln Phe Leu Phe Leu Ile Val Leu Val Thr Gln Ile Cys Gly Val
      50             55             60

gcc acc ttt gtg tac aac tcg aag gcg caa tgc ttt cgt cag tcc gga 240
Ala Thr Phe Val Tyr Asn Ser Lys Ala Gln Cys Phe Arg Gln Ser Gly
      65             70             75             80

ttt ctg cgg ttc tac tcc agc tta gtt ctc att ttt ctg gca ctt ttc 288
Phe Leu Arg Phe Tyr Ser Ser Leu Val Leu Ile Phe Leu Ala Leu Phe
      85             90             95

ctg att gtt acc acg agt aaa atg ttt cat aat ctg caa gct gtg tgg 336
Leu Ile Val Thr Thr Ser Lys Met Phe His Asn Leu Gln Ala Val Trp
      100            105            110

cca tat gtg gta gga agt gtt atc ata ttg gtg gta aga ata cac gga 384
Pro Tyr Val Val Gly Ser Val Ile Ile Leu Val Val Arg Ile His Gly
      115            120            125

ctt ttg gaa agt gca gag atc gtg gag ttg cta aac caa atg ctg aga 432
Leu Leu Glu Ser Ala Glu Ile Val Glu Leu Leu Asn Gln Met Leu Arg
      130            135            140

atc atg agg cag gtg aat cta atg gcc agg cac ccg aat ctg ttt cgc 480
Ile Met Arg Gln Val Asn Leu Met Ala Arg His Pro Asn Leu Phe Arg
      145            150            155            160

```

```

ctg aaa cat ttg ctg ctc ctt ctt ttg gcc ctg caa aat ctt tta aga 528
Leu Lys His Leu Leu Leu Leu Leu Ala Leu Gln Asn Leu Leu Arg
165 170 175

tca ctg aat acg ata gtg gga ata agt aac cac tct gct gaa gct tat 576
Ser Leu Asn Thr Ile Val Gly Ile Ser Asn His Ser Ala Glu Ala Tyr
180 185 190

gac tct ttt ctt aat agc gtt atc cta tta att ata ctg gcc gtc ctg 624
Asp Ser Phe Leu Asn Ser Val Ile Leu Leu Ile Ile Leu Ala Val Leu
195 200 205

ctg agc ttt ctt ctt cag atc acc atc aat att tgc ctc ttt gta gtg 672
Leu Ser Phe Leu Leu Gln Ile Thr Ile Asn Ile Cys Leu Phe Val Val
210 215 220

ctc att gcc acg tat agc gaa cta cac cat tgc act cga cga atc tca 720
Leu Ile Ala Thr Tyr Ser Glu Leu His His Cys Thr Arg Arg Ile Ser
225 230 235 240

aat gat atg gat aag ctc aga ctt cat tct gtc cat gaa agt ggt caa 768
Asn Asp Met Asp Lys Leu Arg Leu His Ser Val His Glu Ser Gly Gln
245 250 255

ttt atg gtg ttg gta aaa caa ctt caa gga atc act gaa aaa tta att 816
Phe Met Val Leu Val Lys Gln Leu Gln Gly Ile Thr Glu Lys Leu Ile
260 265 270

cga ctg cgt caa aat gtc ttt cat att acc gtc aga atc ata cgg cat 864
Arg Leu Arg Gln Asn Val Phe His Ile Thr Val Arg Ile Ile Arg His
275 280 285

ttc cga ttt cat tgg ctg tgt gct att atc tac gga tta tta cca ttc 912
Phe Arg Phe His Trp Leu Cys Ala Ile Ile Tyr Gly Leu Leu Pro Phe
290 295 300

ttt agt tta aca gct aaa gat caa aat ggt ttt aac ttc ctc atc att 960
Phe Ser Leu Thr Ala Lys Asp Gln Asn Gly Phe Asn Phe Leu Ile Ile
305 310 315 320

tcc gca ttg aac ata ata ttc cag tgg act att ttt gcg att ctt tct 1008
Ser Ala Leu Asn Ile Ile Phe Gln Trp Thr Ile Phe Ala Ile Leu Ser
325 330 335

cgt gaa tca aga atc acc cgg agt tta tgc act ttt cac ttg acc aat 1056
Arg Glu Ser Arg Ile Thr Arg Ser Leu Cys Thr Phe His Leu Thr Asn
340 345 350

tac cat aag gaa act gct aga acg 1080
Tyr His Lys Glu Thr Ala Arg Thr
355 360

```

<210> 68

<211> 360

<212> PRT

<213> *Drosophila melanogaster*

<400> 68

Met Arg Val His Gln Arg Gln Ser Ala Val Ile Ile Gln Met Gly His
 1 5 10 15

Pro Pro Phe Met Ser Leu Lys Gly Gly Lys Ser Gly Phe Gly Ser Ile
 20 25 30

Val Trp Pro Ser Ala Met Arg Glu Val Asn Leu Leu Asn Arg Phe Thr
 35 40 45

Arg Gln Phe Leu Phe Leu Ile Val Leu Val Thr Gln Ile Cys Gly Val
 50 55 60

Ala Thr Phe Val Tyr Asn Ser Lys Ala Gln Cys Phe Arg Gln Ser Gly
 65 70 75 80

Phe Leu Arg Phe Tyr Ser Ser Leu Val Leu Ile Phe Leu Ala Leu Phe
 85 90 95

Leu Ile Val Thr Thr Ser Lys Met Phe His Asn Leu Gln Ala Val Trp
 100 105 110

Pro Tyr Val Val Gly Ser Val Ile Ile Leu Val Val Arg Ile His Gly
 115 120 125

Leu Leu Glu Ser Ala Glu Ile Val Glu Leu Leu Asn Gln Met Leu Arg
 130 135 140

Ile Met Arg Gln Val Asn Leu Met Ala Arg His Pro Asn Leu Phe Arg
 145 150 155 160

Leu Lys His Leu Leu Leu Leu Leu Ala Leu Gln Asn Leu Leu Arg
 165 170 175

Ser Leu Asn Thr Ile Val Gly Ile Ser Asn His Ser Ala Glu Ala Tyr
 180 185 190

Asp Ser Phe Leu Asn Ser Val Ile Leu Leu Ile Ile Leu Ala Val Leu
 195 200 205

Leu Ser Phe Leu Leu Gln Ile Thr Ile Asn Ile Cys Leu Phe Val Val
 210 215 220

Leu Ile Ala Thr Tyr Ser Glu Leu His His Cys Thr Arg Arg Ile Ser
 225 230 235 240

Asn Asp Met Asp Lys Leu Arg Leu His Ser Val His Glu Ser Gly Gln
 245 250 255

Phe Met Val Leu Val Lys Gln Leu Gln Gly Ile Thr Glu Lys Leu Ile
 260 265 270

Arg Leu Arg Gln Asn Val Phe His Ile Thr Val Arg Ile Ile Arg His
 275 280 285

119

Phe Arg Phe His Trp Leu Cys Ala Ile Ile Tyr Gly Leu Leu Pro Phe
 290 295 300

Phe Ser Leu Thr Ala Lys Asp Gln Asn Gly Phe Asn Phe Leu Ile Ile
 305 310 315 320

Ser Ala Leu Asn Ile Ile Phe Gln Trp Thr Ile Phe Ala Ile Leu Ser
 325 330 335

Arg Glu Ser Arg Ile Thr Arg Ser Leu Cys Thr Phe His Leu Thr Asn
 340 345 350

Tyr His Lys Glu Thr Ala Arg Thr
 355 360

<210> 69
 <211> 1431
 <212> DNA
 <213> Drosophila melanogaster

<220>
 <221> CDS
 <222> (1)..(1431)
 <223> Coding region GR66C.1

<400> 69
 atg gcg cag gcg gag gac gca gtg caa cca cta ttg cag cag ttc cag 48
 Met Ala Gln Ala Glu Asp Ala Val Gln Pro Leu Leu Gln Gln Phe Gln
 1 5 10 15

caa ctg ttc ttc ata tcc aag ata gct gga att ctg cca cag gat ctc 96
 Gln Leu Phe Phe Ile Ser Lys Ile Ala Gly Ile Leu Pro Gln Asp Leu
 20 25 30

gag aag ttt cga tct agg aat ctg ctg gag aaa tcc cgt aat ggc atg 144
 Glu Lys Phe Arg Ser Arg Asn Leu Leu Glu Lys Ser Arg Asn Gly Met
 35 40 45

att tac atg ctg agt act tta ata ctc tac gtt gtg ctc tat aat att 192
 Ile Tyr Met Leu Ser Thr Leu Ile Leu Tyr Val Val Leu Tyr Asn Ile
 50 55 60

ttg ata tat tcc ttt gga gag gag gac cgc agc cta aag gcc tcg cag 240
 Leu Ile Tyr Ser Phe Gly Glu Glu Asp Arg Ser Leu Lys Ala Ser Gln
 65 70 75 80

agc acc ttg act ttc gtg att ggc ttg ttc ctg acc tat atc ggt ctg 288
 Ser Thr Leu Thr Phe Val Ile Gly Leu Phe Leu Thr Tyr Ile Gly Leu
 85 90 95

att atg atg gtc tcc gac cag ttg acc gcg tta cga aac cag ggt aga 336
 Ile Met Met Val Ser Asp Gln Leu Thr Ala Leu Arg Asn Gln Gly Arg
 100 105 110

```

att gga gag ctt tac gag cgc atc cgt ctg gtg gat gag cgc ctt tac 384
Ile Gly Glu Leu Tyr Glu Arg Ile Arg Leu Val Asp Glu Arg Leu Tyr
    115                120                125

aaa gag ggg tgt gtt atg gac aac agt aca att gga cgg cgc ata cga 432
Lys Glu Gly Cys Val Met Asp Asn Ser Thr Ile Gly Arg Arg Ile Arg
    130                135                140

att atg ctg atc atg acg gtc atc ttt gag ttg tcc att ttg gtg agc 480
Ile Met Leu Ile Met Thr Val Ile Phe Glu Leu Ser Ile Leu Val Ser
    145                150                155                160

acc tat gtc aag ctg gtg gac tat agt caa tgg atg tcc ttg tta tgg 528
Thr Tyr Val Lys Leu Val Asp Tyr Ser Gln Trp Met Ser Leu Leu Trp
    165                170                175

ata gtg tcc gcc att ccc acg ttc atc aac acg cta gac aag atc tgg 576
Ile Val Ser Ala Ile Pro Thr Phe Ile Asn Thr Leu Asp Lys Ile Trp
    180                185                190

ttc gct gtt tcg tta tat gcg ttg aaa gaa cgc ttc gag gcc ata aac 624
Phe Ala Val Ser Leu Tyr Ala Leu Lys Glu Arg Phe Glu Ala Ile Asn
    195                200                205

gcc acc cta gag gaa ctg gtg gac acg cac gag aag cat aag ctg tgg 672
Ala Thr Leu Glu Glu Leu Val Asp Thr His Glu Lys His Lys Leu Trp
    210                215                220

ctg cga ggc aat caa gag gtt ccg cct cct ctg gac agc tcc cag ccg 720
Leu Arg Gly Asn Gln Val Pro Pro Pro Leu Asp Ser Ser Gln Pro
    225                230                235                240

cct cag tat gac agc aac ctg gag tat ctg tac aag gaa cta gga ggt 768
Pro Gln Tyr Asp Ser Asn Leu Glu Tyr Leu Tyr Lys Glu Leu Gly Gly
    245                250                255

atg gac ata ggt tcc att ggc aag agt tca gta gag gag aaa cta aac 816
Met Asp Ile Gly Ser Ile Gly Lys Ser Ser Val Glu Glu Lys Leu Asn
    260                265                270

aac ctg tgc cag gtg cac gac gag atc tgt gaa atc gga aaa gct ttg 864
Asn Leu Cys Gln Val His Asp Glu Ile Cys Glu Ile Gly Lys Ala Leu
    275                280                285

aac gag ctg tgg agc tat ccc att cta tct cta atg gcc tat ggt ttt 912
Asn Glu Leu Trp Ser Tyr Pro Ile Leu Ser Leu Met Ala Tyr Gly Phe
    290                295                300

ctg att ttc act gct caa ctt tat ttc ctc tac tgc gct aca cag tac 960
Leu Ile Phe Thr Ala Gln Leu Tyr Phe Leu Tyr Cys Ala Thr Gln Tyr
    305                310                315                320

caa tcg ata cca tcg ctt ttc cgt tcc gcc aag aat ccc ttc atc act 1008
Gln Ser Ile Pro Ser Leu Phe Arg Ser Ala Lys Asn Pro Phe Ile Thr
    325                330                335

```

```

gtt ata gtt cta agt tat acg tct gga aaa tgc gtg tac ctc atc tac 1056
Val Ile Val Leu Ser Tyr Thr Ser Gly Lys Cys Val Tyr Leu Ile Tyr
      340      345      350

ctg agt tgg aaa acg tcg cag gcc tcc aag cgc acg gga atc agt ctg 1104
Leu Ser Trp Lys Thr Ser Gln Ala Ser Lys Arg Thr Gly Ile Ser Leu
      355      360      365

cac aaa tgt ggc gtg gtg gcc gat gat aat ctt ctc tac gaa att gtt 1152
His Lys Cys Gly Val Val Ala Asp Asp Asn Leu Leu Tyr Glu Ile Val
      370      375      380

aac cac cta tcg cta aaa ttg ctc aac cac tcg gtg gac ttt tcg gct 1200
Asn His Leu Ser Leu Lys Leu Leu Asn His Ser Val Asp Phe Ser Ala
      385      390      395      400

tgc ggc ttc ttt acc ctg gac atg gaa aca ttg tat ggt gtg agt ggc 1248
Cys Gly Phe Phe Thr Leu Asp Met Glu Thr Leu Tyr Gly Val Ser Gly
      405      410      415

ggg atc act agc tac ctg atc atc ctg att cag ttc aat ttg gcc gcc 1296
Gly Ile Thr Ser Tyr Leu Ile Ile Leu Ile Gln Phe Asn Leu Ala Ala
      420      425      430

cag cag gcc aaa gag gct ata cag acg ttc aac tcg ctt aat gac acc 1344
Gln Gln Ala Lys Glu Ala Ile Gln Thr Phe Asn Ser Leu Asn Asp Thr
      435      440      445

gcc ggc ttg gtt ggt gcc gcc acc gat atg gat aat att agc tcc acg 1392
Ala Gly Leu Val Gly Ala Ala Thr Asp Met Asp Asn Ile Ser Ser Thr
      450      455      460

ctg cgt gat ttc gtc acc acg acc atg aca ccg gcg gtc 1431
Leu Arg Asp Phe Val Thr Thr Thr Met Thr Pro Ala Val
      465      470      475

```

<210> 70

<211> 477

<212> PRT

<213> *Drosophila melanogaster*

<400> 70

