



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 315 227**

51 Int. Cl.:

C07H 21/04 (2006.01)

C12P 21/06 (2006.01)

C07K 1/00 (2006.01)

C12N 15/00 (2006.01)

C12N 1/20 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **00915898 .1**

96 Fecha de presentación : **25.02.2000**

97 Número de publicación de la solicitud: **1208111**

97 Fecha de publicación de la solicitud: **29.05.2002**

54 Título: **Genes que codifican los receptores de odorizantes de insecto y usos de los mismos.**

30 Prioridad: **25.02.1999 US 257706**

45 Fecha de publicación de la mención BOPI:
01.04.2009

45 Fecha de la publicación del folleto de la patente:
01.04.2009

73 Titular/es: **The Trustees of Columbia University in
the City of New York
Broadway and West 116th Street
New York, New York 10027-6699, US**

72 Inventor/es: **Vosshall, Leslie, B.;
Amrein, Hubert, O. y
Axel, Richard**

74 Agente: **Elzaburu Márquez, Alberto**

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

Genes que codifican los receptores de odorizantes de insecto y usos de los mismos.

5 Esta solicitud reivindica prioridad y es una solicitud de continuación en parte del documento de Estados Unidos de N° de Serie 09/257.706, presentado el 25 de febrero de 1999.

La invención descrita en este documento se realizó con financiación gubernamental bajo NIH:NIMH, 5P50, MH50733-05 y el NINDS, NS29832-07 del Departamento de Salud y Servicios Humanos. Por consiguiente, el go-
10 bierno de Estados Unidos tiene ciertos derechos en esta invención.

A lo largo de toda esta solicitud, se hace referencia a diversas publicaciones mediante numeración arábica entre paréntesis. Las cintas completas para estas publicaciones se presentan inmediatamente antes de las reivindicaciones.

15 Antecedentes de la invención

Todos los animales poseen una “nariz”, un órgano sensorial olfativo que permite el reconocimiento y la discri-
minación de la información quimiosensorial presente en el entorno. Se cree que los seres humanos, por ejemplo,
reconocen más de 10.000 olores diferentes con un poder discriminatorio exquisito de tal modo que diferencias sutiles
20 en la estructura química pueden conducir con frecuencia a diferencias profundas en la calidad de olor percibida. ¿Qué
mecanismos han evolucionado para permitir el reconocimiento y la discriminación de información olfativa compleja y
cómo se traduce en última instancia la percepción olfativa en respuestas conductuales apropiadas?. El reconociemien-
to de los olores se consigue mediante receptores de odorizantes que residen en cilios olfativos, una especialización
de las dendritas de las neuronas sensoriales olfativas. Los genes de receptores de odorizantes codifican nuevos re-
ceptores serpentina que atraviesan la membrana siete veces. En varias especies de vertebrados y en el invertebrado
25 *Caenorhabditis elegans*, hasta 1.000 genes codifican receptores de odorizantes, sugiriendo que 1-5% del potencial
codificante del genoma en estos organismos está dedicado al reconocimiento de estímulos sensoriales olfativos (Buck
y Axel, 1991; Levy *et al.*, 1991; Parmentier *et al.*, 1992; Ben-Arie *et al.*, 1994; Troemel *et al.*, 1995; Sengupta *et al.*,
1996; Robertson, 1998). Por lo tanto, a diferencia de la visión de color en la que tres fotorreceptores pueden absorber
30 luz por todo el espectro visible, estos datos sugieren que un pequeño número de receptores de odorizantes son insu-
ficientes para reconocer el espectro completo de estructuras moleculares diferentes percibidas por el sistema olfativo.
En su lugar, el sistema sensorial olfativo emplea un número extremadamente grande de receptores, siendo cada uno
capaz de reconocer un pequeño número de ligandos olorosos.

La discriminación de la información olfativa requiere que el cerebro diferencie cuáles de los numerosos receptores
se han activado por un odorizante. En mamíferos, neuronas sensoriales olfativas individuales expresan sólo uno de
miles de genes de receptores de tal modo que las neuronas son funcionalmente diferentes (Ngai *et al.*, 1993; Ressler
et al., 1993; Vassar *et al.*, 1993; Chess *et al.*, 1994; Dulac y Axel, no publicado). Los axones de las neuronas olfativas
que expresan un receptor específico convergen en dos glomérulos espacialmente invariantes entre los 1800 glomérulos
40 dentro del bulbo olfatorio (Ressler *et al.*, 1994; Vassar *et al.*, 1994; Mombaerts *et al.*, 1996; Wang *et al.*, 1998). Por lo
tanto, el bulbo proporciona un mapa espacial que identifica cuáles de los numerosos receptores se han activado dentro
del epitelio sensorial. La calidad de un estímulo olfativo estaría codificada por lo tanto por combinaciones específicas
de glomérulos activados por un odorizante dado.

La lógica de la discriminación olfativa es muy diferente en el nematodo *C. elegans*. A pesar del gran tamaño de
la familia de genes de receptores de odorizantes, los odorizantes volátiles se reconocen únicamente por tres pares
de células quimiosensoriales, de las que cada una expresa probablemente un gran número de genes de receptores
(Bargmann y Horvitz, 1991; Colbert y Bargmann, 1995; Troemel *et al.*, 1995). La activación de uno cualquiera de
los múltiples receptores en una célula conducirá a la quimioatracción, mientras que la activación de receptores en una
50 segunda célula dará como resultado una quimiorrepulsión (Troemel *et al.*, 1997). El circuito neural específico activado
por una neurona sensorial dada es, por lo tanto, el determinante de la respuesta conductual. Por lo tanto, este sistema
sensorial olfativo de invertebrados conserva la capacidad de reconocer una amplia serie de odorizantes pero tiene sólo
un limitado poder discriminatorio.

Los vertebrados generan una representación interna del mundo olfativo externo que debe traducir las característi-
cas del estímulo en información neural. A pesar de elucidarse un mapa espacial preciso, ha sido difícil en vertebrados
discernir cómo se descodifica esta información para relacionar el reconocimiento de olores con respuestas conductua-
les específicas. El análisis genético del comportamiento dirigido por el olfato en invertebrados puede proporcionar en
última instancia un sistema para comprender el vínculo mecánico entre el reconocimiento de olores y el comporta-
60 miento. Los insectos proporcionan un sistema modelo atractivo para el estudio de los acontecimientos periféricos y
centrales en el olfato porque presentan comportamientos dirigidos por el olfato sofisticados bajo el control de un siste-
ma sensorial olfativo que es significativamente más sencillo anatómicamente que el de los vertebrados (Siddiqi, 1987;
Carlson, 1996). El aprendizaje asociativo basado en el olfato, por ejemplo, es fuerte en insectos y da como resultado
modificaciones discernibles en la representación neural de olores en el cerebro (Faber *et al.*, 1998). Por lo tanto, puede
65 ser posible asociar modificaciones en conexiones olfativas definidas con paradigmas *in vivo* para el aprendizaje y la
memoria.

El reconocimiento olfativo en la mosca de la fruta *Drosophila* se consigue mediante pelos sensoriales distribuidos sobre la superficie del tercer segmento antenal y el palpo maxilar. Las neuronas olfativas dentro de pelos sensoriales envían proyecciones a uno de 43 glomérulos dentro del lóbulo antenal del cerebro (Stocker, 1994; Laissue, *et al.*, 1999). Los glomérulos están inervados por dendritas de las neuronas de proyección, el equivalente en insectos de las células mitrales en el bulbo olfatorio de los vertebrados, cuyos cuerpos celulares rodean a los glomérulos. Estas neuronas del lóbulo antenal, a su vez, se proyectan hacia el cuerpo pedunculado y el cuerno lateral del protocerebro (revisado en Stocker, 1994). La localización de la 2-desoxiglucosa en el mapa genómico de la mosca de la fruta (Rodríguez, 1988) y la formación de imágenes con calcio en la abeja (Joerges *et al.*, 1997; Faber *et al.*, 1998) demuestran que diferentes odorizantes provocan patrones definidos de actividad glomerular, lo cual sugiere que en insectos, así como en vertebrados, se representa un mapa topográfico de la calidad del olor en el lóbulo antenal. Sin embargo, en ausencia de los genes que codifican las moléculas receptoras, no ha sido posible definir una base física para este mapa espacial.

En este estudio se identifica una gran familia de genes que probablemente codifican los receptores de odorizantes de *Drosophila melanogaster*. La clonación diferencial, junto con el análisis de secuencias genómicas de *Drosophila*, ha conducido a la identificación de una familia de supuestos receptores de 7 dominios transmembrana codificados probablemente por 100 a 200 genes dentro del genoma de *Drosophila* de acuerdo con la descripción de Clyne *et al.*, 1999, Neuron 22, 327-338. Cada receptor se expresa en un pequeño subconjunto de células sensoriales (0,5-1,5%) que está espacialmente definido dentro de la antena y el palpo maxilar. Además, diferentes neuronas expresan complementos diferentes de genes de receptores de tal modo que las neuronas individuales son funcionalmente diferentes.