```

Met Ala Gln Ala Glu Asp Ala Val Gln Pro Leu Leu Gln Gln Phe Gln
 1          5          10          15

Gln Leu Phe Phe Ile Ser Lys Ile Ala Gly Ile Leu Pro Gln Asp Leu
 20          25          30

Glu Lys Phe Arg Ser Arg Asn Leu Leu Glu Lys Ser Arg Asn Gly Met
 35          40          45

Ile Tyr Met Leu Ser Thr Leu Ile Leu Tyr Val Val Leu Tyr Asn Ile
 50          55          60

Leu Ile Tyr Ser Phe Gly Glu Glu Asp Arg Ser Leu Lys Ala Ser Gln

```

65	70	75	80
Ser Thr Leu Thr Phe Val Ile Gly Leu Phe Leu Thr Tyr Ile Gly Leu	85	90	95
Ile Met Met Val Ser Asp Gln Leu Thr Ala Leu Arg Asn Gln Gly Arg	100	105	110
Ile Gly Glu Leu Tyr Glu Arg Ile Arg Leu Val Asp Glu Arg Leu Tyr	115	120	125
Lys Glu Gly Cys Val Met Asp Asn Ser Thr Ile Gly Arg Arg Ile Arg	130	135	140
Ile Met Leu Ile Met Thr Val Ile Phe Glu Leu Ser Ile Leu Val Ser	145	150	155
Thr Tyr Val Lys Leu Val Asp Tyr Ser Gln Trp Met Ser Leu Leu Trp	165	170	175
Ile Val Ser Ala Ile Pro Thr Phe Ile Asn Thr Leu Asp Lys Ile Trp	180	185	190
Phe Ala Val Ser Leu Tyr Ala Leu Lys Glu Arg Phe Glu Ala Ile Asn	195	200	205
Ala Thr Leu Glu Glu Leu Val Asp Thr His Glu Lys His Lys Leu Trp	210	215	220
Leu Arg Gly Asn Gln Glu Val Pro Pro Pro Leu Asp Ser Ser Gln Pro	225	230	235
Pro Gln Tyr Asp Ser Asn Leu Glu Tyr Leu Tyr Lys Glu Leu Gly Gly	245	250	255
Met Asp Ile Gly Ser Ile Gly Lys Ser Ser Val Glu Glu Lys Leu Asn	260	265	270
Asn Leu Cys Gln Val His Asp Glu Ile Cys Glu Ile Gly Lys Ala Leu	275	280	285
Asn Glu Leu Trp Ser Tyr Pro Ile Leu Ser Leu Met Ala Tyr Gly Phe	290	295	300
Leu Ile Phe Thr Ala Gln Leu Tyr Phe Leu Tyr Cys Ala Thr Gln Tyr	305	310	315
Gln Ser Ile Pro Ser Leu Phe Arg Ser Ala Lys Asn Pro Phe Ile Thr	325	330	335
Val Ile Val Leu Ser Tyr Thr Ser Gly Lys Cys Val Tyr Leu Ile Tyr	340	345	350
Leu Ser Trp Lys Thr Ser Gln Ala Ser Lys Arg Thr Gly Ile Ser Leu	355	360	365

His Lys Cys Gly Val Val Ala Asp Asp Asn Leu Leu Tyr Glu Ile Val
 370 375 380

Asn His Leu Ser Leu Lys Leu Leu Asn His Ser Val Asp Phe Ser Ala
 385 390 395 400

Cys Gly Phe Phe Thr Leu Asp Met Glu Thr Leu Tyr Gly Val Ser Gly
 405 410 415

Gly Ile Thr Ser Tyr Leu Ile Ile Leu Ile Gln Phe Asn Leu Ala Ala
 420 425 430

Gln Gln Ala Lys Glu Ala Ile Gln Thr Phe Asn Ser Leu Asn Asp Thr
 435 440 445

Ala Gly Leu Val Gly Ala Ala Thr Asp Met Asp Asn Ile Ser Ser Thr
 450 455 460

Leu Arg Asp Phe Val Thr Thr Thr Met Thr Pro Ala Val
 465 470 475

<210> 71
 <211> 1126
 <212> DNA
 <213> Drosophila melanogaster

<220>
 <221> CDS
 <222> (1)..(1125)
 <223> Coding region GR68D.1

<400> 71
 atg aag atc tac cag gat ata tat cct ata tcc aag ccc tcg cag atc 48
 Met Lys Ile Tyr Gln Asp Ile Tyr Pro Ile Ser Lys Pro Ser Gln Ile
 1 5 10 15

ttt gct ata tta cca ttc tat agt gga gac gtc gat gat ggc ttt cgc 96
 Phe Ala Ile Leu Pro Phe Tyr Ser Gly Asp Val Asp Asp Gly Phe Arg
 20 25 30

ttt gga gga ttg ggt cgc tgg tac gga cga ctg gtg gcc ctg att ata 144
 Phe Gly Gly Leu Gly Arg Trp Tyr Gly Arg Leu Val Ala Leu Ile Ile
 35 40 45

cta atc gga tca ctg aca ttg ggg gaa gat gtg ctc ttc gcc tcc aag 192
 Leu Ile Gly Ser Leu Thr Leu Gly Glu Asp Val Leu Phe Ala Ser Lys
 50 55 60

gag tac aga ttg gtg gcc agt gcc caa ggt gat acc gag gag atc aac 240
 Glu Tyr Arg Leu Val Ala Ser Ala Gln Gly Asp Thr Glu Glu Ile Asn
 65 70 75 80

agg acc atc gag acg ctt ctg tgt atc atc agc tat acc atg gtg gta 288
 Arg Thr Ile Glu Thr Leu Leu Cys Ile Ile Ser Tyr Thr Met Val Val

	85	90	95	
tta tgc agt gtg cag aat gca tcc aga cac ttt cgc act ctt cac gat				336
Leu Ser Ser Val Gln Asn Ala Ser Arg His Phe Arg Thr Leu His Asp				
	100	105	110	
att gcc aaa atc gat gag tat ctg ctg gcc aat gga ttc cgg gaa acc				384
Ile Ala Lys Ile Asp Glu Tyr Leu Leu Ala Asn Gly Phe Arg Glu Thr				
	115	120	125	
tac agt tgt cgc aat ctg acc att ctg gtc acc tct gca gca ggc ggt				432
Tyr Ser Cys Arg Asn Leu Thr Ile Leu Val Thr Ser Ala Ala Gly Gly				
	130	135	140	
gtc ctc gct gtg gcc ttc tac tat att cac tat cga agt ggc att ggt				480
Val Leu Ala Val Ala Phe Tyr Tyr Ile His Tyr Arg Ser Gly Ile Gly				
	145	150	155	160
gcc aaa agg caa atc atc ctg ctt ctg atc tac ttc ctc caa ctg ctg				528
Ala Lys Arg Gln Ile Ile Leu Leu Leu Ile Tyr Phe Leu Gln Leu Leu				
	165	170	175	
tac tcc act ctc ctg gcc tta tat cta agg aca ctc atg atg aat ttg				576
Tyr Ser Thr Leu Leu Ala Leu Tyr Leu Arg Thr Leu Met Met Asn Leu				
	180	185	190	
gca cag aga att gga ttt ctt aac cag aaa ctg gat acg ttt aat ctt				624
Ala Gln Arg Ile Gly Phe Leu Asn Gln Lys Leu Asp Thr Phe Asn Leu				
	195	200	205	
cag gat tgt ggc cac atg gaa aac tgg cgt gag ttg agc aac ctc ata				672
Gln Asp Cys Gly His Met Glu Asn Trp Arg Glu Leu Ser Asn Leu Ile				
	210	215	220	
gag gtc ctt tgc aaa ttc cgc tac atc acc gag aat att aac tgc gtg				720
Glu Val Leu Cys Lys Phe Arg Tyr Ile Thr Glu Asn Ile Asn Cys Val				
	225	230	235	240
gcc gga gtt tca ttg cta ttc tac ttt ggc ttc tcc ttc tac acg gtc				768
Ala Gly Val Ser Leu Leu Phe Tyr Phe Gly Phe Ser Phe Tyr Thr Val				
	245	250	255	
aca aat cag agt tac ttg gcc ttt gcc aca ttg aca gcc ggc tgc ttg				816
Thr Asn Gln Ser Tyr Leu Ala Phe Ala Thr Leu Thr Ala Gly Ser Leu				
	260	265	270	
agt tcc aaa aca gaa gtg gct gat acc ata gga cta tca tgt atc tgg				864
Ser Ser Lys Thr Glu Val Ala Asp Thr Ile Gly Leu Ser Cys Ile Trp				
	275	280	285	
gtt ttg gct gaa act atc aca atg att gtg att tgc agc gca tgt gac				912
Val Leu Ala Glu Thr Ile Thr Met Ile Val Ile Cys Ser Ala Cys Asp				
	290	295	300	
ggc ctg gca tcc gag gtg aat ggc acg gca cag atc ctg gcg aga att				960
Gly Leu Ala Ser Glu Val Asn Gly Thr Ala Gln Ile Leu Ala Arg Ile				

125

```
<210> 72
<211> 375
<212> PRT
<213> Drosophila melanogaster
```

```

<400> 72
Met Lys Ile Tyr Gln Asp Ile Tyr Pro Ile Ser Lys Pro Ser Gln Ile
  1             5             10             15
Phe Ala Ile Leu Pro Phe Tyr Ser Gly Asp Val Asp Asp Gly Phe Arg
      20             25             30
Phe Gly Gly Leu Gly Arg Trp Tyr Gly Arg Leu Val Ala Leu Ile Ile
      35             40             45
Leu Ile Gly Ser Leu Thr Leu Gly Glu Asp Val Leu Phe Ala Ser Lys
      50             55             60
Glu Tyr Arg Leu Val Ala Ser Ala Gln Gly Asp Thr Glu Glu Ile Asn
  65             70             75             80
Arg Thr Ile Glu Thr Leu Leu Cys Ile Ile Ser Tyr Thr Met Val Val
      85             90             95
Leu Ser Ser Val Gln Asn Ala Ser Arg His Phe Arg Thr Leu His Asp
      100            105            110
Ile Ala Lys Ile Asp Glu Tyr Leu Leu Ala Asn Gly Phe Arg Glu Thr
      115            120            125
Tyr Ser Cys Arg Asn Leu Thr Ile Leu Val Thr Ser Ala Ala Gly Gly
      130            135            140
Val Leu Ala Val Ala Phe Tyr Tyr Ile His Tyr Arg Ser Gly Ile Gly
  145            150            155            160
Ala Lys Arg Gln Ile Ile Leu Leu Leu Ile Tyr Phe Leu Gln Leu Leu

```

	165	170	175
Tyr Ser Thr	Leu Leu Ala Leu Tyr	Leu Arg Thr Leu Met Met Asn Leu	
	180	185	190
Ala Gln Arg	Ile Gly Phe Leu Asn Gln Lys Leu Asp Thr Phe Asn Leu		
	195	200	205
Gln Asp Cys	Gly His Met Glu Asn Trp Arg Glu Leu Ser Asn Leu Ile		
	210	215	220
Glu Val Leu	Cys Lys Phe Arg Tyr Ile Thr Glu Asn Ile Asn Cys Val		
	225	230	235
Ala Gly Val	Ser Leu Leu Phe Tyr Phe Gly Phe Ser Phe Tyr Thr Val		
	245	250	255
Thr Asn Gln	Ser Tyr Leu Ala Phe Ala Thr Leu Thr Ala Gly Ser Leu		
	260	265	270
Ser Ser Lys	Thr Glu Val Ala Asp Thr Ile Gly Leu Ser Cys Ile Trp		
	275	280	285
Val Leu Ala	Glu Thr Ile Thr Met Ile Val Ile Cys Ser Ala Cys Asp		
	290	295	300
Gly Leu Ala	Ser Glu Val Asn Gly Thr Ala Gln Ile Leu Ala Arg Ile		
	305	310	315
Tyr Gly Lys	Ser Lys Gln Phe Gln Asn Leu Ile Asp Lys Phe Leu Thr		
	325	330	335
Lys Ser Ile	Lys Gln Asp Leu Gln Phe Thr Ala Tyr Gly Phe Phe Ser		
	340	345	350
Ile Asp Asn	Ser Thr Leu Phe Lys Ile Phe Ser Ala Val Thr Thr Tyr		
	355	360	365
Leu Val Ile	Leu Ile Gln Phe		
	370	375	

<210> 73
 <211> 1347
 <212> DNA
 <213> *Drosophila melanogaster*

<220>
 <221> CDS
 <222> (1) .. (1347)
 <223> Coding region GR77E.1

<400> 73
 atg cca ttg cct ttg ggc gac cca cta gca cta gca gtt tcg cct cag 48
 Met Pro Leu Pro Leu Gly Asp Pro Leu Ala Leu Ala Val Ser Pro Gln

1	5	10	15	
ctc gga tac att cga atc aca gcg atg ccg cgt tgg ctt caa ctc ccg				96
Leu Gly Tyr Ile Arg Ile Thr Ala Met Pro Arg Trp Leu Gln Leu Pro	20	25	30	
ggt atg tcg gcc ctg ggt ata ctg tac agt ctg acc agg gtg ttt ggc				144
Gly Met Ser Ala Leu Gly Ile Leu Tyr Ser Leu Thr Arg Val Phe Gly	35	40	45	
cta atg gcc acc gcc aat tgg agt cct cga ggc att aag agg gtc aga				192
Leu Met Ala Thr Ala Asn Trp Ser Pro Arg Gly Ile Lys Arg Val Arg	50	55	60	
cag agc ctc tat ctt cgc atc cac ggc tgc gtg atg ctg atc ttt gtg				240
Gln Ser Leu Tyr Leu Arg Ile His Gly Cys Val Met Leu Ile Phe Val	65	70	75	80
ggc tgt ttc tcg ccc ttc gcc ttc tgg tgc att ttc caa agg atg gcc				288
Gly Cys Phe Ser Pro Phe Ala Phe Trp Cys Ile Phe Gln Arg Met Ala	85	90	95	
ttt ctg cgc caa aac cga att ctg cta atg att ggc ttt aat cga tat				336
Phe Leu Arg Gln Asn Arg Ile Leu Leu Met Ile Gly Phe Asn Arg Tyr	100	105	110	
gtc ctg ctg ctg gtc tgt gct ttt atg act ctg tgg ata cac tgc ttc				384
Val Leu Leu Leu Val Cys Ala Phe Met Thr Leu Trp Ile His Cys Phe	115	120	125	
aag cag gcg gag att att ggg tgt ctc aat cga ttg ctg aag tgt cga				432
Lys Gln Ala Glu Ile Ile Gly Cys Leu Asn Arg Leu Leu Lys Cys Arg	130	135	140	
agg cgg tta aga aga ctc atg cac act cgg aag ctg aag gat tcc atg				480
Arg Arg Leu Arg Arg Leu Met His Thr Arg Lys Leu Lys Asp Ser Met	145	150	155	160
gat tgt ctg gcc act aag gga cac ctg cta gaa gtg gtg gtg cta ctc				528
Asp Cys Leu Ala Thr Lys Gly His Leu Leu Glu Val Val Val Leu Leu	165	170	175	
agt tcc tat ctg ctc tcc atg gct cag ccc att cag ata ctc aag gat				576
Ser Ser Tyr Leu Leu Ser Met Ala Gln Pro Ile Gln Ile Leu Lys Asp	180	185	190	
gat ccg gaa gtg cga agg aac ttc atg tac gcc tgc tct ttg gtt ttc				624
Asp Pro Glu Val Arg Arg Asn Phe Met Tyr Ala Cys Ser Leu Val Phe	195	200	205	
gta tcc gta tgc cag gcc atc ctg cag ctt tcc ctg ggg atg tac aca				672
Val Ser Val Cys Gln Ala Ile Leu Gln Leu Ser Leu Gly Met Tyr Thr	210	215	220	
atg gcc ata cta ttt ctg ggt cac ctc gtt cgt cac tcc aat ctg ctg				720
Met Ala Ile Leu Phe Leu Gly His Leu Val Arg His Ser Asn Leu Leu				

225	230	235	240	
ctg gcc aaa att	ctg gcg gat gcc gag	cat att ttt gaa agt	tcc cag	768
Leu Ala Lys Ile	Leu Ala Asp Ala Glu His	Ile Phe Glu Ser	Ser Gln	
	245	250	255	
aag gcc gga ttt	tgg ccc aac cgc cag gag	ctg tac aaa ggc	caa caa	816
Lys Ala Gly Phe	Trp Pro Asn Arg Gln Glu	Leu Tyr Lys Gly	Gln Gln	
	260	265	270	
aag tgg ttg gcc	ctc gaa ttg tgg aga	tta ctt cat gta	cat cat caa	864
Lys Trp Leu Ala	Leu Glu Leu Trp Arg	Leu Leu His Val	His His Gln	
	275	280	285	
ttg ttg aag ctc	cat cgg tcc ata tgc	tcg ttg tgc gca	gtt cag gcc	912
Leu Leu Lys Leu	His Arg Ser Ile Cys	Ser Leu Cys Ala	Val Gln Ala	
	290	295	300	
gtc tgc ttt ttg	gga ttt gtg ccc ttg	gag tgc acg att	cac ctg ttc	960
Val Cys Phe Leu	Gly Phe Val Pro Leu	Glu Cys Thr Ile	His Leu Phe	
	305	310	315	320
ttc aca tac ttc	atg aag tac agc aag	ttt att ttg cga	aaa tat gga	1008
Phe Thr Tyr Phe	Met Lys Tyr Ser Lys	Phe Ile Leu Arg	Lys Tyr Gly	
	325	330	335	
aga tcg ttt ccg	ttg aac tac ttc gcc	atc gcc ttt ttg	gtg ggg ctc	1056
Arg Ser Phe Pro	Leu Asn Tyr Phe Ala	Ile Ala Phe Leu	Val Gly Leu	
	340	345	350	
ttc acc aac ctt	ttg ctg gtc att tta	ccc acc tac tac	tca gaa aga	1104
Phe Thr Asn Leu	Leu Leu Val Ile Leu	Pro Thr Tyr Tyr	Ser Glu Arg	
	355	360	365	
aga ttc aac tgt	acc agg gaa atc atc	aaa ggc gga gga	ttg gcc ttt	1152
Arg Phe Asn Cys	Thr Arg Glu Ile Ile	Lys Gly Gly Gly	Leu Ala Phe	
	370	375	380	
cca tcc agg atc	act gtt aag caa ctc	agg cac act atg	cac ttc tac	1200
Pro Ser Arg Ile	Thr Val Lys Gln Leu	Arg His Thr Met	His Phe Tyr	
	385	390	395	400
ggc ctg tat ctg	aag aat gtg gag cat	gta ttt gcc gtc	agc gcc tgt	1248
Gly Leu Tyr Leu	Lys Asn Val Glu His	Val Phe Ala Val	Ser Ala Cys	
	405	410	415	
gga ctg ttt aag	ctg aac aat gcg ata	ctc ttt tgc atc	gtg ggt gca	1296
Gly Leu Phe Lys	Leu Asn Asn Ala Ile	Leu Phe Cys Ile	Val Gly Ala	
	420	425	430	
ata cta gaa tac	ctt atg ata ctg ata	cag ttt gat aaa	gtc cta aac	1344
Ile Leu Glu Tyr	Leu Met Ile Leu Ile	Gln Phe Asp Lys	Val Leu Asn	
	435	440	445	
aaa				1347
Lys				

<210> 74

<211> 449

<212> PRT

<213> *Drosophila melanogaster*

<400> 74

```

Met Pro Leu Pro Leu Gly Asp Pro Leu Ala Leu Ala Val Ser Pro Gln
 1           5           10           15

Leu Gly Tyr Ile Arg Ile Thr Ala Met Pro Arg Trp Leu Gln Leu Pro
      20           25           30

Gly Met Ser Ala Leu Gly Ile Leu Tyr Ser Leu Thr Arg Val Phe Gly
      35           40           45

Leu Met Ala Thr Ala Asn Trp Ser Pro Arg Gly Ile Lys Arg Val Arg
      50           55           60

Gln Ser Leu Tyr Leu Arg Ile His Gly Cys Val Met Leu Ile Phe Val
      65           70           75           80

Gly Cys Phe Ser Pro Phe Ala Phe Trp Cys Ile Phe Gln Arg Met Ala
      85           90           95

Phe Leu Arg Gln Asn Arg Ile Leu Leu Met Ile Gly Phe Asn Arg Tyr
      100          105          110

Val Leu Leu Leu Val Cys Ala Phe Met Thr Leu Trp Ile His Cys Phe
      115          120          125

Lys Gln Ala Glu Ile Ile Gly Cys Leu Asn Arg Leu Leu Lys Cys Arg
      130          135          140

Arg Arg Leu Arg Arg Leu Met His Thr Arg Lys Leu Lys Asp Ser Met
      145          150          155          160

Asp Cys Leu Ala Thr Lys Gly His Leu Leu Glu Val Val Val Leu Leu
      165          170          175

Ser Ser Tyr Leu Leu Ser Met Ala Gln Pro Ile Gln Ile Leu Lys Asp
      180          185          190

Asp Pro Glu Val Arg Arg Asn Phe Met Tyr Ala Cys Ser Leu Val Phe
      195          200          205

Val Ser Val Cys Gln Ala Ile Leu Gln Leu Ser Leu Gly Met Tyr Thr
      210          215          220

Met Ala Ile Leu Phe Leu Gly His Leu Val Arg His Ser Asn Leu Leu
      225          230          235          240

Leu Ala Lys Ile Leu Ala Asp Ala Glu His Ile Phe Glu Ser Ser Gln
      245          250          255