La identificación de una gran familia de supuestos receptores de odorizantes en insectos indica que, como en otras especies, la diversidad y especificidad del reconocimiento del olor se proporcionan por una gran familia de genes de receptores. La identificación de la familia de supuestos genes de receptores de odorizantes puede proporcionar una nueva perspectiva de la lógica de la percepción olfativa en *Drosophila*.

Los insectos proporcionan un sistema atractivo para el estudio de la percepción sensorial olfativa. Se ha identificado una nueva familia de proteínas de 7 dominios transmembrana codificadas por 100 a 200 genes, que probablemente representan la familia de receptores de odorizantes de *Drosophila*. Los miembros de esta familia de genes se expresan en subpoblaciones definidas topográficamente de neuronas sensoriales olfativas en la antena o el palpo maxilar. Las neuronas sensoriales expresan diferentes complementos de genes de receptores de tal modo que neuronas individuales son funcionalmente diferentes. El aislamiento de genes de receptores de odorizantes candidatos junto con un análisis genético de comportamientos dirigidos por el olor en insectos puede proporcionar, en última instancia, un sistema para comprender el vínculo mecánico entre el reconocimiento del olor y el comportamiento.

Compendio de la invención

Esta invención proporciona una molécula de ácido nucleico aislada que codifica un receptor de odorizantes de insectos. En una realización, la molécula de ácido nucleico aislada comprende: (a) una de las secuencias de ácido nucleico que se exponen en la Figura 8, (b) una secuencia que es degenerada con respecto a una secuencia de (a) como resultado del código genético; o (c) una secuencia que codifica una de las secuencias de aminoácidos que se exponen en la Figura 8.

Esta invención describe una molécula de ácido nucleico de al menos 12 nucleótidos capaz de hibridar específicamente con la secuencia de la molécula de ácido nucleico descrita anteriormente. Esta invención proporciona un vector que comprende la molécula de ácido nucleico aislada descrita anteriormente. En otra realización, el vector es un plásmido.

Esta invención también proporciona un sistema de vector-hospedador para la producción de un polipéptido que tiene la actividad biológica de un receptor de odorizantes de insecto que comprende el vector descrito anteriormente y una célula hospedadora adecuada.

Esta invención proporciona un método de producción de un polipéptido que tiene la actividad biológica de un receptor de odorizantes de insecto, que comprende cultivar el sistema de vector-hospedador descrito anteriormente en condiciones que permitan la producción del polipéptido y recuperar el polipéptido producido de este modo.

Esta invención también proporciona un receptor de odorizantes de insecto purificado. Esta invención describe además un polipéptido codificado por la molécula de ácido nucleico aislada descrita anteriormente.

Esta invención proporciona un anticuerpo capaz de unirse específicamente a un receptor de odorizantes de insecto. Esta invención también describe un anticuerpo capaz de inhibir de forma competitiva la unión del anticuerpo capaz de unirse específicamente a un receptor de odorizantes de insecto.

Esta invención proporciona un método para identificar un inserto de ADNc que codifica receptores de odorizantes de insecto que comprende: (a) generar una genoteca de ADNc que contiene clones que llevan insertos de ADNc de neuronas sensoriales antenales o del palpo maxilar de un insecto; (b) hibridar ácidos nucleicos de los clones de las genotecas de ADNc generadas en la etapa (a) con una sonda que tiene una secuencia exclusivamente presente en un

ácido nucleico que codifica un polipéptido que comprende aminoácidos consecutivos, cuya secuencia es idéntica a la expuesta para DORA45 en la SEC ID N°: 107; y (c) aislar los ácidos nucleicos hibridados resultantes de tal modo que se identifique el inserto de ADNc que codifica el receptor de odorizantes de insecto.

5 Esta invención también describe insertos de ADNc identificados mediante el método anterior.

Esta invención describe además un método para identificar insertos de ADN que codifican receptores de odorizantes de insectos, que comprende: (a) generar genotecas de ADN que contienen clones que llevan insertos a partir de una muestra que contiene al menos una neurona antenal o del palpo maxilar; (b) poner en contacto los clones de las genotecas de ADNc generadas en la etapa (a) con una molécula de ácido nucleico capaz de hibridar específicamente con la secuencia que codifica un receptor de odorizantes de insecto en condiciones apropiadas que permitan la hibridación de las moléculas de ácido nucleico de los clones y la molécula de ácido nucleico; (c) seleccionar clones que hibriden con la molécula de ácido nucleico; y (d) aislar los clones que lleven los insertos hibridados, identificado de este modo los insertos que codifican los receptores de odorizantes.

15 Esta invención también describe un método para identificar insertos de ADN que codifican receptores de odorizantes de insectos, que comprende: (a) generar genotecas de ADN que contienen clones con insertos de una muestra que contiene al menos una neurona sensorial de la antena o del palpo maxilar; (b) poner en contacto los clones de las genotecas de ADN generadas en la etapa (a) con cebadores de una reacción en cadena de la polimerasa apropiados capaces de unirse específicamente a moléculas de ácido nucleico que codifican receptores de odorizantes en condiciones apropiadas que permitan la amplificación de los insertos hibridados mediante reacción en cadena de la polimerasa; (c) seleccionar los insertos amplificados; y (d) aislar los insertos amplificados, identificando de este modo los insertos que codifican los receptores de odorizantes.

25 Esta invención también proporciona un método para obtener un ácido nucleico que codifica un receptor de odorizantes a partir de un insecto, que comprende: (a) poner en contacto una muestra que contiene ácido nucleico procedente de un insecto con cebadores de una reacción en cadena de la polimerasa que hibridan específicamente con un ácido nucleico que codifica un polipéptido que comprende aminoácidos consecutivos, cuya secuencia es idéntica a la expuesta para DORA45 en la SEC ID N°: 107 en condiciones apropiadas que permitan la hibridación de los cebadores con el ácido nucleico para producir un producto de hibridación; (b) amplificar el producto de hibridación resultante usando una reacción en cadena de la polimerasa; y (c) aislar las moléculas amplificadas, obteniéndose de este modo el ácido nucleico que codifica el receptor de odorizantes a partir del insecto.

35 Esta invención también proporciona un método para transformar células, que comprende transfectar una célula hospedadora con un vector adecuado descrito anteriormente. Esta invención también proporciona células transformadas producidas por el método anterior.

Esta invención proporciona un método para identificar un compuesto capaz de unirse específicamente a un receptor de odorizantes de insecto, que comprende poner en contacto una célula transformada o una fracción de membrana de las células transformadas descritas anteriormente con el compuesto en condiciones que permitan la unión del compuesto con el receptor de odorizantes, detectándose la presencia de cualquiera de dichos compuestos unidos específicamente al receptor y, de este modo, identificándose el compuesto como un compuesto que se une específicamente al receptor de odorizantes de insecto.

45 Esta invención proporciona un método para identificar un compuesto que se une específicamente a un receptor de odorizantes de insecto, que comprende poner en contacto un receptor de odorizantes de insecto purificado con el compuesto en condiciones que permitan la unión del compuesto con dicho receptor purificado, detectándose la presencia de cualquiera de dichos compuestos unidos específicamente al receptor y, de este modo, identificándose el compuesto como un compuesto que se une específicamente a un receptor de odorizantes de insecto.

50 Esta invención también describe un método para identificar un compuesto capaz de activar la actividad de un receptor de odorizantes de insecto, que comprende poner en contacto las células transfectadas o fracciones de membrana de las células transfectadas descritas anteriormente con el compuesto en condiciones que permitan la activación de una respuesta de receptor de odorizantes funcional, indicando la activación del receptor que compuesto es capaz de activar la actividad de un receptor de odorizantes.

Esta invención también describe un método para identificar un compuesto capaz de activar la actividad de un receptor de odorizantes que comprende poner en contacto un receptor de odorizantes de insecto purificado con el compuesto en condiciones que permitan la activación de una respuesta de un receptor de odorizantes funcional, indicando la activación del receptor que el compuesto es capaz de activar la actividad de un receptor de odorizantes. En una realización, el receptor purificado esta incorporado en una bicapa lipídica.

65 Esta invención también describe un método para identificar un compuesto capaz de inhibir la actividad de un receptor de odorizantes, que comprende poner en contacto las células transfectadas o fracciones de membrana de las células transfectadas descritas anteriormente con una cantidad apropiada del compuesto en condiciones que permitan la inhibición de una respuesta de receptor de odorizantes funcional, indicando la inhibición de la respuesta del receptor que el compuesto es capaz de inhibir la actividad de un receptor de odorizantes.

Esta invención describe un método para identificar un compuesto capaz de inhibir la actividad de un receptor de odorizantes, que comprende poner en contacto una cantidad apropiada del receptor de odorizantes de insecto purificado con una cantidad apropiada del compuesto en condiciones que permitan la inhibición de una respuesta de receptor de odorizantes funcional, indicando la inhibición de la respuesta del receptor que el compuesto es capaz de activar la actividad de un receptor de odorizantes. En una realización, el receptor purificado está incorporado en una bicapa lipídica.

Esta invención también describe el compuesto identificado por los métodos descritos anteriormente.

Esta invención describe un método de control de poblaciones de plagas que comprende identificar ligandos odorizantes mediante el método descrito anteriormente que son ligandos odorizantes de alarma y pulverizar el área deseada con los ligandos odorizantes identificados.