Lys Ala Gly Phe Trp Pro Asn Arg Gln Glu Leu Tyr Lys Gly Gln Gln

```

130

260 265 270
 Lys Trp Leu Ala Leu Glu Leu Trp Arg Leu Leu His Val His His Gln
 275 280 285
 Leu Leu Lys Leu His Arg Ser Ile Cys Ser Leu Cys Ala Val Gln Ala
 290 295 300
 Val Cys Phe Leu Gly Phe Val Pro Leu Glu Cys Thr Ile His Leu Phe
 305 310 315 320
 Phe Thr Tyr Phe Met Lys Tyr Ser Lys Phe Ile Leu Arg Lys Tyr Gly
 325 330 335
 Arg Ser Phe Pro Leu Asn Tyr Phe Ala Ile Ala Phe Leu Val Gly Leu
 340 345 350
 Phe Thr Asn Leu Leu Leu Val Ile Leu Pro Thr Tyr Tyr Ser Glu Arg
 355 360 365
 Arg Phe Asn Cys Thr Arg Glu Ile Ile Lys Gly Gly Gly Leu Ala Phe
 370 375 380
 Pro Ser Arg Ile Thr Val Lys Gln Leu Arg His Thr Met His Phe Tyr
 385 390 395 400
 Gly Leu Tyr Leu Lys Asn Val Glu His Val Phe Ala Val Ser Ala Cys
 405 410 415
 Gly Leu Phe Lys Leu Asn Asn Ala Ile Leu Phe Cys Ile Val Gly Ala
 420 425 430
 Ile Leu Glu Tyr Leu Met Ile Leu Ile Gln Phe Asp Lys Val Leu Asn
 435 440 445
 Lys

<210> 75
 <211> 1165
 <212> DNA
 <213> *Drosophila melanogaster*

<220>
 <221> CDS
 <222> (1) .. (1164)
 <223> Coding region GR94E.1

<400> 75
 atg gac ttc acc agc gac tac gcg cat cgg cgt atg gtg aaa ttt ctg 48
 Met Asp Phe Thr Ser Asp Tyr Ala His Arg Arg Met Val Lys Phe Leu
 1 5 10 15
 acg atc ata ctg ata ggc ttt atg acc gtc ttc gga ctc ctg gcc aat 96

Thr Ile Ile Leu Ile Gly Phe Met Thr Val Phe Gly Leu Leu Ala Asn
 20 25 30
 cga tat cgg gcg ggc cgt cgt gaa aga ttc cgc ttc tca aag gca aat 144
 Arg Tyr Arg Ala Gly Arg Arg Glu Arg Phe Arg Phe Ser Lys Ala Asn
 35 40 45
 ctg gcc ttt gct tcg ctg tgg gca att gca ttc agc ttg gtt tac ggg 192
 Leu Ala Phe Ala Ser Leu Trp Ala Ile Ala Phe Ser Leu Val Tyr Gly
 50 55 60
 cgg caa atc tac aag gag tac cag gag ggt cag atc aac ctg aag gac 240
 Arg Gln Ile Tyr Lys Glu Tyr Gln Glu Gly Gln Ile Asn Leu Lys Asp
 65 70 75 80
 gcc acc act ctg tac agc tat atg aac att acg gtg gct gtt att aac 288
 Ala Thr Thr Leu Tyr Ser Tyr Met Asn Ile Thr Val Ala Val Ile Asn
 85 90 95
 tat gtg tcg caa atg ata atc agt gac cat gtg gcc aag gtg ttg agc 336
 Tyr Val Ser Gln Met Ile Ile Ser Asp His Val Ala Lys Val Leu Ser
 100 105 110
 aaa gtg ccc ttc ttt gat acc cta aaa gaa ttc cgt ctg gac agc agg 384
 Lys Val Pro Phe Phe Asp Thr Leu Lys Glu Phe Arg Leu Asp Ser Arg
 115 120 125
 tcg ctg tac ata tcc atc gtt ttg gct ctg gtc aag acc gtg gct ttt 432
 Ser Leu Tyr Ile Ser Ile Val Leu Ala Leu Val Lys Thr Val Ala Phe
 130 135 140
 cct cta aca att gaa gtg gct ttc ata ctg caa cag agg cgg cag cat 480
 Pro Leu Thr Ile Glu Val Ala Phe Ile Leu Gln Gln Arg Arg Gln His
 145 150 155 160
 ccc gag atg agc ttg atc tgg acc ttg tac cgg ctg ttt ccc tta att 528
 Pro Glu Met Ser Leu Ile Trp Thr Leu Tyr Arg Leu Phe Pro Leu Ile
 165 170 175
 att tcg aat ttt ctc aat aac tgc tac ttt ggc gca atg gtg gtg gtg 576
 Ile Ser Asn Phe Leu Asn Asn Cys Tyr Phe Gly Ala Met Val Val Val
 180 185 190
 aag gag att ctg tac gct ctg aac aga cgg ctg gaa gcg cag ctg cag 624
 Lys Glu Ile Leu Tyr Ala Leu Asn Arg Arg Leu Glu Ala Gln Leu Gln
 195 200 205
 gag gtg aat ctg ctg cag agg aag gac cag cta aag ttg tac act aaa 672
 Glu Val Asn Leu Leu Gln Arg Lys Asp Gln Leu Lys Leu Tyr Thr Lys
 210 215 220
 tac tac cgc atg cag cga ttt tgc gcc ttg gcg gat gaa ctc gac cag 720
 Tyr Tyr Arg Met Gln Arg Phe Cys Ala Leu Ala Asp Glu Leu Asp Gln
 225 230 235 240
 ctg gcg tat cgc tat agg ttg ata tat gtg cat tcg gga aag tat ctg 768

Leu Ala Tyr Arg Tyr Arg Leu Ile Tyr Val His Ser Gly Lys Tyr Leu
 245 250 255
 acc cca atg tcc ttg tcc atg att ctg tgc ctc ata tgc cac ctg ctc 816
 Thr Pro Met Ser Leu Ser Met Ile Leu Ser Leu Ile Cys His Leu Leu
 260 265 270
 gga ata acg gtg ggt ttc tac agt ctg tac tat gcc ata gcg gac acc 864
 Gly Ile Thr Val Gly Phe Tyr Ser Leu Tyr Tyr Ala Ile Ala Asp Thr
 275 280 285
 tta atc atg ggc aag ccc gta caa tgg tct tgg atc gct gat caa tct 912
 Leu Ile Met Gly Lys Pro Val Gln Trp Ser Trp Ile Ala Asp Gln Ser
 290 295 300
 ggt ttt cct ctc cat ctc gct ggc gga gat cac att gct cac gca ttt 960
 Gly Phe Pro Leu His Leu Ala Gly Gly Asp His Ile Ala His Ala Phe
 305 310 315 320
 gag atg aat ctc cag cat gcg gac agc cgc tac cgt cag gca gtc cac 1008
 Glu Met Asn Leu Gln His Ala Asp Ser Arg Tyr Arg Gln Ala Val His
 325 330 335
 ggc ttt act ctg ctg gtc acg gtg acc aag tac caa att aaa ccc ttg 1056
 Gly Phe Thr Leu Leu Val Thr Val Thr Lys Tyr Gln Ile Lys Pro Leu
 340 345 350
 ggc ttg tac gag ctg gac atg cga ctg atc agc aat gtc ttc tgc gcg 1104
 Gly Leu Tyr Glu Leu Asp Met Arg Leu Ile Ser Asn Val Phe Ser Ala
 355 360 365
 gtg gcc agc ttc ctg ctg atc ctc gtg cag gcc gat ctg tcc cag cgc 1152
 Val Ala Ser Phe Leu Leu Ile Leu Val Gln Ala Asp Leu Ser Gln Arg
 370 375 380
 ttc aag atg caa t 1165
 Phe Lys Met Gln
 385

<210> 76

<211> 388

<212> PRT

<213> *Drosophila melanogaster*

<400> 76

Met Asp Phe Thr Ser Asp Tyr Ala His Arg Arg Met Val Lys Phe Leu
 1 5 10 15
 Thr Ile Ile Leu Ile Gly Phe Met Thr Val Phe Gly Leu Leu Ala Asn
 20 25 30
 Arg Tyr Arg Ala Gly Arg Arg Glu Arg Phe Arg Phe Ser Lys Ala Asn
 35 40 45
 Leu Ala Phe Ala Ser Leu Trp Ala Ile Ala Phe Ser Leu Val Tyr Gly

50	55	60
Arg Gln Ile Tyr Lys Glu Tyr Gln Glu Gly Gln Ile Asn Leu Lys Asp 65 70 75 80		
Ala Thr Thr Leu Tyr Ser Tyr Met Asn Ile Thr Val Ala Val Ile Asn 85 90 95		
Tyr Val Ser Gln Met Ile Ile Ser Asp His Val Ala Lys Val Leu Ser 100 105 110		
Lys Val Pro Phe Phe Asp Thr Leu Lys Glu Phe Arg Leu Asp Ser Arg 115 120 125		
Ser Leu Tyr Ile Ser Ile Val Leu Ala Leu Val Lys Thr Val Ala Phe 130 135 140		
Pro Leu Thr Ile Glu Val Ala Phe Ile Leu Gln Gln Arg Arg Gln His 145 150 155 160		
Pro Glu Met Ser Leu Ile Trp Thr Leu Tyr Arg Leu Phe Pro Leu Ile 165 170 175		
Ile Ser Asn Phe Leu Asn Asn Cys Tyr Phe Gly Ala Met Val Val Val 180 185 190		
Lys Glu Ile Leu Tyr Ala Leu Asn Arg Arg Leu Glu Ala Gln Leu Gln 195 200 205		
Glu Val Asn Leu Leu Gln Arg Lys Asp Gln Leu Lys Leu Tyr Thr Lys 210 215 220		
Tyr Tyr Arg Met Gln Arg Phe Cys Ala Leu Ala Asp Glu Leu Asp Gln 225 230 235 240		
Leu Ala Tyr Arg Tyr Arg Leu Ile Tyr Val His Ser Gly Lys Tyr Leu 245 250 255		
Thr Pro Met Ser Leu Ser Met Ile Leu Ser Leu Ile Cys His Leu Leu 260 265 270		
Gly Ile Thr Val Gly Phe Tyr Ser Leu Tyr Tyr Ala Ile Ala Asp Thr 275 280 285		
Leu Ile Met Gly Lys Pro Val Gln Trp Ser Trp Ile Ala Asp Gln Ser 290 295 300		
Gly Phe Pro Leu His Leu Ala Gly Gly Asp His Ile Ala His Ala Phe 305 310 315 320		
Glu Met Asn Leu Gln His Ala Asp Ser Arg Tyr Arg Gln Ala Val His 325 330 335		
Gly Phe Thr Leu Leu Val Thr Val Thr Lys Tyr Gln Ile Lys Pro Leu 340 345 350		

Gly Leu Tyr Glu Leu Asp Met Arg Leu Ile Ser Asn Val Phe Ser Ala
 355 360 365

Val Ala Ser Phe Leu Leu Ile Leu Val Gln Ala Asp Leu Ser Gln Arg
 370 375 380

Phe Lys Met Gln
 385

<210> 77

<211> 1269

<212> DNA

<213> Drosophila melanogaster

<220>

<221> CDS

<222> (1)..(1269)

<223> Coding region GR97D.1

<400> 77

atg cga ttc ctg cga agg cag aca cgt cga ctc cgc tcc atc tgg cag 48
 Met Arg Phe Leu Arg Arg Gln Thr Arg Arg Leu Arg Ser Ile Trp Gln
 1 5 10 15

cga agc ctt ccc gtt cgt ttc cgg cga gga aaa ctc cat acc caa ctg 96
 Arg Ser Leu Pro Val Arg Phe Arg Arg Gly Lys Leu His Thr Gln Leu
 20 25 30

gtc acg att tgt ctt tac gcg acc gtt ttc ctt aac ata ctt tac ggc 144
 Val Thr Ile Cys Leu Tyr Ala Thr Val Phe Leu Asn Ile Leu Tyr Gly
 35 40 45

gtt tat ctc ggc cgc ttt tcc ttc agg cgc aag aag ttt gtg ttt tcc 192
 Val Tyr Leu Gly Arg Phe Ser Phe Arg Arg Lys Lys Phe Val Phe Ser
 50 55 60

aaa ggg ctc act atc tat agc tta ttt gtg gcc acg ttc ttt gcg ctg 240
 Lys Gly Leu Thr Ile Tyr Ser Leu Phe Val Ala Thr Phe Phe Ala Leu
 65 70 75 80

ttt tac atc tgg aat att tat aat gaa att tcc act ggt cag atc aat 288
 Phe Tyr Ile Trp Asn Ile Tyr Asn Glu Ile Ser Thr Gly Gln Ile Asn
 85 90 95

ctt cgc gat aca att gga ata tat tgt tat atg aac gtc tgc gtt tgc 336
 Leu Arg Asp Thr Ile Gly Ile Tyr Cys Tyr Met Asn Val Cys Val Cys
 100 105 110

cta ttc aac tat gtg acc caa tgg gaa aaa aca ctg caa ata att cgg 384
 Leu Phe Asn Tyr Val Thr Gln Trp Glu Lys Thr Leu Gln Ile Ile Arg
 115 120 125

ttt cag aat agt gtg cct ctt ttc aag gtc ctc gat tca ctg gac att 432
 Phe Gln Asn Ser Val Pro Leu Phe Lys Val Leu Asp Ser Leu Asp Ile

130	135	140	
tcg gcg atg att gtg tgg cgg gca ttt ata tat ggc ttg ctc aag atc			480
Ser Ala Met Ile Val Trp Arg Ala Phe Ile Tyr Gly Leu Leu Lys Ile			
145	150	155	160
gtg ttc tgt ccc ctt atc acc tat ata acg ctg att cta tat cat aga			528
Val Phe Cys Pro Leu Ile Thr Tyr Ile Thr Leu Ile Leu Tyr His Arg			
	165	170	175
cgc tcc att tcg gaa agt caa tgg acg agt gta acg acc acg aag acc			576
Arg Ser Ile Ser Glu Ser Gln Trp Thr Ser Val Thr Thr Thr Lys Thr			
	180	185	190
atg ttg cca ctg att gtt tcc aat cag ata aat aat tgc ttc ttc ggc			624
Met Leu Pro Leu Ile Val Ser Asn Gln Ile Asn Asn Cys Phe Phe Gly			
	195	200	205
ggc ctg gta ctt gca aat ttg ata ttt gcc gcg gtg aat cgt aag ctg			672
Gly Leu Val Leu Ala Asn Leu Ile Phe Ala Ala Val Asn Arg Lys Leu			
	210	215	220
cac ggc att gtc aag gag gcc aat atg ctg cag tca cct gtc cag atg			720
His Gly Ile Val Lys Glu Ala Asn Met Leu Gln Ser Pro Val Gln Met			
	225	230	235
aat ctc cat aag ccc tac tac cgg atg cgt cgc ttc tgc gag tta gcc			768
Asn Leu His Lys Pro Tyr Tyr Arg Met Arg Arg Phe Cys Glu Leu Ala			
	245	250	255
gat ctt ctg gat gag ttg gct agg aaa tat ggc ttc act gcc agc cgc			816
Asp Leu Leu Asp Glu Leu Ala Arg Lys Tyr Gly Phe Thr Ala Ser Arg			
	260	265	270
tca aag aat tac ctt cgg ttc acg gac tgg tcc atg gtg cta tcg atg			864
Ser Lys Asn Tyr Leu Arg Phe Thr Asp Trp Ser Met Val Leu Ser Met			
	275	280	285
cta atg aac ctt ctc ggt atc acc atg gga tgc tac aat cag tat ctg			912
Leu Met Asn Leu Leu Gly Ile Thr Met Gly Cys Tyr Asn Gln Tyr Leu			
	290	295	300
gcc att gcg gat cac tac atc aac gag gag cct ttc gat ctc ttt ctg			960
Ala Ile Ala Asp His Tyr Ile Asn Glu Glu Pro Phe Asp Leu Phe Leu			
	305	310	315
gcc att gtt ctg gtc gta ttt ctg gct gtt ccc ttc ctg gaa ctc gtc			1008
Ala Ile Val Leu Val Val Phe Leu Ala Val Pro Phe Leu Glu Leu Val			
	325	330	335
atg gtg gct cgg ata agt aac caa acc ctc acc agg aga act gga gaa			1056
Met Val Ala Arg Ile Ser Asn Gln Thr Leu Thr Arg Arg Thr Gly Glu			
	340	345	350
cta ttg cag cga ttc gat ctg cag cac gcc gat gcc cgt ttc aag caa			1104
Leu Leu Gln Arg Phe Asp Leu Gln His Ala Asp Ala Arg Phe Lys Gln			

```

          355          360          365
gtg gtg aat gct ttc tgg ctg cag gtc gtc acc atc aac tac aag ctt 1152
Val Val Asn Ala Phe Trp Leu Gln Val Val Thr Ile Asn Tyr Lys Leu
    370          375          380

atg cca ttg ggc ctc ctg gaa ctg aat acc tcg ctg gtc aat aag gta 1200
Met Pro Leu Gly Leu Leu Glu Leu Asn Thr Ser Leu Val Asn Lys Val
    385          390          395          400

ttc tcg tca gct att gga agc ctg ctg att ctc att caa agt gat ttg 1248
Phe Ser Ser Ala Ile Gly Ser Leu Leu Ile Leu Ile Gln Ser Asp Leu
    405          410          415

aca tta agg ttt tct cta aaa 1269
Thr Leu Arg Phe Ser Leu Lys
    420

```

<210> 78
 <211> 423
 <212> PRT
 <213> *Drosophila melanogaster*

```

<400> 78
Met Arg Phe Leu Arg Arg Gln Thr Arg Arg Leu Arg Ser Ile Trp Gln
  1          5          10          15

Arg Ser Leu Pro Val Arg Phe Arg Arg Gly Lys Leu His Thr Gln Leu
    20          25          30

Val Thr Ile Cys Leu Tyr Ala Thr Val Phe Leu Asn Ile Leu Tyr Gly
    35          40          45

Val Tyr Leu Gly Arg Phe Ser Phe Arg Arg Lys Lys Phe Val Phe Ser
    50          55          60

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

Phe Tyr Ile Trp Asn Ile Tyr Asn Glu Ile Ser Thr Gly Gln Ile Asn
    85          90          95

Leu Arg Asp Thr Ile Gly Ile Tyr Cys Tyr Met Asn Val Cys Val Cys
    100          105          110

Leu Phe Asn Tyr Val Thr Gln Trp Glu Lys Thr Leu Gln Ile Ile Arg
    115          120          125

Phe Gln Asn Ser Val Pro Leu Phe Lys Val Leu Asp Ser Leu Asp Ile
    130          135          140

Ser Ala Met Ile Val Trp Arg Ala Phe Ile Tyr Gly Leu Leu Lys Ile
    145          150          155          160

Val Phe Cys Pro Leu Ile Thr Tyr Ile Thr Leu Ile Leu Tyr His Arg

```

	165	170	175
Arg Ser Ile Ser Glu Ser Gln Trp Thr Ser Val Thr Thr Thr Lys Thr	180	185	190
Met Leu Pro Leu Ile Val Ser Asn Gln Ile Asn Asn Cys Phe Phe Gly	195	200	205
Gly Leu Val Leu Ala Asn Leu Ile Phe Ala Ala Val Asn Arg Lys Leu	210	215	220
His Gly Ile Val Lys Glu Ala Asn Met Leu Gln Ser Pro Val Gln Met	225	230	235
Asn Leu His Lys Pro Tyr Tyr Arg Met Arg Arg Phe Cys Glu Leu Ala	245	250	255
Asp Leu Leu Asp Glu Leu Ala Arg Lys Tyr Gly Phe Thr Ala Ser Arg	260	265	270
Ser Lys Asn Tyr Leu Arg Phe Thr Asp Trp Ser Met Val Leu Ser Met	275	280	285
Leu Met Asn Leu Leu Gly Ile Thr Met Gly Cys Tyr Asn Gln Tyr Leu	290	295	300
Ala Ile Ala Asp His Tyr Ile Asn Glu Glu Pro Phe Asp Leu Phe Leu	305	310	315
Ala Ile Val Leu Val Val Phe Leu Ala Val Pro Phe Leu Glu Leu Val	325	330	335
Met Val Ala Arg Ile Ser Asn Gln Thr Leu Thr Arg Arg Thr Gly Glu	340	345	350
Leu Leu Gln Arg Phe Asp Leu Gln His Ala Asp Ala Arg Phe Lys Gln	355	360	365
Val Val Asn Ala Phe Trp Leu Gln Val Val Thr Ile Asn Tyr Lys Leu	370	375	380
Met Pro Leu Gly Leu Leu Glu Leu Asn Thr Ser Leu Val Asn Lys Val	385	390	395
Phe Ser Ser Ala Ile Gly Ser Leu Leu Ile Leu Ile Gln Ser Asp Leu	405	410	415
Thr Leu Arg Phe Ser Leu Lys	420		

<210> 79

<211> 1299

<212> DNA

<213> Drosophila melanogaster

<220>

<221> CDS

<222> (1) .. (1299)

<223> Coding region GR98B.1

<400> 79

atg gaa gcc aat cgg agt cgt ctg ctg gcc gca gcg cgt cct tac att 48
 Met Glu Ala Asn Arg Ser Arg Leu Leu Ala Ala Ala Arg Pro Tyr Ile
 1 5 10 15

cag att tat tcc att ttc gga ctc acg cgg cca att cag ttt ttt acc 96
 Gln Ile Tyr Ser Ile Phe Gly Leu Thr Pro Pro Ile Gln Phe Phe Thr
 20 25 30

agg acc tta cat aag cga cgt aga gga att gtg ata ttg ggc tac gcc 144
 Arg Thr Leu His Lys Arg Arg Arg Gly Ile Val Ile Leu Gly Tyr Ala
 35 40 45

tgc tat tta att agc att tcc ctg atg gtc atc tat gag tgc tac gcg 192
 Cys Tyr Leu Ile Ser Ile Ser Leu Met Val Ile Tyr Glu Cys Tyr Ala
 50 55 60

aac att gtg gct ctg caa aag gat ata cat aag ttt cac gcc gag gac 240
 Asn Ile Val Ala Leu Gln Lys Asp Ile His Lys Phe His Ala Glu Asp
 65 70 75 80

tct agc aaa gtt atg ggg gaa tac gca gaa agg tcc tgg tgg gta gcc 288
 Ser Ser Lys Val Met Gly Glu Tyr Ala Glu Arg Ser Trp Trp Val Ala
 85 90 95

atg ttc gtt tgg aat caa ttg aac att ctg ctt aac ttt cgg cgc ctt 336
 Met Phe Val Trp Asn Gln Leu Asn Ile Leu Leu Asn Phe Arg Arg Leu
 100 105 110

gct agg att tat gat gat att gcg gat ctg gaa ata gat ttg aat aac 384
 Ala Arg Ile Tyr Asp Asp Ile Ala Asp Leu Glu Ile Asp Leu Asn Asn
 115 120 125

gcc tct agc ggt ttt gtt ggc caa cgg cac tgg tgg cgc ttc cgt ttc 432
 Ala Ser Ser Gly Phe Val Gly Gln Arg His Trp Trp Arg Phe Arg Phe
 130 135 140

cgg ttg gcc ctc tct gtg ggc ctg tgg ata gtg ttg ctg gtg ggt ctc 480
 Arg Leu Ala Leu Ser Val Gly Leu Trp Ile Val Leu Leu Val Gly Leu
 145 150 155 160

acg cca cga ttc acc ctc gtg gca ctc gga ccc tac ctc cac tgg aca 528
 Thr Pro Arg Phe Thr Leu Val Ala Leu Gly Pro Tyr Leu His Trp Thr
 165 170 175

aat aaa gtg ctc acc gaa atc att ctg ata atg cta caa ctt aag tgt 576
 Asn Lys Val Leu Thr Glu Ile Ile Leu Ile Met Leu Gln Leu Lys Cys
 180 185 190

aca gag tat tgt gtg ttt gtg ctc ctg atc tat gaa ctg atc ctc cga 624
 Thr Glu Tyr Cys Val Phe Val Leu Leu Ile Tyr Glu Leu Ile Leu Arg

195	200	205	
ggg cgc cac atc ctt cag cag atc agt gtg gag ctc gag ggt aac cag	672		
Gly Arg His Ile Leu Gln Gln Ile Ser Val Glu Leu Glu Gly Asn Gln			
210 215 220			
tca agg gac agt gtt cag gag ctg tgt gtg gcc ttg aaa cgc aat cag	720		
Ser Arg Asp Ser Val Gln Glu Leu Cys Val Ala Leu Lys Arg Asn Gln			
225 230 235 240			
ttg ctg gct gga cgc att tgg ggc ttg gtg aat gag gtc agc ttg tat	768		
Leu Leu Ala Gly Arg Ile Trp Gly Leu Val Asn Glu Val Ser Leu Tyr			
245 250 255			
ttt acc cta tcc ttg acg ctt ttg ttt ctc tac aat gaa ctg acc att	816		
Phe Thr Leu Ser Leu Thr Leu Leu Phe Leu Tyr Asn Glu Leu Thr Ile			
260 265 270			
ctg caa att gtc aat tgg gct ctc att aaa tcc gtc aat cca aac gaa	864		
Leu Gln Ile Val Asn Trp Ala Leu Ile Lys Ser Val Asn Pro Asn Glu			
275 280 285			
tgc tgt caa tat agt aag tta gtt ttc aag ttc aaa aga aac ttt acc	912		
Cys Cys Gln Tyr Ser Lys Leu Val Phe Lys Phe Lys Arg Asn Phe Thr			
290 295 300			
tat aaa caa gtt att ttc ata ata ggg cgc gtt ggt act tgc ctc ttg	960		
Tyr Lys Gln Val Ile Phe Ile Ile Gly Arg Val Gly Thr Cys Leu Leu			
305 310 315 320			
ctg tca atc aat att ttt cta tcc tgt tta tac agc gag ttc tgc att	1008		
Leu Ser Ile Asn Ile Phe Leu Ser Cys Leu Tyr Ser Glu Phe Cys Ile			
325 330 335			
caa aca tat aat agc att tca cga gtt ctt cac caa atg tat tgc ctt	1056		
Gln Thr Tyr Asn Ser Ile Ser Arg Val Leu His Gln Met Tyr Cys Leu			
340 345 350			
tct gca gcc gaa gat tat cta ata tta aaa atg ggc ctg agg gaa tac	1104		
Ser Ala Ala Glu Asp Tyr Leu Ile Leu Lys Met Gly Leu Arg Glu Tyr			
355 360 365			
tcg ctg caa atg gag cat tta aag ctg att ttc aca tgc ggt ggc ctc	1152		
Ser Leu Gln Met Glu His Leu Lys Leu Ile Phe Thr Cys Gly Gly Leu			
370 375 380			
ttt gac atc aat ctt aag ttc ttc gga ggg atg gta gtc acc tta ttc	1200		
Phe Asp Ile Asn Leu Lys Phe Phe Gly Gly Met Val Val Thr Leu Phe			
385 390 395 400			
ggg tat atc att att ctc gtg caa ttt aaa att caa ttt ttt gct caa	1248		
Gly Tyr Ile Ile Ile Leu Val Gln Phe Lys Ile Gln Phe Phe Ala Gln			
405 410 415			
tca aat ttt atg caa aat att aac agc acc gaa ctg aaa gca tat acc	1296		
Ser Asn Phe Met Gln Asn Ile Asn Ser Thr Glu Leu Lys Ala Tyr Thr			

420

425

430

gCg
Ala

1299

<210> 80
 <211> 433
 <212> PRT
 <213> Drosophila melanogaster

<400> 80
 Met Glu Ala Asn Arg Ser Arg Leu Leu Ala Ala Ala Arg Pro Tyr Ile
 1 5 10 15
 Gln Ile Tyr Ser Ile Phe Gly Leu Thr Pro Pro Ile Gln Phe Phe Thr
 20 25 30
 Arg Thr Leu His Lys Arg Arg Arg Gly Ile Val Ile Leu Gly Tyr Ala
 35 40 45
 Cys Tyr Leu Ile Ser Ile Ser Leu Met Val Ile Tyr Glu Cys Tyr Ala
 50 55 60
 Asn Ile Val Ala Leu Gln Lys Asp Ile His Lys Phe His Ala Glu Asp
 65 70 75 80
 Ser Ser Lys Val Met Gly Glu Tyr Ala Glu Arg Ser Trp Trp Val Ala
 85 90 95
 Met Phe Val Trp Asn Gln Leu Asn Ile Leu Leu Asn Phe Arg Arg Leu
 100 105 110
 Ala Arg Ile Tyr Asp Asp Ile Ala Asp Leu Glu Ile Asp Leu Asn Asn
 115 120 125
 Ala Ser Ser Gly Phe Val Gly Gln Arg His Trp Trp Arg Phe Arg Phe
 130 135 140
 Arg Leu Ala Leu Ser Val Gly Leu Trp Ile Val Leu Leu Val Gly Leu
 145 150 155 160
 Thr Pro Arg Phe Thr Leu Val Ala Leu Gly Pro Tyr Leu His Trp Thr
 165 170 175
 Asn Lys Val Leu Thr Glu Ile Ile Leu Ile Met Leu Gln Leu Lys Cys
 180 185 190
 Thr Glu Tyr Cys Val Phe Val Leu Leu Ile Tyr Glu Leu Ile Leu Arg
 195 200 205
 Gly Arg His Ile Leu Gln Gln Ile Ser Val Glu Leu Glu Gly Asn Gln
 210 215 220
 Ser Arg Asp Ser Val Gln Glu Leu Cys Val Ala Leu Lys Arg Asn Gln
 225 230 235 240

141

WO 00/77208

PCT/US00/16211

Leu Leu Ala Gly Arg Ile Trp Gly Leu Val Asn Glu Val Ser Leu Tyr
245 250 255

Phe Thr Leu Ser Leu Thr Leu Leu Phe Leu Tyr Asn Glu Leu Thr Ile
260 265 270

Leu Gln Ile Val Asn Trp Ala Leu Ile Lys Ser Val Asn Pro Asn Glu
275 280 285

Cys Cys Gln Tyr Ser Lys Leu Val Phe Lys Phe Lys Arg Asn Phe Thr
290 295 300

Tyr Lys Gln Val Ile Phe Ile Ile Gly Arg Val Gly Thr Cys Leu Leu
305 310 315 320

Leu Ser Ile Asn Ile Phe Leu Ser Cys Leu Tyr Ser Glu Phe Cys Ile
325 330 335

Gln Thr Tyr Asn Ser Ile Ser Arg Val Leu His Gln Met Tyr Cys Leu
340 345 350

Ser Ala Ala Glu Asp Tyr Leu Ile Leu Lys Met Gly Leu Arg Glu Tyr
355 360 365

Ser Leu Gln Met Glu His Leu Lys Leu Ile Phe Thr Cys Gly Gly Leu
370 375 380

Phe Asp Ile Asn Leu Lys Phe Phe Gly Gly Met Val Val Thr Leu Phe
385 390 395 400

Gly Tyr Ile Ile Ile Leu Val Gln Phe Lys Ile Gln Phe Phe Ala Gln
405 410 415

Ser Asn Phe Met Gln Asn Ile Asn Ser Thr Glu Leu Lys Ala Tyr Thr
420 425 430

Ala

<210> 81
<211> 1068
<212> DNA
<213> Drosophila melanogaster

<220>
<221> CDS
<222> (1)..(1068)
<223> Coding region GR98B.4

<400> 81
atg gaa caa atg tcg gga gaa ctg cac gct gcc tcg ttg ctt tac atg 48
Met Glu Gln Met Ser Gly Glu Leu His Ala Ala Ser Leu Leu Tyr Met
1 5 10 15