Por último, esta invención describe un método de control de una población de plaga que comprende identificar ligandos odorizantes mediante el método descrito anteriormente que interfieren con la interacción entre los ligandos odorizantes y los receptores de odorizantes que se asocian con la fertilidad.

Breve descripción de las figuras

Figura 1

Identificación de genes específicos de la antena y el palpo maxilar poco comunes

Se sometieron fagos específicos de la antena/palpo maxilar candidatos a una escisión *in vivo*, digestión de los ADN plasmídicos de pBLUESCRIPT resultantes con BamHI/Asp718 y electroforesis en geles de agarosa al 1,5%. Se hibridaron transferencias de southern con sondas de ADNc marcadas con ³²P generadas a partir de ARNm de la antena/palpo maxilar (Panel A), ARNm de cabeza menos antena/palpo maxilar (Panel B) o ARNm de cuerpo de hembra virgen (Panel C). El gel teñido con bromuro de etidio se muestra en el Panel D. De los trece clones que se presentan en esta figura, cuatro parecen ser específicos de la antena/palpo maxilar (calles 5, 7, 9 y 11). Sin embargo, sólo dos se expresan de forma selectiva en subconjuntos de células en órganos quimiosensoriales de la mosca adulta. DOR104, un supuesto receptor de odorizantes del palpo maxilar está en la Calle 9. El clon de la Calle 11 (RN106) es homólogo a la lipoproteína y las triglicérido lipasas y se expresa en un dominio restringido en la antena (no se muestran los datos).

Figura 2

Expresión de DOR104 en un subconjunto de neuronas del palpo maxilar

(A) Una sección frontal de un palpo maxilar adulto se hibridó con una sonda de ARN antisentido marcada con digoxigenina y se visualizó con anti-digoxigenina conjugado con fosfatasa alcalina. Se pueden ver siete células que expresan DOR104 en esta sección de 15 µm, que representa aproximadamente un tercio del diámetro del palpo maxilar. Se puntuaron secciones seriadas de múltiples palpos maxilares en relación con la expresión de DOR104 y, en promedio, son positivas para este receptor 20 células por palpo maxilar.

(B) Se tiñeron moscas transgénicas que llevan un transgén indicador de DOR104-lacZ con X-GAL en una preparación montada completa. Los palpos maxilares se diseccionaron de la cabeza y se observaron en una preparación con cubreobjetos aplanada bajo óptica de Normaski, que permite la visualización de las 20 células que expresan DOR104-lacZ.

(C) En esta sección horizontal de un palpo maxilar son visibles dendritas y axones de neuronas que expresan DOR104-lacZ. La expresión de lacZ se visualizó con un anticuerpo primario anti-β-galactosidasa policlonal y un anticuerpo secundario conjugado con CY3. Las secciones se observaron bajo epifluorescencia y se fotografiaron sobre película negra y blanca.

Figura 3

Secuencias de aminoácidos predichas de genes de receptores de odorizantes de Drosophila

Se alinean secuencias de aminoácidos deducidas de 12 genes DOR usando ClustalW (MacVector, Oxford Molecular). Las posiciones predichas de las regiones transmembrana (I-VII) se indican mediante barras por encima del alineamiento. Las identidades de los aminoácidos se marcan con un sombreado oscuro y las similitudes se indican con un sombreado claro. Las secuencias proteicas de DOR87, 53, 67, 104 y 64 se obtuvieron a partir de clones de ADNc. Todas las demás se obtuvieron a partir de predicciones de GENSCAN de organizaciones intrón-exón en ADN genómico, como se indica mediante la letra "g" después del nombre del gen. Se obtuvo un clon de ADNc parcial para DOR62 y se descubrió que tenía una identidad del 100% con la proteína de GENSCAN en la región de los aminoácidos 245-381. Se predijo una extensión de 40 aminoácidos para DOR19 mediante el análisis del GENSCAN. Ésta se reemplazó por un asterisco en el alineamiento, y el aislamiento de clones de ADNc para este receptor resolverá si esta extensión está físicamente presente en la proteína.

Figura 4

Expresión de genes de receptores en regiones espacialmente restringidas de la antena

Cada una de las sondas de ARN antisentido marcadas con digoxigenina contra 8 genes DOR hibrida con un pequeño número de células distribuidas en regiones diferentes de la antena. El número total de células por antena que expresan un receptor dado se obtuvo por recuento de las células positivas en secciones seriadas de múltiples antenas. Existen aproximadamente 20 células positivas por antena para DOR67 (A), 53 (B) y 24 (no se muestran los datos); 15 células positivas para DOR62 (C) y 87 (D); y 10 células positivas para DOR64 (E). El número real de células que se tiñen en estas secciones es un subconjunto de este número total. Con la excepción de DOR53 y DOR67, que presentan una fuerte hibridación cruzada, los genes de receptores probablemente identifican diferentes neuronas olfativas, de tal modo que el número de células que se tiñen con una sonda mixta (F) es igual a la suma de las que se tiñen con las sondas individuales (A-E). La mezcla de DOR53, 67, 62, 87 y 64 marca un total de aproximadamente 60 células por antena. Un total de 34 células se tiñen con la sonda mixta en esta sección de 15 μm . La expresión de los genes ligados DOR71, DOR72 y DOR73 se muestra en los paneles (G), (H) e (I), respectivamente. DOR71 se expresa en aproximadamente 10 células en el palpo maxilar. Se observan cinco células positivas en la sección horizontal en el panel (G). También se examinó la expresión de los otros miembros de este grupo de ligamiento y se descubrió DOR72 en aproximadamente 15 células (de las que 3 se marcan en esta sección) (H) y DOR73 en de 1 a 2 células por antena (I).

Figura 5

Los receptores de odorizantes están restringidos a diferentes poblaciones de neuronas olfativas

(A-C) Las moscas del genotipo C155 *elav-GAL4*; UAS-lacZ expresan lacZ citoplasmático en todas las células neuronales. Los paneles (A-C) muestran imágenes confocales de una sección del palpo maxilar horizontal de una mosca incubada con una sonda de ARN antisentido contra DOR104 (rojo) y un anticuerpo anti- β -galactosidasa (verde). DOR104 reconoce cinco células en esta sección del palpo maxilar (A), expresando todas ellas también *elav-lacZ* (B), como se demuestra por las células amarillas en la imagen combinada en el panel (C).

(D, E) DOR64 y DOR87 se expresan en neuronas no solapantes en la punta de la antena. Se hibridaron sondas de ARN antisentido para DOR64 (digoxigenina-ARN; rojo) y DOR87 (FITC-ARN; verde) en las mismas secciones antenales y se observaron mediante microscopía confocal. El Panel (D) es una superposición digital de imágenes confocales tomadas a intervalos de 0,5 μm a través de una sección de 10 μm de la antena. Las células en diferentes planes focales expresan ambos receptores, pero no se encuentran células con doble marcaje.

(F, G) Una hibridación *in situ* de ARN de dos colores con receptores de odorizantes y proteínas de unión a odorizantes demuestra que estas proteínas se expresan en diferentes poblaciones de células. DOR53 (FITC-ARN; verde) marca unas pocas células interna a la cutícula en el borde medio proximal, mientras que PBPR2 (digoxigenina-ARN; rojo) marca un gran número de células yuxtapuestas a la cutícula por toda la antena (F). La proteína de unión a odorizante más restringida OS-F (digoxigenina-ARN, rojo) también tiñe a células diferentes de las que expresan DOR67 (FITC-ARN; verde) (G).

Figura 6

La expresión de receptores está conservada entre individuos

Se hibridaron secciones frontales de antenas de seis individuos diferentes con sondas de ARN antisentido marcadas con digoxigenina contra DOR53 (A-C) o DOR87 (D-F). DOR53 marca aproximadamente 20 células en el borde medio proximal de la antena, de las que aproximadamente 5 se muestran marcadas en estas secciones. DOR87 se expresa en aproximadamente el mismo número de células en la punta distal. Tanto la posición como el número de células que se tiñen se conservan entre diferentes individuos y no presentan dimorfismo sexual.

Figura 7

Los receptores de odorizantes de Drosophila son muy divergentes

Se digirió ADN genómico de Oregon R aislado de moscas completas con BamHI (B), EcoRI (E) o HindIII (H), se sometió a electroforesis en geles de agarosa al 0,8% y se transfirió a membranas de nitrocelulosa. Las manchas de transferencia se hibridaron con sondas marcadas con ^{32}P obtenidas a partir de ADNc de DOR53 (A), ADNc de DOR67 (B) o fragmentos de ADN generados por RT-PCR a partir de ARNm antenal para DOR24 (C), DOR62 (D) y DOR72 (E). Se observa una fuerte hibridación cruzada de DOR53 y DOR67 tanto a rigurosidad elevada como reducida (A, B), mientras que DOR24, 62 y 72 revelan sólo una única banda de hibridación en cada calle tanto a rigurosidad reducida (C-E) como a rigurosidad elevada (no se muestran los datos).

Figura 8

DOR62, 104, 87, 53, 67, 64, 71g, 72g, 73g, 46, 19g y 24g

- 5 Se describe tanto la secuencia de ácido nucleico de cada DOR como su secuencia de aminoácidos codificada.