```

cgg cga ctg atg aag tgt ttg gga atg ctg cct ttc ggt cag aat ctg   96
Arg Arg Leu Met Lys Cys Leu Gly Met Leu Pro Phe Gly Gln Asn Leu
      20              25              30

ttc tcg aaa ggg ttt tgc tat gtg cta ctc ttt gtg tca ctg gga ttt   144
Phe Ser Lys Gly Phe Cys Tyr Val Leu Leu Phe Val Ser Leu Gly Phe
      35              40              45

tcg agc tac tgg cgc ttc agc ttc gat tat gaa ttt gac tat gat ttc   192
Ser Ser Tyr Trp Arg Phe Ser Phe Asp Tyr Glu Phe Asp Tyr Asp Phe
      50              55              60

cta aac gat cga ttc tcc agc acc atc gac ctg agc aac ttt gtt gcc   240
Leu Asn Asp Arg Phe Ser Ser Thr Ile Asp Leu Ser Asn Phe Val Ala
      65              70              75              80

cta gtt ttg ggt cat gct att atc gta ttg gag cta ttg tgg gga aat   288
Leu Val Leu Gly His Ala Ile Ile Val Leu Glu Leu Leu Trp Gly Asn
      85              90              95

tgc agc aag gat gtg gat agg caa ctg cag gcg atc cat tcc caa ata   336
Cys Ser Lys Asp Val Asp Arg Gln Leu Gln Ala Ile His Ser Gln Ile
      100             105             110

aaa ctg caa ctg ggt act tcg aat agc acg gat cgt gtg cgg aaa tac   384
Lys Leu Gln Leu Gly Thr Ser Asn Ser Thr Asp Arg Val Arg Lys Tyr
      115             120             125

tgc aat tgg atc tat gga tcc cta atc ata cga tgg cta ata ttt atc   432
Cys Asn Trp Ile Tyr Gly Ser Leu Ile Ile Arg Trp Leu Ile Phe Ile
      130             135             140

gtg ggg act ata tat aaa ctt atc gtt gga ggc tcc aat gtt ctg gat   480
Val Gly Thr Ile Tyr Lys Leu Ile Val Gly Gly Ser Asn Val Leu Asp
      145             150             155             160

gaa ttg tac aga acc cga tat gag atg tgg tcc ata cgc cgt tta tcc   528
Glu Leu Tyr Arg Thr Arg Tyr Glu Met Trp Ser Ile Arg Arg Leu Ser
      165             170             175

ctg cag aaa ttg gcc aag tta cag gca atc cat aac tct ttg tgg cag   576
Leu Gln Lys Leu Ala Lys Leu Gln Ala Ile His Asn Ser Leu Trp Gln
      180             185             190

gct atc cgt tgt ctg gag tgc tac ttt caa ctg agc ctg atc aca ctg   624
Ala Ile Arg Cys Leu Glu Cys Tyr Phe Gln Leu Ser Leu Ile Thr Leu
      195             200             205

ctc atg aag ttc ttc atc gat act tct gct ttg cca tac tgg ctc tac   672
Leu Met Lys Phe Phe Ile Asp Thr Ser Ala Leu Pro Tyr Trp Leu Tyr
      210             215             220

ctc agc aga gtt gag cac aca agg gtg gct gtg cag cac tac gtc gct   720
Leu Ser Arg Val Glu His Thr Arg Val Ala Val Gln His Tyr Val Ala
      225             230             235             240

```

acg gtt gag tgc atc aaa ctc tta gag att gta gtg ccc tgc tat ctc 768
 Thr Val Glu Cys Ile Lys Leu Leu Glu Ile Val Val Pro Cys Tyr Leu
 245 250 255
 tgc acg cga tgt gat gca atg cag cga aag ttc cta tgc atg ttc tac 816
 Cys Thr Arg Cys Asp Ala Met Gln Arg Lys Phe Leu Ser Met Phe Tyr
 260 265 270
 aca gtc act acc gat cga cgt agt agt caa cta aat gca gct cta aga 864
 Thr Val Thr Thr Asp Arg Arg Ser Ser Gln Leu Asn Ala Ala Leu Arg
 275 280 285
 agt cta aac ctt cag ttg agt cag gag aaa tat aaa ttt agt gcc gga 912
 Ser Leu Asn Leu Gln Leu Ser Gln Glu Lys Tyr Lys Phe Ser Ala Gly
 290 295 300
 gga atg gtg gac ata aac aca gaa atg ctg gga aag ttc ttt ttt gga 960
 Gly Met Val Asp Ile Asn Thr Glu Met Leu Gly Lys Phe Phe Phe Gly
 305 310 315 320
 atg atc agc tat att gta atc tgc att cag ttc agc atc aac ttc agg 1008
 Met Ile Ser Tyr Ile Val Ile Cys Ile Gln Phe Ser Ile Asn Phe Arg
 325 330 335
 gcc aaa aaa atg agc aat gag caa atg agt caa aac atc aca tcc aca 1056
 Ala Lys Lys Met Ser Asn Glu Gln Met Ser Gln Asn Ile Thr Ser Thr
 340 345 350
 agt gcc ccc att 1068
 Ser Ala Pro Ile
 355

<210> 82
 <211> 356
 <212> PRT
 <213> Drosophila melanogaster

<400> 82
 Met Glu Gln Met Ser Gly Glu Leu His Ala Ala Ser Leu Leu Tyr Met
 1 5 10 15
 Arg Arg Leu Met Lys Cys Leu Gly Met Leu Pro Phe Gly Gln Asn Leu
 20 25 30
 Phe Ser Lys Gly Phe Cys Tyr Val Leu Leu Phe Val Ser Leu Gly Phe
 35 40 45
 Ser Ser Tyr Trp Arg Phe Ser Phe Asp Tyr Glu Phe Asp Tyr Asp Phe
 50 55 60
 Leu Asn Asp Arg Phe Ser Ser Thr Ile Asp Leu Ser Asn Phe Val Ala
 65 70 75 80
 Leu Val Leu Gly His Ala Ile Ile Val Leu Glu Leu Leu Trp Gly Asn
 85 90 95

Cys Ser Lys Asp Val Asp Arg Gln Leu Gln Ala Ile His Ser Gln Ile
 100 105 110
 Lys Leu Gln Leu Gly Thr Ser Asn Ser Thr Asp Arg Val Arg Lys Tyr
 115 120 125
 Cys Asn Trp Ile Tyr Gly Ser Leu Ile Ile Arg Trp Leu Ile Phe Ile
 130 135 140
 Val Gly Thr Ile Tyr Lys Leu Ile Val Gly Gly Ser Asn Val Leu Asp
 145 150 155 160
 Glu Leu Tyr Arg Thr Arg Tyr Glu Met Trp Ser Ile Arg Arg Leu Ser
 165 170 175
 Leu Gln Lys Leu Ala Lys Leu Gln Ala Ile His Asn Ser Leu Trp Gln
 180 185 190
 Ala Ile Arg Cys Leu Glu Cys Tyr Phe Gln Leu Ser Leu Ile Thr Leu
 195 200 205
 Leu Met Lys Phe Phe Ile Asp Thr Ser Ala Leu Pro Tyr Trp Leu Tyr
 210 215 220
 Leu Ser Arg Val Glu His Thr Arg Val Ala Val Gln His Tyr Val Ala
 225 230 235 240
 Thr Val Glu Cys Ile Lys Leu Leu Glu Ile Val Val Pro Cys Tyr Leu
 245 250 255
 Cys Thr Arg Cys Asp Ala Met Gln Arg Lys Phe Leu Ser Met Phe Tyr
 260 265 270
 Thr Val Thr Thr Asp Arg Arg Ser Ser Gln Leu Asn Ala Ala Leu Arg
 275 280 285
 Ser Leu Asn Leu Gln Leu Ser Gln Glu Lys Tyr Lys Phe Ser Ala Gly
 290 295 300
 Gly Met Val Asp Ile Asn Thr Glu Met Leu Gly Lys Phe Phe Phe Gly
 305 310 315 320
 Met Ile Ser Tyr Ile Val Ile Cys Ile Gln Phe Ser Ile Asn Phe Arg
 325 330 335
 Ala Lys Lys Met Ser Asn Glu Gln Met Ser Gln Asn Ile Thr Ser Thr
 340 345 350
 Ser Ala Pro Ile
 355

<210> 83
 <211> 381
 <212> DNA

<213> *Drosophila melanogaster*

<220>

<221> CDS

<222> (1)..(381)

<223> Partial coding region GR27F.1

<400> 83

cac tgt gct ttt gtg ccc tct ttc tcg tcc cac ttg ctc tcg tat ctg 48
 His Cys Ala Phe Val Pro Ser Phe Ser Ser His Leu Leu Ser Tyr Leu
 1 5 10 15

atc gcc acc att ccc atc gtc gag gac agt aat cga gcc atc aca aag 96
 Ile Ala Thr Ile Pro Ile Val Glu Asp Ser Asn Arg Ala Ile Thr Lys
 20 25 30

agc gag aaa act gga ggc ata gtg cac tcc cta ctc aat aaa ccc aaa 144
 Ser Glu Lys Thr Gly Gly Ile Val His Ser Leu Leu Asn Lys Pro Lys
 35 40 45

agt gct gag tac aag gag aaa ctg cag caa ttc tcc atg cag ttg atg 192
 Ser Ala Glu Tyr Lys Glu Lys Leu Gln Gln Phe Ser Met Gln Leu Met
 50 55 60

cat ctg aaa atc aat ttt act gca gct ggt ctg ttc aac atc gac cgc 240
 His Leu Lys Ile Asn Phe Thr Ala Ala Gly Leu Phe Asn Ile Asp Arg
 65 70 75 80

aca ttg tat ttc acg atc agc ggg gcc ttg acc act tat ctc atc atc 288
 Thr Leu Tyr Phe Thr Ile Ser Gly Ala Leu Thr Thr Tyr Leu Ile Ile
 85 90 95

ttg ctg cag ttc aca tcc aat tcc ccg aac aat ggt tat ggg aat ggc 336
 Leu Leu Gln Phe Thr Ser Asn Ser Pro Asn Asn Gly Tyr Gly Asn Gly
 100 105 110

agc tct tgc tgt gag ccc ttc aat aat atg acg aat cat acg ctt 381
 Ser Ser Cys Cys Glu Pro Phe Asn Asn Met Thr Asn His Thr Leu
 115 120 125

<210> 84

<211> 127

<212> PRT

<213> *Drosophila melanogaster*

<400> 84

His Cys Ala Phe Val Pro Ser Phe Ser Ser His Leu Leu Ser Tyr Leu
 1 5 10 15

Ile Ala Thr Ile Pro Ile Val Glu Asp Ser Asn Arg Ala Ile Thr Lys
 20 25 30

Ser Glu Lys Thr Gly Gly Ile Val His Ser Leu Leu Asn Lys Pro Lys
 35 40 45

WO 00/77208

PCT/US00/16211

Ser Ala Glu Tyr Lys Glu Lys Leu Gln Gln Phe Ser Met Gln Leu Met
50 55 60
His Leu Lys Ile Asn Phe Thr Ala Ala Gly Leu Phe Asn Ile Asp Arg
65 70 75 80
Thr Leu Tyr Phe Thr Ile Ser Gly Ala Leu Thr Thr Tyr Leu Ile Ile
85 90 95
Leu Leu Gln Phe Thr Ser Asn Ser Pro Asn Asn Gly Tyr Gly Asn Gly
100 105 110
Ser Ser Cys Cys Glu Pro Phe Asn Asn Met Thr Asn His Thr Leu
115 120 125

<210> 85
<211> 186
<212> DNA
<213> Drosophila melanogaster

<220>
<221> CDS
<222> (1)..(186)
<223> Partial coding region GR93F.1

<400> 85
ggt gaa aca ttt ctc ggt cag ctg caa acc caa cga ctg gag atc aaa 48
Val Glu Thr Phe Leu Gly Gln Leu Gln Thr Gln Arg Leu Glu Ile Lys
1 5 10 15
gta ttg gga ttt ttc cat cta aat aat gag ttc att ctt ctc att ctg 96
Val Leu Gly Phe Phe His Leu Asn Asn Glu Phe Ile Leu Leu Ile Leu
20 25 30
tct gcc ata ata tcg tac ctg ttt atc ctt att cag ttc ggc att aca 144
Ser Ala Ile Ile Ser Tyr Leu Phe Ile Leu Ile Gln Phe Gly Ile Thr
35 40 45
ggt ggc ttt gag gcg tcc gag gac att aaa aat cgt ttt gat 186
Gly Gly Phe Glu Ala Ser Glu Asp Ile Lys Asn Arg Phe Asp
50 55 60

<210> 86
<211> 62
<212> PRT
<213> Drosophila melanogaster

<400> 86
Val Glu Thr Phe Leu Gly Gln Leu Gln Thr Gln Arg Leu Glu Ile Lys
1 5 10 15
Val Leu Gly Phe Phe His Leu Asn Asn Glu Phe Ile Leu Leu Ile Leu
20 25 30

147

Ser Ala Ile Ile Ser Tyr Leu Phe Ile Leu Ile Gln Phe Gly Ile Thr
 35 40 45

Gly Gly Phe Glu Ala Ser Glu Asp Ile Lys Asn Arg Phe Asp
 50 55 60

<210> 87

<211> 864

<212> DNA

<213> *Drosophila melanogaster*

<220>

<221> CDS

<222> (1)..(864)

<223> Partial coding region GR93F.4

<400> 87

tac tta tcg ata att cac ttg aag atc tgt cat ggc ccg gaa gtg acc 48
 Tyr Leu Ser Ile Ile His Leu Lys Ile Cys His Gly Pro Glu Val Thr
 1 5 10 15

aag ttg gtt aac caa tat ctg cac atc ttt cgc ctg gga acc ctg gac 96
 Lys Leu Val Asn Gln Tyr Leu His Ile Phe Arg Leu Gly Thr Leu Asp
 20 25 30

atc cgt aga aga agt caa ttt gga ggc ggt aga gag ttg ttc ctg cta 144
 Ile Arg Arg Arg Ser Gln Phe Gly Gly Gly Arg Glu Leu Phe Leu Leu
 35 40 45

att ctt tcg gtt tgt tgt cag att cat gaa tat gtt ttt ata ttg gtc 192
 Ile Leu Ser Val Cys Cys Gln Ile His Glu Tyr Val Phe Ile Leu Val
 50 55 60

ata gcg agc aga ctt tgt ggt ttt caa cat att att tgg tgg gtt agc 240
 Ile Ala Ser Arg Leu Cys Gly Phe Gln His Ile Ile Trp Trp Val Ser
 65 70 75 80

tat aca tat gtg ttt att ata tgt aat tcg atc atg tgt ttc ggt ttt 288
 Tyr Thr Tyr Val Phe Ile Ile Cys Asn Ser Ile Met Cys Phe Gly Phe
 85 90 95

att tgg cac cta agc ctg gga gtt ctc tat gct gag ctg aac gat aat 336
 Ile Trp His Leu Ser Leu Gly Val Leu Tyr Ala Glu Leu Asn Asp Asn
 100 105 110

ctt cgc ttt gag tca ggc ttc caa acg gca ttt tta agg aaa cag caa 384
 Leu Arg Phe Glu Ser Gly Phe Gln Thr Ala Phe Leu Arg Lys Gln Gln
 115 120 125

aga att agg gta caa aaa tct atg gca ctt ttc aag gag ata tcc tct 432
 Arg Ile Arg Val Gln Lys Ser Met Ala Leu Phe Lys Glu Ile Ser Ser
 130 135 140

gtg gtt acc tcc ttg cag gac att ttc aat gta cac cta ttt ttg agc 480

148

Val Val Thr Ser Leu Gln Asp Ile Phe Asn Val His Leu Phe Leu Ser
 145 150 155 160

gca ctc ttg aca ctt ctg caa gtt cta gtc gtt tgg tat aaa atg atc 528
 Ala Leu Leu Thr Leu Leu Gln Val Leu Val Val Trp Tyr Lys Met Ile
 165 170 175

atc gat tta ggg ttt tct gac ttt cgg att tgg tca ttc tcg ctt aaa 576
 Ile Asp Leu Gly Phe Ser Asp Phe Arg Ile Trp Ser Phe Ser Leu Lys
 180 185 190

aat ttg atc caa acg ctc ctt cct gtt ttg gcc att caa gaa gcg gca 624
 Asn Leu Ile Gln Thr Leu Leu Pro Val Leu Ala Ile Gln Glu Ala Ala
 195 200 205

aat caa ttt aaa caa act cgg gaa cgt gct ctg gat att ttc ctc gtt 672
 Asn Gln Phe Lys Gln Thr Arg Glu Arg Ala Leu Asp Ile Phe Leu Val
 210 215 220

gga aag tca aag cat tgg atg aaa tcg gtg gaa ata ttt gtc acc cat 720
 Gly Lys Ser Lys His Trp Met Lys Ser Val Glu Ile Phe Val Thr His
 225 230 235 240

ctc aac cta agc gag ttc agg gtt aac ttg ttg ggc ctc ttc aat gtt 768
 Leu Asn Leu Ser Glu Phe Arg Val Asn Leu Leu Gly Leu Phe Asn Val
 245 250 255

tcc aac gaa cta ttt cta ata att gta tct gca atg ttt tgc tat ctg 816
 Ser Asn Glu Leu Phe Leu Ile Ile Val Ser Ala Met Phe Cys Tyr Leu
 260 265 270

gtt ttt gta aca caa tgt gta atc gtg tat cgc aga cgc tat gtt att 864
 Val Phe Val Thr Gln Cys Val Ile Val Tyr Arg Arg Arg Tyr Val Ile
 275 280 285

<210> 88

<211> 288

<212> PRT

<213> Drosophila melanogaster

<400> 88

Tyr Leu Ser Ile Ile His Leu Lys Ile Cys His Gly Pro Glu Val Thr
 1 5 10 15

Lys Leu Val Asn Gln Tyr Leu His Ile Phe Arg Leu Gly Thr Leu Asp
 20 25 30

Ile Arg Arg Arg Ser Gln Phe Gly Gly Gly Arg Glu Leu Phe Leu Leu
 35 40 45

Ile Leu Ser Val Cys Cys Gln Ile His Glu Tyr Val Phe Ile Leu Val
 50 55 60

Ile Ala Ser Arg Leu Cys Gly Phe Gln His Ile Ile Trp Trp Val Ser
 65 70 75 80

Tyr Thr Tyr Val Phe Ile Ile Cys Asn Ser Ile Met Cys Phe Gly Phe
 85 90 95
 Ile Trp His Leu Ser Leu Gly Val Leu Tyr Ala Glu Leu Asn Asp Asn
 100 105 110
 Leu Arg Phe Glu Ser Gly Phe Gln Thr Ala Phe Leu Arg Lys Gln Gln
 115 120 125
 Arg Ile Arg Val Gln Lys Ser Met Ala Leu Phe Lys Glu Ile Ser Ser
 130 135 140
 Val Val Thr Ser Leu Gln Asp Ile Phe Asn Val His Leu Phe Leu Ser
 145 150 155 160
 Ala Leu Leu Thr Leu Leu Gln Val Leu Val Val Trp Tyr Lys Met Ile
 165 170 175
 Ile Asp Leu Gly Phe Ser Asp Phe Arg Ile Trp Ser Phe Ser Leu Lys
 180 185 190
 Asn Leu Ile Gln Thr Leu Leu Pro Val Leu Ala Ile Gln Glu Ala Ala
 195 200 205
 Asn Gln Phe Lys Gln Thr Arg Glu Arg Ala Leu Asp Ile Phe Leu Val
 210 215 220
 Gly Lys Ser Lys His Trp Met Lys Ser Val Glu Ile Phe Val Thr His
 225 230 235 240
 Leu Asn Leu Ser Glu Phe Arg Val Asn Leu Leu Gly Leu Phe Asn Val
 245 250 255
 Ser Asn Glu Leu Phe Leu Ile Ile Val Ser Ala Met Phe Cys Tyr Leu
 260 265 270
 Val Phe Val Thr Gln Cys Val Ile Val Tyr Arg Arg Arg Tyr Val Ile
 275 280 285

<210> 89

<211> 678

<212> DNA

<213> Drosophila melanogaster

<220>

<221> CDS

<222> (1)..(678)

<223> Partial coding region GR97D.1

<400> 89

cgg cga gga aaa ctc cat acc caa ctg gtc acg att tgt ctt tac gcg 48
 Arg Arg Gly Lys Leu His Thr Gln Leu Val Thr Ile Cys Leu Tyr Ala
 1 5 10 15