Figura 9

- 10 *Análisis de las proyecciones axonales de neuronas receptoras olfativas que expresan un receptor de odorizantes de Drosophila dado*

Resultado: todas las neuronas que expresan un receptor dado envían sus axones a un único glomérulo o una estructura sináptica separada en el centro de procesamiento olfativo del cerebro de la mosca. Este resultado es idéntico al obtenido con receptores de odorizantes del ratón: cada glomérulo se destina a recibir la entrada axonal de neuronas que expresan un receptor de odorizantes dado. Por lo tanto, este resultado refuerza el argumento de que estos genes funcionan de hecho como receptores de odorizantes en *Drosophila*.

20 Figura 10

Alineamientos de ClustalW de dos subfamilias de los receptores de odorizantes de Drosophila las familias DOR53 (A-1 y A-2) y DOR64 (B)

- 25 Esta figura destaca las similitudes de secuencia entre genes DOR que son distintivos de diagnóstico de las proteínas. Los restos que son idénticos en diferentes genes DOR se destacan en negro, mientras que los restos que son similares se destacan en gris.

30 Descripción detallada de la invención

Para facilitar la comprensión de la sección de Procedimientos Experimentales que sigue, se describen ciertos métodos y/o términos que aparecen frecuentemente en Sambrook, *et al.*, (1989).

- 35 A lo largo de toda esta solicitud se usan las siguientes abreviaturas convencionales en toda la memoria descriptiva para indicar nucleótidos específicos:

C = citosina A = adenosina

- 40 T = timidina G = guanosina

Esta invención proporciona una molécula de ácido nucleico aislada que codifica un receptor de odorizantes de insecto. El ácido nucleico incluye, pero no se limita a ADN, ADNc, ADN genómico, ADN sintético o ARN. En una realización, la molécula de ácido nucleico codifica un receptor de odorizantes de *Drosophila*.

- 45 En una realización adicional, la molécula de ácido nucleico aislada comprende: (a) una de las secuencias de ácido nucleico que se exponen en la Figura 8, (b) una secuencia que es degenerada con respecto a una secuencia de (a) como resultado del código genético; o (c) una secuencia que codifica uno de las secuencias de aminoácidos que se expone en la Figura 8.

- 50 Las moléculas de ácido nucleico que codifican un receptor de insecto incluyen moléculas que codifican análogos polipeptídicos, fragmentos o derivados de polipéptidos antigénicos que difieren de las formas de origen natural en términos de la identidad o localización de uno o más restos aminoacídicos (análogos de delección que contienen menos de todos los restos especificados para la proteína, análogos de sustitución en los que uno o más restos especificados se reemplazan por otros restos y análogos de adición en los que uno o más restos aminoacídicos se añaden a una parte terminal o media de los polipéptidos) y que comparten algunas o todas las propiedades de las formas naturales.

- Estas moléculas incluyen, pero sin limitación: la incorporación de codones “preferidos” para la expresión por hospedadores no mamíferos seleccionados; el suministro de sitios para escisión por enzimas endonucleasas de restricción; y el suministro de secuencias iniciales, terminales o intermedias adicionales que faciliten la construcción de vectores expresados fácilmente. Por consiguiente, estos cambios pueden dar como resultado un receptor de odorizantes de insecto modificado. Esta invención tiene la intención de incluir moléculas de ácido nucleico que codifican receptores de odorizantes de insectos modificados. Además, para facilitar la expresión del receptor en diferentes células hospedadoras, puede ser necesario modificar la molécula de tal modo que los receptores expresados puedan alcanzar la superficie de las células hospedadoras. El receptor de odorizantes de insecto modificado debe tener actividades biológicas similares a las del receptor de odorizantes de insecto no modificado. Las moléculas también pueden modificarse para aumentar la actividad biológica del receptor expresado.

Esta invención describe una molécula de ácido nucleico de al menos 12 nucleótidos capaz de hibridar específicamente con la secuencia de la molécula de ácido nucleico descrita anteriormente. En una realización, la molécula de ácido nucleico hibrida con una secuencia única dentro de la secuencia de la molécula de ácido nucleico descrita anteriormente. Esta molécula de ácido nucleico puede ser ADN, ADNc, ADN genómico, ADN sintético o ARN.

Esta invención proporciona un vector que comprende la molécula de ácido nucleico aislada descrita anteriormente. En otra realización, el vector es un plásmido.

En una realización, la molécula de ácido nucleico aislada descrita anteriormente está unida operativamente a un elemento regulador.

Los elementos reguladores necesarios para la expresión incluyen secuencias promotoras para unirse a una ARN polimerasa y secuencias de inicio de la transcripción para la unión al ribosoma. Por ejemplo, un vector de expresión bacteriano incluye un promotor tal como el promotor lac y para el inicio de la transcripción la secuencia de Shine-Dalgarno y el codón de inicio AUG. De forma similar, un vector de expresión eucariota incluye un promotor heterólogo u homólogo para la ARN polimerasa II, una señal de poliadenilación cadena abajo, el codón de inicio AUG y un codón de terminación para la separación del ribosoma. Dichos vectores pueden obtenerse en el mercado o ensamblarse a partir de las secuencias descritas por métodos bien conocidos en la técnica, por ejemplo, los métodos descritos anteriormente para la construcción de vectores en general.

Esta invención también proporciona un sistema de vector-hospedador para la producción de un polipéptido que tiene la actividad biológica de un receptor de odorizantes de insecto que comprende el vector descrito anteriormente y un hospedador adecuado.

Esta invención también proporciona un sistema de vector-hospedador en el que el hospedador adecuado es una célula bacteriana, célula de levadura, célula de insecto o célula animal. La célula hospedadora del sistema de expresión anterior puede seleccionarse del grupo que consiste en células en las que la proteína de interés se expresa normalmente o células extrañas tales como células bacterianas (tales como *E. coli*), células de levadura, células fúngicas, células de insecto, células de nematodo, células vegetales o animales en las que proteína de interés no se expresa normalmente. Las células animales adecuadas incluyen, pero sin limitación, células Vero, células HeLa, células Cos, células Cos, células CV1 y diversas células primarias de mamífero.

Esta invención proporciona un método de producción de un polipéptido que tiene la actividad biológica de un receptor de odorizantes de insecto que comprende cultivar el sistema de vector-hospedador descrito anteriormente en condiciones que permitan la producción del polipéptido y recuperar el polipéptido producido de este modo.

Esta invención también proporciona un receptor de odorizantes de insecto purificado. Esta invención describe además un polipéptido codificado por la molécula de ácido nucleico aislada descrita anteriormente.

Esta invención proporciona un anticuerpo capaz de unirse específicamente a un receptor de odorizantes de insecto. Esta invención también describe un anticuerpo capaz de inhibir de forma competitiva la unión del anticuerpo capaz de unirse específicamente a un receptor de odorizantes de insecto. En una realización, el anticuerpo es monoclonal. En otra realización, el anticuerpo es policlonal.

El anticuerpo monoclonal dirigido contra un receptor de odorizantes de insecto puede comprender, por ejemplo, un anticuerpo monoclonal dirigido contra un epítipo de un receptor de odorizantes de insecto presente en la superficie de una célula. Las secuencias de aminoácidos pueden analizarse por métodos bien conocidos por los especialistas en la técnica para determinar si producen regiones hidrófobas o hidrófilas en las proteínas con las que se construyen. En el caso de proteínas de membrana celular, se sabe bien que las regiones hidrófobas forman la parte de la proteína que se inserta en la bicapa lipídica que forma la membrana celular, mientras que las regiones hidrófilas se localizan en la superficie celular en un entorno acuoso.

Los anticuerpos dirigidos contra un receptor de odorizantes de insecto pueden ser procedentes del suero o monoclonales y se preparan usando métodos bien conocidos en la técnica. Por ejemplo, se preparan anticuerpos monoclonales usando la tecnología de hibridomas por fusión de células B productoras de anticuerpo de animales inmunizados con células de mieloma y selección de la línea celular de hibridoma resultante que produce el anticuerpo deseado. Pueden usarse células tales como células NIH3T3 o células 293 que expresan el receptor como inmunógenos para generar dicho anticuerpo. Como alternativa, pueden prepararse péptidos sintéticos usando máquinas disponibles en el mercado.

Como otra alternativa adicional, puede clonarse y expresarse el ADN, tal como un ADNc o un fragmento del mismo, que codifica el receptor o una parte del receptor. El polipéptido expresado se recupera y se usa como un inmunógeno.

Los anticuerpos resultantes son útiles para detectar la presencia de receptores de odorizantes de insecto o para inhibir la función del receptor en animales vivos, en seres humanos o en tejidos o fluidos biológicos aislados a partir de animales o seres humanos.

Estos anticuerpos pueden ser útiles para identificar o aislar otros receptores de odorizantes de insecto. Por ejemplo, pueden usarse anticuerpos contra el receptor de odorizantes de *Drosophila* para explorar una biblioteca de expresión de cucaracha para un receptor de odorizantes de cucaracha. Dichos anticuerpos pueden ser anticuerpos monoclonales o policlonales mono-específicos contra un receptor de odorizantes de insectos seleccionado. Diferentes bibliotecas de expresión de insectos están fácilmente disponibles y pueden generarse usando tecnologías bien conocidas en la técnica.