```

acc gtt ttc ctt aac ata ctt tac ggc gtt tat ctc ggc cgc ttt tcc 96
Thr Val Phe Leu Asn Ile Leu Tyr Gly Val Tyr Leu Gly Arg Phe Ser
      20                25                30

ttc agg cgc aag aag ttt gtg ttt tcc aaa ggg ctc act atc tat agc 144
Phe Arg Arg Lys Lys Phe Val Phe Ser Lys Gly Leu Thr Ile Tyr Ser
      35                40                45

tta ttt gtg gcc acg ttc ttt gcg ctg ttt tac atc tgg aat att tat 192
Leu Phe Val Ala Thr Phe Phe Ala Leu Phe Tyr Ile Trp Asn Ile Tyr
      50                55                60

aat gaa att tcc act ggt cag atc aat ctt cgc gat aca att gga ata 240
Asn Glu Ile Ser Thr Gly Gln Ile Asn Leu Arg Asp Thr Ile Gly Ile
      65                70                75                80

tat tgt tat atg aac gtc tgc gtt tgc cta ttc aac tat gtg acc caa 288
Tyr Cys Tyr Met Asn Val Cys Val Cys Leu Phe Asn Tyr Val Thr Gln
      85                90                95

tgg gaa aaa aca ctg caa ata att cgg ttt cag aat agt gtg cct ctt 336
Trp Glu Lys Thr Leu Gln Ile Ile Arg Phe Gln Asn Ser Val Pro Leu
      100               105               110

ttc aag gtc ctc gat tca ctg gac att tcg gcg atg att gtg tgg cgg 384
Phe Lys Val Leu Asp Ser Leu Asp Ile Ser Ala Met Ile Val Trp Arg
      115               120               125

gca ttt ata tat ggc ttg ctc aag atc gtg ttc tgt ccc ctt atc acc 432
Ala Phe Ile Tyr Gly Leu Leu Lys Ile Val Phe Cys Pro Leu Ile Thr
      130               135               140

tat ata acg ctg att cta tat cat aga cgc tcc att tcg gaa agt caa 480
Tyr Ile Thr Leu Ile Leu Tyr His Arg Arg Ser Ile Ser Glu Ser Gln
      145               150               155               160

tgg acg agt gta acg acc acg aag acc atg ttg cca ctg att gtt tcc 528
Trp Thr Ser Val Thr Thr Lys Thr Met Leu Pro Leu Ile Val Ser
      165               170               175

aat cag ata aat aat tgc ttc ttc ggc ggc ctg gta ctt gca aat ttg 576
Asn Gln Ile Asn Asn Cys Phe Phe Gly Gly Leu Val Leu Ala Asn Leu
      180               185               190

ata ttt gcc gcg gtg aat cgt aag ctg cac ggc att gtc aag gag gcc 624
Ile Phe Ala Ala Val Asn Arg Lys Leu His Gly Ile Val Lys Glu Ala
      195               200               205

aat atg ctg cag tca cct gtc cag atg aat ctc cat aag ccc tac tac 672
Asn Met Leu Gln Ser Pro Val Gln Met Asn Leu His Lys Pro Tyr Tyr
      210               215               220

cgg atg
Arg Met
225
678

```

<210> 90
 <211> 226
 <212> PRT
 <213> Drosophila melanogaster

<400> 90
 Arg Arg Gly Lys Leu His Thr Gln Leu Val Thr Ile Cys Leu Tyr Ala
 1 5 10 15
 Thr Val Phe Leu Asn Ile Leu Tyr Gly Val Tyr Leu Gly Arg Phe Ser
 20 25 30
 Phe Arg Arg Lys Lys Phe Val Phe Ser Lys Gly Leu Thr Ile Tyr Ser
 35 40 45
 Leu Phe Val Ala Thr Phe Phe Ala Leu Phe Tyr Ile Trp Asn Ile Tyr
 50 55 60
 Asn Glu Ile Ser Thr Gly Gln Ile Asn Leu Arg Asp Thr Ile Gly Ile
 65 70 75 80
 Trp Cys Tyr Met Asn Val Cys Val Cys Leu Phe Asn Tyr Val Thr Gln
 85 90 95
 Trp Glu Lys Thr Leu Gln Ile Ile Arg Phe Gln Asn Ser Val Pro Leu
 100 105 110
 Phe Lys Val Leu Asp Ser Leu Asp Ile Ser Ala Met Ile Val Trp Arg
 115 120 125
 Ala Phe Ile Tyr Gly Leu Leu Lys Ile Val Phe Cys Pro Leu Ile Thr
 130 135 140
 Tyr Ile Thr Leu Ile Leu Tyr His Arg Arg Ser Ile Ser Glu Ser Gln
 145 150 155 160
 Trp Thr Ser Val Thr Thr Lys Thr Met Leu Pro Leu Ile Val Ser
 165 170 175
 Asn Gln Ile Asn Asn Cys Phe Phe Gly Gly Leu Val Leu Ala Asn Leu
 180 185 190
 Ile Phe Ala Ala Val Asn Arg Lys Leu His Gly Ile Val Lys Glu Ala
 195 200 205
 Asn Met Leu Gln Ser Pro Val Gln Met Asn Leu His Lys Pro Tyr Tyr
 210 215 220
 Arg Met
 225

<210> 91
 <211> 381
 <212> DNA

<213> Drosophila melanogaster

<220>

<221> CDS

<222> (1)..(381)

<223> Partial coding region GR98B.2

<400> 91

```

gga cgc gct gga ccc ttt ttt cac tgg gtg aat caa gta tta acc caa   48
Gly Arg Ala Gly Pro Phe Phe His Trp Val Asn Gln Val Leu Thr Gln
  1             5             10             15

atc att ctt ata atg ctt caa ctc aaa ggc ccc gag tat tgc cta ttc   96
Ile Ile Leu Ile Met Leu Gln Leu Lys Gly Pro Glu Tyr Cys Leu Phe
      20             25             30

gtc cta ctg gtt tat gaa ctg att cta aga acg cgc cat gtc ctt gag   144
Val Leu Leu Val Tyr Glu Leu Ile Leu Arg Thr Arg His Val Leu Glu
      35             40             45

cag cta aag gat gat ctc gaa gac ttc gac tgc gga gcc agg att cag   192
Gln Leu Lys Asp Asp Leu Glu Asp Phe Asp Cys Gly Ala Arg Ile Gln
      50             55             60

gag ctg tgc gtg act tta aag cag aac caa ttg ctc att gga cga ata   240
Glu Leu Cys Val Thr Leu Lys Gln Asn Gln Leu Leu Ile Gly Arg Ile
      65             70             75             80

tgg aga ttg gtg gat gag att gga gca tat ttc aga tgg tcc atg act   288
Trp Arg Leu Val Asp Glu Ile Gly Ala Tyr Phe Arg Trp Ser Met Thr
      85             90             95

ctg ctg ttt ctc tac aat gga ctt acc att ctg cac gtt gtc aac tgg   336
Leu Leu Phe Leu Tyr Asn Gly Leu Thr Ile Leu His Val Val Asn Trp
      100            105            110

gct atc att aga tcc atc gat cca aac gat tgc tgt caa ctc agt       381
Ala Ile Ile Arg Ser Ile Asp Pro Asn Asp Cys Cys Gln Leu Ser
      115            120            125

```

<210> 92

<211> 127

<212> PRT

<213> Drosophila melanogaster

<400> 92

```

Gly Arg Ala Gly Pro Phe Phe His Trp Val Asn Gln Val Leu Thr Gln
  1             5             10             15

Ile Ile Leu Ile Met Leu Gln Leu Lys Gly Pro Glu Tyr Cys Leu Phe
      20             25             30