Un medio para aislar una molécula de ácido nucleico que codifica un receptor de odorizantes de insecto es sondear una biblioteca con una sonda natural o diseñada artificialmente usando métodos bien conocidos en la técnica. Las sondas pueden ser de ADN o ARN. La biblioteca puede ser de ADNc o ADN genómico.

Esta invención proporciona un método para identificar un inserto de ADNc que codifica receptores de odorizantes de insecto, que comprende: (a) generar una genoteca de ADNc que contiene clones que llevan insertos de ADNc a partir de neuronas sensoriales de la antena o del palpo maxilar de un insecto; (b) hibridar ácidos nucleicos de los clones de las genotecas de ADNc generadas en la etapa (a) con una sonda que tiene una secuencia exclusivamente presente en un ácido nucleico que codifica un polipéptido que comprende aminoácidos consecutivos, cuya secuencia es idéntica a la expuesta para DORA45 en la SEC ID N°: 107; y (c) aislar los ácidos nucleicos hibridados resultantes para identificar de este modo el inserto de ADNc que codifica el receptor de odorizantes de insecto.

En una realización del método anterior, después de la etapa (c), comprende además: (a) amplificar los insertos de los clones seleccionados mediante reacción en cadena de la polimerasa; (b) hibridar los insertos amplificados con sondas a partir de las neuronas de la antena o del palpo maxilar; y (c) aislar los clones que lleven los insertos hibridados identificando de este modo los insertos que codifican los receptores de odorizantes. En una realización, las sondas son sondas de ADNc.

Los cebadores de la reacción en cadena de la polimerasa apropiados pueden seleccionarse a partir de las regiones conservadas de las secuencias de receptores de odorizantes de insectos conocidas. Como alternativa, los cebadores pueden seleccionarse a partir de las regiones que son los sitios activos para la unión de ligandos.

Esta invención también describe insertos de ADNc identificados por el método anterior.

Esta invención describe además un método para identificar insertos de ADN que codifican receptores de odorizantes de insectos, que comprende: (a) generar genotecas de ADN que contienen clones que llevan insertos a partir de una muestra que contiene al menos una neurona antenal o del palpo maxilar; (b) poner en contacto clones de las genotecas de ADNc generadas en la etapa (a) con una molécula de ácido nucleico capaz de hibridar específicamente con la secuencia que codifica un receptor de odorizantes de insecto en condiciones apropiadas que permitan la hibridación de las moléculas de ácido nucleico de los clones y la molécula de ácido nucleico; (c) seleccionar clones que hibriden con la molécula de ácido nucleico; y (d) aislar los clones que llevan los insertos hibridados, identificándose de este modo los insertos que codifican los receptores de odorizantes.

Esta invención también describe un método para identificar insertos de ADN que codifican receptores de odorizantes de insecto, que comprende: (a) generar genotecas de ADN que contienen clones con insertos a partir de una muestra que contiene al menos una neurona sensorial de la antena o del palpo maxilar; (b) poner en contacto los clones de las genotecas de ADN generadas en la etapa (a) con cebadores de reacción en cadena de la polimerasa apropiados capaces de unirse específicamente a moléculas de ácido nucleico que codifican receptores de odorizantes en condiciones apropiadas que permiten la amplificación de los insertos hibridados mediante reacción en cadena de la polimerasa; (c) seleccionar los insertos amplificados; y (d) aislar los insertos amplificados, identificándose de este modo los insertos que codifican los receptores de odorizantes.

Esta invención también proporciona un método para obtener un ácido nucleico que codifica un receptor de odorizantes a partir del insecto, que comprende: (a) poner en contacto una muestra que contiene un ácido nucleico de origen de insecto con cebadores de reacción en cadena de la polimerasa que hibridan específicamente con un ácido nucleico que codifica un polipéptido que comprende aminoácidos consecutivos, cuya secuencia es idéntica a la expuesta para DORA45 en la SEC ID N°: 107 en condiciones apropiadas que permitan la hibridación de los cebadores con el ácido nucleico para producir un producto de hibridación; (b) amplificar el producto de hibridación resultante usando una reacción en cadena de la polimerasa; y (c) aislar las moléculas amplificadas, obteniéndose de este modo el ácido nucleico que codifica el receptor de odorizantes del insecto.

Esta invención también proporciona un método para transformar células que comprende transfectar una célula hospedadora con un vector adecuado descrito anteriormente.

Esta invención también proporciona células transformadas producidas por el método anterior. En una realización, las células hospedadoras son células que habitualmente no expresan receptores de odorizantes. En otra realización, las células hospedadoras están expresando receptores de odorizantes.

Esta invención proporciona un método de identificación de un compuesto capaz de unirse específicamente a un receptor de odorizantes de insecto, que comprende poner en contacto una célula transformada o una fracción de membrana de la célula transformada descrita anteriormente con el compuesto en condiciones que permiten la unión del compuesto con el receptor de odorizantes, detectar la presencia de cualquiera de dichos compuestos unidos específicamente.

mente al receptor e identificar de este modo el compuesto como un compuesto que se une específicamente al receptor de odorizantes de insecto.

Esta invención proporciona un método de identificación de un compuesto que se une específicamente a un receptor de odorizantes de insecto, que comprende poner en contacto un receptor de odorizantes de insecto purificado con el compuesto en condiciones que permiten la unión del compuesto con dicho receptor purificado, detectar la presencia de cualquiera de dichos compuestos unidos específicamente al receptor e identificar de este modo el compuesto como un compuesto que se une específicamente a un receptor de odorizantes de insecto. En una realización, el receptor purificado está incorporado en una bicapa lipídica. El receptor purificado puede estar incorporado en los liposomas con una orientación adecuada para llevar a cabo funciones normales. La tecnología de liposomas se conoce bien en la técnica.

Esta invención también describe un método de identificación de un compuesto capaz de activar la actividad de un receptor de odorizantes de insecto, que comprende poner en contacto las células transfectadas o fracciones de membrana de las células transfectadas descritas anteriormente con el compuesto en condiciones que permiten la activación de una respuesta de receptor de odorizantes funcional, indicando la activación del receptor que el compuesto es capaz de activar la actividad de un receptor de odorizantes.

Esta invención también describe un método de identificación de un compuesto capaz de activar la actividad de un receptor de odorizantes, que comprende poner en contacto un receptor de odorizantes de insecto purificado con el compuesto en condiciones que permitan la activación de una respuesta de receptor de odorizantes funcional, indicando la activación del receptor que el compuesto es capaz de activar la actividad de un receptor de odorizantes. En una realización, el receptor purificado está incorporado en una bicapa lipídica.

Esta invención también describe un método de identificación de un compuesto capaz de inhibir la actividad de un receptor de odorizantes, que comprende poner en contacto las células transfectadas o fracciones de membrana de las células transfectadas descritas anteriormente con una cantidad apropiada del compuesto en condiciones que permitan la inhibición de una respuesta de receptor de odorizantes funcional, indicando la inhibición de la respuesta del receptor que el compuesto es capaz de inhibir la actividad de un receptor de odorizantes.

Esta invención describe un método de identificación de un compuesto capaz de inhibir la actividad de un receptor de odorizantes, que comprende poner en contacto una cantidad apropiada del receptor de odorizantes de insecto purificado con una cantidad apropiada del compuesto en condiciones que permiten la inhibición de una respuesta de un receptor de odorizantes funcional, indicando la inhibición de la respuesta del receptor que el compuesto es capaz de activar la actividad de un receptor de odorizantes. En una realización, el receptor purificado está incorporado en una bicapa lipídica.

En una realización separada del método anterior, el compuesto no se conoce previamente. Esta invención también proporciona el compuesto identificado por los métodos descritos anteriormente.

Esta invención describe un método de control de poblaciones de plagas que comprende identificar ligandos odorizantes mediante el método descrito anteriormente que son ligandos odorizantes de alarma y pulverizar el área deseada con los ligandos odorizantes identificados.

Por último, esta invención describe un método de control de una población de plagas, que comprende identificar ligandos odorizantes mediante el método descrito anteriormente que interfieren con la interacción entre los ligandos odorizantes y los receptores de odorizantes que están asociados con la fertilidad.

Esta invención se entenderá mejor a partir de los procedimientos experimentales que se proporcionan a continuación. Sin embargo, un especialista en la técnica apreciará fácilmente que los métodos específicos y resultados analizados son simplemente ilustrativos de la invención como se describe más completamente en las reivindicaciones que vienen a continuación.

Procedimientos experimentales

Animales de experimentación

Se criaron moscas Oregon R (*Drosophila melanogaster*) sobre medio convencional de harina de maíz-agar-melazas a 25°C. Se inyectaron construcciones transgénicas en embriones yw. Se obtuvieron moscas C155 *elav*-GAL4 de Corey Goodman (Lin y Goodman, 1994) y Gary Struhl suministró la reserva de UAS-lacZ (citoplasmático).