Val Leu Leu Val Tyr Glu Leu Ile Leu Arg Thr Arg His Val Leu Glu
      35             40             45

```

WO 00/77208

PCT/US00/16211

Gln Leu Lys Asp Asp Leu Glu Asp Phe Asp Cys Gly Ala Arg Ile Gln
50 55 60
Glu Leu Cys Val Thr Leu Lys Gln Asn Gln Leu Leu Ile Gly Arg Ile
65 70 75 80
Trp Arg Leu Val Asp Glu Ile Gly Ala Tyr Phe Arg Trp Ser Met Thr
85 90 95
Leu Leu Phe Leu Tyr Asn Gly Leu Thr Ile Leu His Val Val Asn Trp
100 105 110
Ala Ile Ile Arg Ser Ile Asp Pro Asn Asp Cys Cys Gln Leu Ser
115 120 125

<210> 93
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Positive
control primer for preparing cDNA

<400> 93
cggatcccta tgtcaagggtg 20

<210> 94
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Positive
control primer for preparing cDNA

<400> 94
gaagagcttc gtgctggtct 20

Claims defining the invention are as follows:

1. An isolated nucleic acid molecule that encodes the amino acid sequence of a *Drosophila* Gustatory Receptor protein or fragment thereof, wherein:

- (i) said *Drosophila* Gustatory Receptor protein is expressed in chemosensory neurons of *Drosophila* labellum or labral sense organ;
- (ii) said nucleic acid molecule is not expressed in *Drosophila* leg or wing chemosensory hairs;
- (iii) said *Drosophila* Gustatory Receptor protein contains seven transmembrane domains as defined by hydropathy analysis; and
- (iv) said *Drosophila* Gustatory Receptor protein or fragment thereof causes activation of a taste chemosensory neuron when stimulated.

2. The isolated nucleic acid molecule of claim 1, wherein the encoded protein or protein fragment contains at least one conserved amino acid as compared to SEQ ID No. 16 by alignment using a BLAST algorithm, wherein said conserved amino acid is selected from the group consisting of:

- a) Serine in the first extracellular amino terminal domain at a position corresponding to position 29 of SEQ ID No. 16;
- b) Phenylalanine in the first transmembrane domain at a position corresponding to position 48 of SEQ ID No. 16;
- c) Arginine in the first extracellular loop at a position corresponding to position 95 of SEQ ID No. 16;
- d) Leucine in the third transmembrane domain at a position corresponding to position 126 of SEQ ID No. 16;
- e) Leucine in the fourth transmembrane domain at a position corresponding to position 151 of SEQ ID No. 16;
- f) Leucine in the third extracellular loop at a position corresponding to position 220 of SEQ ID No. 16;
- g) Glycine in the fifth transmembrane domain at a position corresponding to position 232 of SEQ ID No. 16;
- h) Tyrosine in the fifth transmembrane domain at a position corresponding to position 251 of SEQ ID No. 16;
- i) Phenylalanine in the fourth extracellular loop at a position corresponding to position 321 of SEQ ID No. 16;
- j) Leucine in the fourth extracellular loop at a position corresponding to position 323 of SEQ ID No. 16;
- k) Alanine in the seventh transmembrane domain at a position corresponding to position 334 of SEQ ID No. 16;



- l) Glycine in the seventh transmembrane domain at a position corresponding to position 336 of SEQ ID No. 16;
- m) Phenylalanine in the seventh transmembrane domain at a position corresponding to position 338 of SEQ ID No. 16;
- n) Leucine in the seventh transmembrane domain at a position corresponding to position 340 of SEQ ID No. 16;
- o) Aspartate in the seventh transmembrane domain at a position corresponding to position 341 of SEQ ID No. 16;
- p) Leucine in the seventh transmembrane domain at a position corresponding to position 345 of SEQ ID No. 16;
- q) Alanine in the seventh transmembrane domain at a position corresponding to position 351 of SEQ ID No. 16;
- r) Threonine in the seventh transmembrane domain at a position corresponding to position 354 of SEQ ID No. 16;
- s) Tyrosine in the seventh transmembrane domain at a position corresponding to position 355 of SEQ ID No. 16;
- t) Valine in the seventh transmembrane domain at a position corresponding to position 357 of SEQ ID No. 16;
- u) Leucine in the seventh transmembrane domain at a position corresponding to position 359 of SEQ ID No. 16;
- v) Glutamine in the fourth intracellular carboxy terminal domain at a position corresponding to position 361 of SEQ ID No. 16; and
- w) Phenylalanine in the fourth intracellular carboxy terminal domain at a position corresponding to position 362 of SEQ ID No. 16.

3. The isolated nucleic acid molecule of claim 1 or claim 2, wherein the nucleic acid molecule comprises a sequence selected from the group consisting of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89 and 91.

4. The isolated nucleic acid molecule of any one of claims 1-3, wherein said amino acid sequence of a Drosophila Gustatory Receptor protein comprises a sequence selected from the group consisting of SEQ ID Nos: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90 and 92.

5. The isolated nucleic acid molecule of any one of claims 1-4, wherein said nucleic acid molecule is operably linked to one or more expression control elements.

6. An isolated host cell comprising the nucleic acid molecule of any one of claims 1-5.
7. The isolated host cell of claim 6, wherein the host cell is a prokaryotic host cell or a eukaryotic host cell.
8. A vector comprising the isolated nucleic acid molecule of any one of claims 1-5.
9. An isolated host cell comprising the vector of claim 8.
10. The isolated host cell of claim 9, wherein the host cell is a prokaryotic host cell or a eukaryotic host cell.
11. A method for producing a protein or polypeptide comprising the step of culturing a host cell transformed with the nucleic acid molecule of any one of claims 1-5 under conditions in which the protein or polypeptide encoded by the nucleic acid molecule is expressed.
12. An isolated protein or polypeptide produced by the method of claim 11.
13. An isolated protein or polypeptide encoded by the nucleic acid of any one of claims 1-5.
14. An isolated protein or polypeptide encoded by the nucleic acid of claim 13, wherein protein comprises an amino acid sequence selected from the group consisting of SEQ ID Nos: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90 and 92.
15. An isolated antibody that binds to the polypeptide of claim 13 or 14.
16. The antibody of claim 15 wherein said antibody is a monoclonal or polyclonal antibody.
17. A method of measuring the activity of an agent which modulates the expression of a protein or protein fragment of claim 13 or 14 comprising the steps of:
 - a) exposing cells which express the protein or protein fragment to the agent; and
 - b) determining whether the agent modulates expression of said protein or protein fragment, thereby identifying an agent which modulates the expression of a protein or protein fragment of claim 13 or 14.
18. A method of measuring the activity an agent which modulates at least one activity of a protein or protein fragment of claim 13 or 14 comprising the steps of:

a) exposing cells which express the protein or protein fragment to the agent; and
b) determining whether the agent modulates at least one activity of said protein or protein fragment, thereby identifying an agent which modulates the activity of a protein or protein fragment of claim 13 or 14.

19. A method of measuring the activity an agent which modulates the transcription of the nucleic acid molecule of any one of claims 1-5 comprising the steps of:

a) exposing cells which transcribe the nucleic acid to the agent; and
b) determining whether the agent modulates transcription of said nucleic acid, thereby identifying an agent which modulates the transcription of the nucleic acid molecule of any one of claims 1-5.

20. A method of identifying binding partners for a protein or protein fragment of claim 13 or 14 comprising the steps of:

a) exposing said protein or protein fragment to a potential binding partner; and
b) determining if the potential binding partner binds to said protein or protein fragment, thereby identifying binding partners for the protein or protein fragment.

21. A method of modulating the expression of a nucleic acid encoding a protein or protein fragment of claim 13 or 14 comprising administering an effective amount of an agent which modulates the expression of a nucleic acid encoding the protein or protein fragment.

22. A transgenic insect modified to contain the nucleic acid molecule of any one of claims 1-5.

23. The transgenic insect of claim 22, wherein the nucleic acid molecule contains a mutation that alters expression of the encoded protein.

24. An isolated nucleic acid molecule that hybridizes to the nucleic acid of any one of claims 1-5, under the following conditions: 7% SDS, 0.5 M sodium phosphate buffer at pH 7.2, 1 nM EDTA, pH 8.0 and 55°C, wherein the nucleic acid molecule encodes an amino acid sequence which causes the firing of an gustatory neuron when stimulated.

25. An isolated nucleic acid molecule that hybridizes the nucleic acid of any one of claims 1-5, under the following conditions: 7% SDS, 0.5 M sodium phosphate buffer at pH 7.2, 1 nM EDTA, pH 8.0 and 65°C, wherein the nucleic acid molecule encodes an amino acid sequence which causes the firing of an gustatory neuron when stimulated.

26. An isolated nucleic acid molecule encoding a fragment of at least 25 consecutive amino acids of the protein of claim 13 or 14, wherein the fragment has gustatory receptor activity, or is capable of generating an antibody which binds to the protein of claim 13 or 14.

27. A method of identifying novel gustatory receptor genes comprising the steps of:

a) selecting candidate gustatory receptor genes by screening a nucleic acid database using an algorithm trained to identify seven transmembrane receptors genes;

b) screening said selected candidate gustatory receptor genes by identifying nucleic acid sequences with conserved amino acid residues and intron-exon boundaries common to gustatory receptors, and having open reading frames of sufficient size so as to encode a seven transmembrane receptor; and

c) identifying the novel gustatory receptor genes and measuring the expression of gustatory receptor genes wherein the detection of expression confirms said candidate gustatory gene as an gustatory gene.

28. A method of identifying novel *Drosophila* gustatory receptor genes comprising the steps of:

a) selecting candidate gustatory receptor genes by screening a nucleic acid database for nucleic acid sequences with sufficient homology to at least one known *Drosophila* gustatory receptor gene;

b) screening said selected candidate gustatory receptor genes by identifying nucleic acids with conserved amino acid residues and intron-exon boundaries common to gustatory receptors, and having open reading frames of sufficient size so as to encode a seven transmembrane receptor; and

c) identifying the novel gustatory receptor genes and measuring the expression of gustatory receptor genes wherein the detection of expression in one or both of the labellum and/or labral sense organ confirms said candidate gustatory gene as an gustatory gene.

Dated this 25th day of November 2004.

Yale University

Patent Attorneys for the Applicant:

ALLENS ARTHUR ROBINSON

Patent & Trade Marks Attorneys

Figure 1

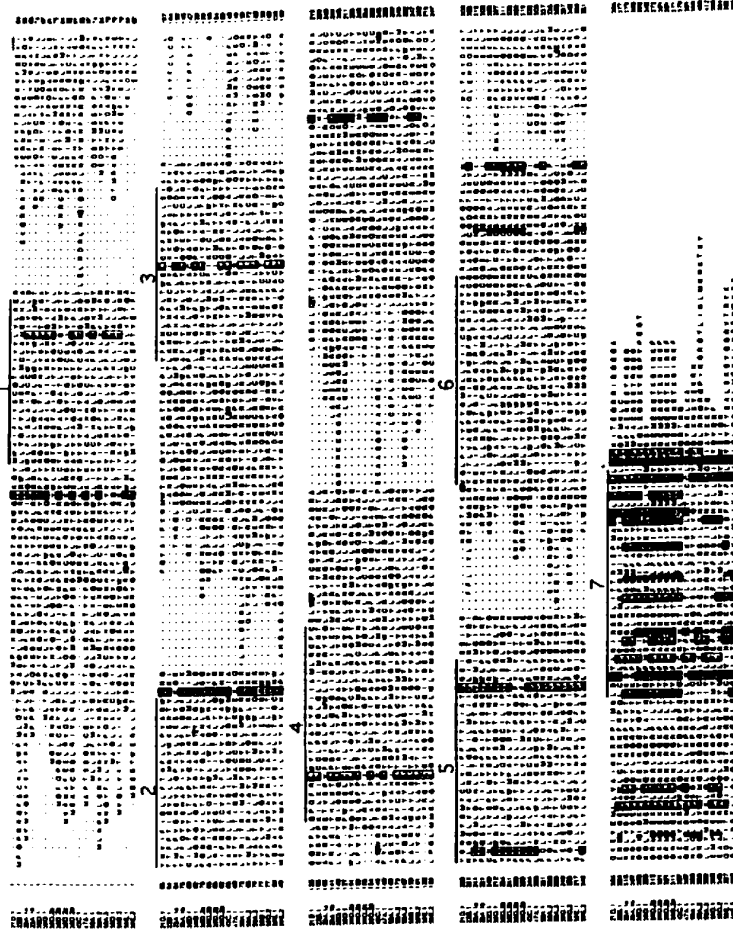


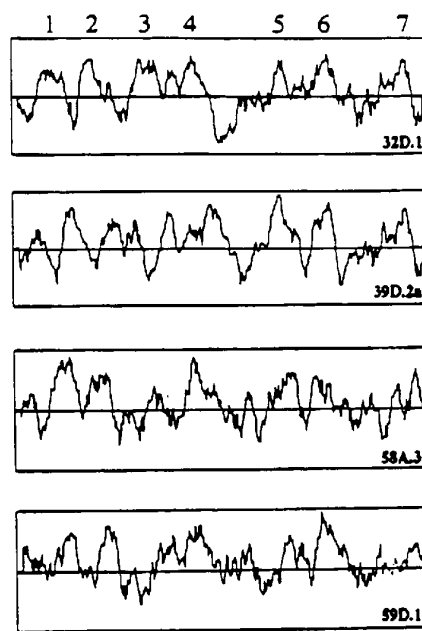
Figure 2

Figure 3

3/5

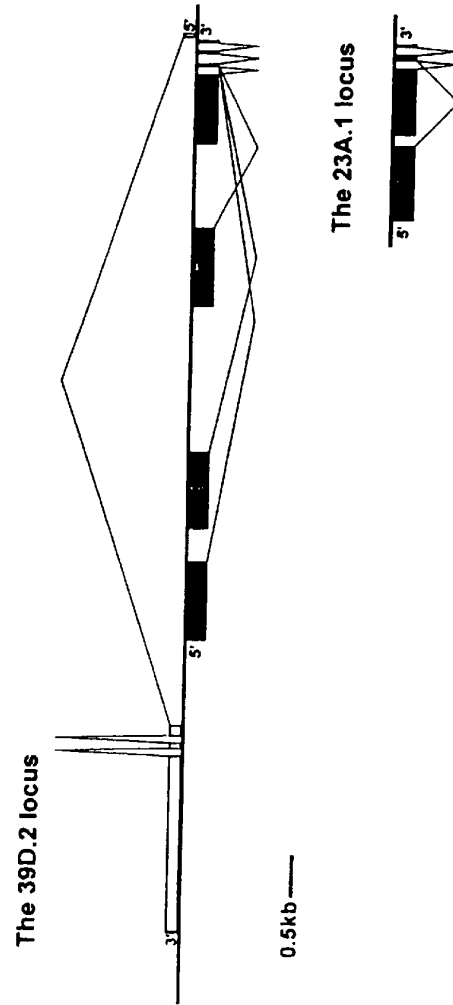


FIG. 4

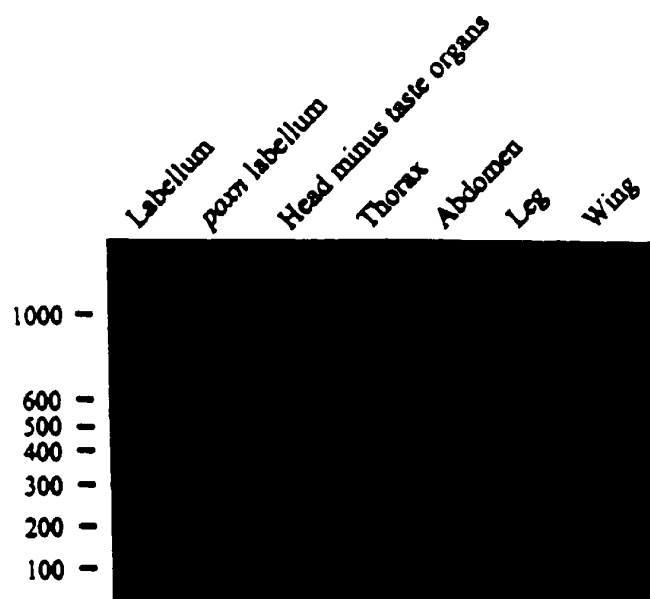


FIG. 5