Preparación y exploración diferencial de una genoteca de ADNc de antena/palpo maxilar de Drosophila

Se obtuvieron antenas y palpos maxilares de *Drosophila* por decapitación manual y congelación de 500 moscas adultas y agitación de las antenas y palpos maxilares a través de un tamiz metálico fino. Se preparó ARNm usando un kit de purificación de ARN poliA+ (Stratagene). Se preparó una genoteca de ADNc de antena/palpo maxilar a partir de 0,5 µg de ARNm usando el kit LambdaZA-PIIXR de Stratagene.

En resumen, se cultivaron en placas fagos a baja densidad (500-1000 pfu/placa de 150 mm) y se entrecruzaron por luz UV después de elevar los cultivos por triplicado a Hybond-N+ (Amersham). Se generaron sondas complejas por marcaje con cebado aleatorio (PrimeltII, Stratagene) de ARNm transcrito por transcripción inversa (kit de RT-PCR, Stratagene) a partir de ARNm corporal de hembras adultas vírgenes y los cultivos por duplicado se hibridaron a rigurosidad elevada durante 36 horas (65°C en tampón fosfato sódico 0,5 M [pH 7,3] que contenía albúmina de suero bovino al 1%, SDS al 4% y ADN de esperma de arenque a 0,5 mg/ml). Se exploró previamente el tercer cultivo con una mezcla de todas las OBP/PBP clonadas anteriormente (McKenna *et al.*, 1994; Pikielny *et al.*, 1994; Kim *et al.*, 1998) para eliminar una fuente de clones específicos olfativos abundantes pero no deseada. Se aislaron aproximadamente 5.000 OBP/PBP individuales y clones de fagos negativos de cuerpos de hembras vírgenes, se amplificaron sus insertos mediante PCR con cebadores T3 y T7 y aproximadamente 3 µg de ADN se sometieron a electroforesis en geles de agarosa al 1,5%. Los geles se transfirieron por duplicado a Hybond-N+ (Amersham), los filtros se entrecruzaron por UV y las transferencias de Southern resultantes se sometieron a análisis de Northern inverso usando sondas complejas generadas a partir de ARNm de cuerpo de hembra virgen. Se identificaron aproximadamente 500 clones que no hibridaban con sondas de cuerpo de hembra virgen y se consolidaron en transferencias de Southern secundarias por triplicado. Estas manchas de transferencia se sondaron con sondas complejas obtenidas de ARNm de la antena/palpo maxilar, de la cabeza menos la antena/palpo maxilar y de cuerpo de hembra virgen. Un total de 210 clones negativos con sondas de cabeza menos antena/palpo maxilar y cuerpos de hembra virgen y fuertemente positivos, débilmente positivos o negativos con sondas de la antena/palpo maxilar se analizaron adicionalmente por secuenciación e hibridación *in situ*.

Análisis de secuencias del proyecto genoma de *Drosophila* para proteínas transmembrana

Todas las secuencias genómicas de *Drosophila* se descargaron de manera discontinua en abril de 1998 del Proyecto Genoma de *Drosophila* de Berkeley (Berkeley Drosophila Genoma Project, sin publicar). Las secuencias P1 genómicas se analizaron primero con el programa GENSCAN (Burge y Karlin, 1997; <http://CCR-081.mit.edu/GENSCAN.html>), que predice estructuras de intrones-exones y genera secuencias codificante hipotéticas (CDS) y fases de lectura abiertas. Las proteínas predichas por GENSCAN más cortas de 50 aminoácidos se descartaron. Las fases de lectura abiertas restantes se usaron para buscar supuestas regiones transmembrana mayores de 15 aminoácidos con dos programas que se obtuvieron de los autores y se usaron localmente de modo autónomo (véase Persson y Argos, 1994; Cserzo *et al.*, 1997). El programa Dense Surface Alignment (DAS) está disponible en <http://www.biokemi.su.se/-server/DAS/> o de M. Cserzo (miklos@pugh.bip.bham.ac.uk). TMAP está disponible en <ftp://ftp.ebi.ac.uk/pub/software/unix/> o por contacto con el autor, Bengt Persson (bpn@mmb.ki.se). Se escribieron guiones para aplicar los programas DAS y TMAP de forma repetida a conjuntos de secuencias a escala de genoma. Se eliminaron los genes que mostraban una similitud de secuencias significativa con la base de datos de proteínas no redundante del NCBI usando análisis BLAST (Altschul *et al.*, 1990; Altschul *et al.*, 1997). Todos los guiones necesarios para estos cálculos se escribieron en ANSI C convencional y se procesaron en un SUN Enterprise 3000.

De las 229 nuevas proteínas de *Drosophila* con tres o más regiones transmembrana predichas, 35 no mostraban una similitud de secuencia clara con ninguna proteína conocida y se seleccionaron para un análisis adicional mediante hibridación *in situ*. Se generaron sondas para hibridación *in situ* mediante RT-PCR usando ARNm de la antena/palpo maxilar como molde.

Posiciones de mapa de genes DOR

La posición cromosómica de DOR104 se determinó mediante hibridación *in situ* de una sonda marcada con biotina con material obtenido por la técnica de aplastamiento de cromosoma politeno de glándula salival como se describe (Amrein *et al.*, 1988).

Las posiciones cromosómicas de todos los demás genes DOR se basaron en las asignaciones de cromosomas de los clones P1 con los que mapean, como se determinó por el Proyecto Genoma de *Drosophila* de Berkeley (comunicación personal; <http://www.fruitfly.org>; véase también Hartl *et al.*, 1994; Kimmerly *et al.*, 1996). DOR62 se localiza un cósmido secuenciado por el Proyecto Genoma de *Drosophila* Europeo (sin publicar; <http://edgp.ebi.ac.uk/>; Siden-Klamos *et al.*, 1990).

RECEPTOR	MAPA	POSICIÓN	Nº DE ACCESO DEL CLON P1
DOR62	(X)	2F	62D9 (cósmido de EDGP)
DOR67	(2L)	22A3	DS00676
DOR53	(2L)	22A2-3	DS05342
DOR64	(2L)	23A1-2	DS06400
DOR71	(2L)	33B1-2	DS07071

ES 2 315 227 T3

	DOR72	(2L)	33B1-2	DS07071
	DOR73	(2L)	33B1-2	DS07071
5	DOR87	(2R)	43B1-2	DS08779
	DOR19	(2R)	46F5-6	DS01913
	DOR24	(2R)	47D6-E2	DS00724
10	DOR46	(2R)	56D5-7	DS07462
	DOR104	(3L)	85B	no aplicable

15 *El aislamiento de clones de ADNc de DOR y transferencia de Southern*

Se exploraron 3×10^6 clones de la genoteca de la antena/palpo maxilar descrita anteriormente con sondas de PCR para los genes DOR87, DOR53, DOR67, DOR64 y DOR62. Estaban presentes ADNc a una frecuencia que variaba de 1:200.000 (DOR67) a 1:1.000.000 (DOR62) en la genoteca y sus secuencias eran destacadamente similares a las CDS hipotéticas predichas por el programa GENSCAN. La frecuencia de estos genes es similar a la de DOR104, que está presente en 1:125.000 en la genoteca de la antena/ palpo maxilar. Toda la secuenciación se realizó con kits de secuenciación de ciclo ABI y las reacciones se procesaron en un sistema de secuenciación ABI 310 ó 377.

35 Cinco μg de ADN genómico de Oregon R aislado a partir de moscas completas se digirieron con BamHI, EcoRI o HindIII, se sometieron a electroforesis en geles de agarosa al 0,8% y se transfirieron a membranas de nitrocelulosa Nitropure (Micron Separations Inc.). Las manchas de transferencia se hornearon y se hibridaron con sondas marcadas con ^{32}P obtenidas a partir de sondas de ADNc de DOR53 y DOR67, o fragmentos de PCR de DOR24, DOR62 y DOR72. La hibridación se realizó a 42°C durante 36 horas en SSCP 5 X, Denhardts 10X, ADN de esperma de arenque 500 $\mu\text{g}/\text{ml}$ y formamida al 50% (rigurosidad elevada) o al 25% (rigurosidad reducida) (Sambrook *et al.*, 1989). Las manchas de transferencia se lavaron durante 1 hora en SSC 0,2X, SDS al 0,5% a 65°C (rigurosidad elevada) o SSC 1X, SDS al 0,5% a 42°C (rigurosidad reducida).

Hibridación in situ

35 Se llevó a cabo una hibridación *in situ* de ARN esencialmente como se describe (Schaeren-Wiemers y Gerfin-Moser, 1993). Este protocolo se modificó para incluir detergentes en la mayoría de las etapas para aumentar la sensibilidad y reducir el efecto de fondo. El tampón de hibridación contenía formamida al 50%, SSC 5X, Denhardts 5X, ARNt de levadura a 250 $\mu\text{g}/\text{ml}$, ADN de esperma de arenque a 500 $\mu\text{g}/\text{ml}$, heparina a 50 $\mu\text{g}/\text{ml}$, EDTA 2,5 mM, Tween-20 al 0,1% y CHAPS al 0,25%. Todas las etapas de anticuerpo se realizaron en presencia de Tritón X-100 al 0,1% y la reacción se desarrolló en tampón que contenía Tween-20 al 0,1%. Los portaobjetos se montaron en Glycergel (DAKO) y se observaron con óptica de Normarski.

45 Se llevó a cabo una hibridación *in situ* fluorescente como anteriormente con sondas de ARN marcadas con digoxigenina o FITC. La sonda de digoxigenina se visualizó con anti-digoxigenina de oveja (Boehringer) seguido de anticuerpo de burro anti-CY3 de oveja (Jackson). Se visualizaron las sondas con anticuerpo de ratón anti-FITC (Boehringer) y anticuerpo de cabra anti-ratón Alexa 488 (Molecular Probes) después de la preincubación con suero de cabra normal. Las secciones se montaron en reactivo de Vectashield (Vector Labs) y se observaron en un Microscopio Confocal Biorad 1024.

50 Para el marcaje doble con un marcador neural, los animales del genotipo C155 *elav-Gal4*; UAS-lacZ se seccionaron y se hibridaron primero con una sonda de ARN de DOR104 antisentido marcada con digoxigenina y se desarrollaron como se ha descrito anteriormente. La expresión específica de neurona de lacZ dirigida por el potenciador trap de *elav-Gal4* se visualizó con un anticuerpo policlonal de conejo anti- β -galactosidasa (Organon-Technika/Cappel), visualizado mediante un anticuerpo secundario conjugado de cabra anti-conejo Alexa 488 (Molecular Probes) después de la preincubación con suero de cabra normal.

60 La proporción de neuronas en el tercer segmento antenal se calculó por comparación del número de núcleos que se teñían con el anticuerpo monoclonal 44C11 ELAV (suministrado por cortesía de Lily Jan) y los que se teñían con TOTO-3 (Molecular Probes), un tinte de contraste de ácido nucleico, en varias secciones confocales de múltiples antenas. En promedio, 36% de los núcleos en la antena eran positivos a ELAV.

Construcción de transgén DOR104-lacZ y tinción histoquímica

65 Se aisló un clon genómico que contenía la región codificante de DOR104 y varias kb de secuencia cadena arriba a partir de una biblioteca genómica preparada a partir de moscas isogénicas para el tercer cromosoma (por cortesía de Kevin Moses y Gerry Rubin). Se aislaron aproximadamente 3 kb de ADN inmediatamente cadena arriba del supuesto sitio de inicio de la traducción de DOR104 por PCR y se subclonaron en el vector pCasperUAG β Gal (Thummel *et*

al., 1988). Se llevó a cabo la tinción de actividad β -galactosidasa con preparaciones montadas de cabeza completa esencialmente como se describe en Wang *et al.* (1998). Se incubaron secciones congeladas de palpos maxilares de DOR104-lacZ con un anticuerpo policlonal de conejo anti- β -galactosidasa y como se ha descrito anteriormente.

5 Resultados experimentales

Clonación de receptores de odorizantes candidatos

En experimentos iniciales, se aisló un ADNc que codificaba un supuesto receptor de odorizantes mediante una estrategia de clonación diferencial diseñada para detectar copias de ADNc de ARNm presente en frecuencias extremadamente bajas en una población de ARNm. En la antena y el palpo maxilar, aproximadamente 30% de las células son neuronas olfativas. Si cada neurona expresase sólo uno de 100 genes diferentes de receptor de odorizantes posibles a un nivel de 0,1% del ARNm en una neurona sensorial, entonces un ARNm de un receptor dado se encontraría a una frecuencia de uno en 300.000 en el ARNm de la antena. Si se expresasen 100 genes de receptores diferentes, entonces la familia completa de genes de receptores estaría representada a una frecuencia de uno en 3.000 ARNm. Por lo tanto, se introdujeron modificaciones experimentales en una clonación diferencial convencional para permitir la identificación de ARNm extremadamente poco frecuente cuya expresión se limita a la antena o al palpo maxilar.

En resumen, se exploraron previamente 5000 entradas de una genoteca de ADNc de la antena/palpo maxilar (véase la sección de Procedimientos Experimentales) y después se sometieron a hibridación de transferencia de Southern con sondas de ADNc de la antena/palpo maxilar, de la cabeza menos la antena/palpo maxilar o de ARNm de cuerpo de hembra virgen (véase la Figura 1). Esta hibridación de transferencia de Southern (o Northern inverso) para ADNc candidatos permite la detección de secuencias presentes a una frecuencia de uno en 100.000 en la sonda, una sensibilidad aproximadamente 100 veces mayor que la de la exploración de placas (véase la sección de Procedimientos Experimentales). Este procedimiento condujo a la identificación de múltiples ADNc específicos de la antena/palpo maxilar que se analizaron mediante secuenciación de ADN e hibridación *in situ*. Un ADNc, DOR104 (por *Drosophila* Odorant Receptor) (Figura 1, Calle 9), codifica una supuesta proteína de siete dominios transmembrana con una similitud de secuencia no evidente con receptores serpentina conocidos (Figura 3). La hibridación *in situ* reveló que este ADNc hibrida con aproximadamente 15% de las 120 neuronas sensoriales dentro del palpo maxilar pero no hibrida con neuronas del cerebro ni de la antena. Se muestran siete células que expresan DOR104 en la sección del palpo maxilar frontal en la Figura 2A.

Estas observaciones sugerían que DOR104 puede ser un miembro de una gran familia de genes de receptores de odorizantes dentro del genoma de *Drosophila*. Sin embargo, los presentes inventores fueron incapaces de identificar genes adicionales homólogos a DOR104 mediante hibridación de rigurosidad reducida con genoteca de ADN genómico y ADNc o tras el análisis de genes ligados en un paseo genómico. Por lo tanto, se analizó la base de datos del genoma de *Drosophila* para familias de proteínas de múltiples dominios transmembrana que comparten una similitud de secuencia con DOR104. Se descargaron secuencias que representaban aproximadamente 10% del genoma de *Drosophila* (Proyecto Genoma de *Drosophila* de Berkeley) y se sometieron a análisis GENSCAN (Burge y Karlin, 1997) para predecir la estructura de intrones-exones de todas las secuencias dentro de la base de datos. Se buscaron fases de lectura abiertas de más de 50 aminoácidos para proteínas con 3 o más regiones transmembrana predichas usando los algoritmos de superficie de alineamiento densa (DAS) y TMAP (Persson y Argos, 1994; Cserzo *et al.*, 1997; véase también la sección de Procedimientos Experimentales). De los 229 genes candidatos identificados de esta forma, once codificaban proteínas que definen una nueva familia divergente de supuestas proteínas de siete dominios transmembrana con similitud de secuencia con la secuencia de DOR104. Esta familia de receptores de odorizantes candidatos no comparte ningún motivo de secuencia conservado con familias previamente identificadas de receptores de siete dominios transmembrana. Los clones de ADNc que contenían las regiones codificantes para cinco de los 11 genes identificados por análisis GENSCAN se habían aislado a partir de una genoteca de ADNc de la antena/palpo maxilar y sus secuencias se proporcionan en la Figura 3. Las 6 secuencias proteicas restantes proceden de las predicciones del GENSCAN para la organización de intrones-exones. Su organización se adapta bien a la estructura real determinada a partir de las secuencias de ADNc de otros miembros de la familia de genes (Figura 3).

Los receptores consisten en un dominio N-terminal extracelular corto (habitualmente menor de 50 aminoácidos) y siete supuestos dominios transmembrana. El análisis de los supuestos dominios transmembrana (Kyte y Doolittle, 1982; Persson y Argos, 1994; Cserzo *et al.*, 1997) revela múltiples segmentos hidrófobos, pero no es posible a partir de estos análisis determinar inequívocamente ni el número ni la colocación de los dominios transmembrana. En la actualidad, la presente asignación de dominios transmembrana es por lo tanto provisional.

Los miembros individuales de la familia son divergentes y la mayoría presentan una identidad de aminoácidos del 17-26%. Dos agrupamientos ligados de genes de receptores constituyen pequeñas subfamilias de genes con una conservación de secuencia significativamente mayor. Dos genes ligados, DOR53 y DOR67, presentan una identidad de aminoácidos de 76%, mientras que los tres genes ligados, DOR71, 72 y 73, revelan una identidad de 30-55% (Figura 3; véase a continuación). A pesar de la divergencia, cada uno de los genes comparte motivos comunes cortos en posiciones fijas dentro de la supuesta estructura de siete dominios transmembrana que definen estas secuencias como miembros altamente divergentes de una nueva familia de supuestas moléculas receptoras.

Expresión de la familia de genes DOR en neuronas olfativas

Si esta familia de genes codifica supuestos receptores de odorizantes en la mosca, sería de esperar que otros miembros de la familia además de DOR104 se expresaran también en neuronas sensoriales olfativas. Por lo tanto, se realizó una hibridación *in situ* para examinar el patrón de expresión de receptor de cada uno de los 11 miembros adicionales de la familia de genes en organismos adultos y en desarrollo. En *Drosophila*, las neuronas sensoriales olfativas están restringidas al palpo maxilar y al tercer segmento de la antena. El tercer segmento de la antena está cubierto con aproximadamente 500 cerdas sensoriales o sensilas finas (Stocker, 1994), conteniendo cada una de una a cuatro neuronas (Venkatesh y Singh, 1984). El palpo maxilar está cubierto con aproximadamente 60 sensilas, estando cada una innervada por dos o tres neuronas (Singh y Nayak, 1985). Por lo tanto, el tercer segmento de la antena y el palpo maxilar contienen aproximadamente 1500 y 120 neuronas sensoriales, respectivamente.

Se realizaron experimentos de hibridación *in situ* de ARN con sondas antisentido de ARN marcadas con digoxigenina con cada uno de los 11 nuevos miembros de la familia de genes en condiciones de rigurosidad elevada. Una pareja ligada de genes homólogos, DOR53 y DOR67, presenta hibridación cruzada mientras que los 10 genes restantes no presentan hibridación cruzada en estas condiciones (véase a continuación). Cada uno de los 11 genes hibrida con una pequeña subpoblación (0,5-1,5%) de las 1500 neuronas sensoriales olfativas en el tercer segmento antenal (Figura 4). Un gen, DOR71, expresa en aproximadamente 10% de las neuronas sensoriales en el palpo maxilar pero no en la antena (Figura 4G). No se ha detectado expresión de DOR46 o DOR19 en la antena o el palpo maxilar. La expresión de esta familia de genes sólo se observa en células dentro de la antena y del palpo maxilar. No se observó hibridación en neuronas del cerebro, ni tampoco se observó hibridación en ninguna sección en otra parte de la mosca adulta o en ningún tejido de ninguna fase durante el desarrollo embrionario. Sin embargo, se descubre hibridación en un pequeño número de células en las antenas en desarrollo en la fase de pupa avanzada (no se muestran los datos). No se ha determinado todavía si esta familia de receptores se expresa en el aparato olfativo larvario. Sólo aproximadamente un tercio de las células en el tercer segmento antenal y en el palpo maxilar son neuronas (no se muestran los datos) que están intercaladas con células de soporte sensilares no neuronales y glía. Se han realizado dos experimentos para demostrar que la familia de genes de receptores de siete dominios transmembrana se expresa en neuronas sensoriales en lugar de en células de soporte o glía dentro de la antena y del palpo maxilar. En primer lugar, se desarrollaron programas de detección de anticuerpos fluorescentes de dos colores para colocalizar la expresión de receptores en células que expresan la proteína de unión a ARN específica de neuronas ELAV (Robinow y White, 1988). Una línea de potenciador trap que lleva una inserción de GAL4 en el locus *elav* expresa altos niveles de lacZ en neuronas cuando se cruza con una línea de respuesta UAS-lacZ transgénica (Lin y Goodman, 1994). La detección por anticuerpo fluorescente de lacZ identifica las neuronas sensoriales en una sección horizontal del palpo maxilar (Figura 5B). La hibridación con la sonda del receptor DOR104 revela expresión en 5 de las 12 células positivas para lacZ en una sección horizontal del palpo maxilar (Figura 5A). Todas las células que expresan DOR104 también son positivas para lacZ (Figura 5C), indicando que este receptor se expresa sólo en neuronas.

En un segundo experimento se demostró que los genes de receptores no se expresan en células no neuronales. Las células de soporte de la antena expresan diferentes miembros de una familia de proteínas de unión a odorizantes (McKenna *et al.*, 1994; Pikielny *et al.*, 1994; Kim *et al.*, 1998). Estos genes codifican abundantes proteínas de bajo peso molecular que se piensa que transportan odorizantes a través de la linfa sensilar (revisado en Pelosi, 1994). Ciertos experimentos *in situ* de dos colores con una sonda para la proteína de unión a odorizante *PBPRP2* (Pikielny *et al.*, 1994), revelan hibridación con un gran número de células ampliamente distribuidas por toda la antena (Figura 5F). En la misma sección, sin embargo, la sonda DOR53 hibrida con una subpoblación no solapante de neuronas restringidas al dominio medio-proximal de la antena. En un experimento similar, la hibridación *in situ* con la proteína de unión a odorizante, OS-F (McKenna *et al.*, 1994), identifica una subpoblación espacialmente restringida de células de soporte en la antena, mientras que la sonda DOR67 identifica una población diferente de neuronas en un dominio medio-proximal (Figura 5G). Por lo tanto, los supuestos genes de receptores de odorizantes se expresan en una subpoblación de neuronas sensoriales diferente de las células de soporte que expresan las proteínas de unión a odorizante. En su conjunto, estos datos demuestran que 10 de los 12 miembros de la familia que se han identificado se expresan en pequeñas subpoblaciones de neuronas sensoriales olfativas en la antena y el palpo maxilar.

Patrones definidos espacialmente de expresión de receptores

Los experimentos de hibridación *in situ* revelan que cada receptor se expresa en una subpoblación de neuronas restringida espacialmente en la antena o el palpo maxilar (Figura 4). El número total de células que expresan cada receptor por antena se obtuvo por recuento de las células positivas en secciones seriadas de antenas de múltiples moscas. Estos números se presentan en las leyendas de la Figura 4. DOR67 y 53, por ejemplo, hibridan con aproximadamente 20 neuronas en el borde medio proximal de la antena (Figura 4A y B), mientras que DOR62 y 87 hibridan con subpoblaciones de 20 células en el borde distal de la antena (Figura 4C-D). Aproximadamente 10 células en el dominio distal expresan DOR64 (Figura 4E). Cada uno de los tres genes ligados DOR71, 72 y 73 se expresa en diferentes neuronas. DOR72 se expresa en aproximadamente 15 células antenales (Figura 4H), mientras que DOR73 se expresa en 1 a 2 células en el borde distal de la antena (Figura 4I). Por el contrario, DOR71 se expresa en aproximadamente 10 neuronas del palpo maxilar pero no se detecta en la antena (Figura 4G). Los tres tipos sensilares se representan en un mapa topográfico general por todo el tercer segmento antenal. La región proximal-media, por ejemplo, contiene principalmente sensilas basicónicas. Los receptores expresados en esta región (DOR53 y 67) están por lo tanto probablemente limitados a las sensilas basicónicas grandes. Las regiones más distales contienen una mezcla de los tres tipos de sensilas y, por lo tanto, no es posible a partir de estos datos asignar receptores específicos a tipos sensilares específicos.

El patrón espacial de neuronas que expresan un receptor dado está conservado entre individuos. La hibridación *in situ* con dos sondas de receptores con tres moscas individuales revela que tanto la frecuencia como la distribución espacial de las neuronas que hibridan están conservadas en individuos diferentes (Figura 6). En la actualidad no se puede determinar la precisión de este mapa topográfico y sólo se puede argumentar que receptores dados se expresan en dominios localizados.

En experimentos preliminares se ha demostrado que el patrón espacial de expresión de un receptor, DOR104, puede recapitularse en moscas transgénicas con un fragmento promotor que flanquea el gen de DOR104. La fusión del supuesto promotor de DOR104 (que consiste en 3 kb de ADN 5' inmediatamente adyacente a la región codificante) con el gen indicador lacZ ha permitido visualizar una subpoblación de neuronas que expresan DOR104 dentro del palpo maxilar. Preparaciones montadas completas de las cabezas de moscas transgénicas revelan una pequeña subpoblación de neuronas sensoriales dentro del palpo maxilar cuyos cuerpos celulares presentan un color azul después de la tinción con X-gal (Figura 2B). El número de células positivas, aproximadamente 20 por palpo maxilar, se corresponde bien con el observado para expresión de ARN de DOR104. La tinción inmunofluorescente de secciones con anticuerpos dirigidos contra β -galactosidasa revela más claramente las dendritas y axones de estas neuronas bipolares en el palpo maxilar (Figura 2C). Los niveles de expresión de lacZ en estas líneas transgénicas son reducidos y la amplificación adicional será necesaria para permitir rastrear los axones hasta los glomérulos en el lóbulo antenal. No obstante, los datos sugieren que la información que gobierna el patrón espacial de la expresión de DOR104 en una subpoblación restringida de neuronas del palpo maxilar reside dentro de 3 kb de ADN 5' con respecto al gen de DOR104.

Neuronas individuales expresan complementos diferentes de receptores

El entendimiento de la lógica de la discriminación olfativa en *Drosophila* requerirá una determinación de la diversidad y especificidad de la expresión de receptores en neuronas individuales. En el epitelio olfativo de vertebrados, es probable que una neurona dada exprese sólo un receptor de la familia de 1.000 genes (Ngai *et al.*, 1993; Ressler *et al.*, 1993; Vassar *et al.*, 1993; Chess *et al.*, 1994; Dulac y Axel, sin publicar). En el nematodo *C. elegans*, sin embargo, se piensa que neuronas quimiosensoriales individuales expresan múltiples genes de receptores (Troemel *et al.*, 1995). Las observaciones de los presentes inventores con los supuestos receptores de odorizantes de *Drosophila* indican que una sonda de un receptor dado hibrida con 0,5-1,5% de las neuronas antenales, lo que sugiere que cada célula expresa sólo un subconjunto de genes de receptores. Si se demuestra que cada una de las sondas de receptores diferentes hibrida con subpoblaciones diferentes no solapantes de neuronas, esto proporcionaría pruebas de que las neuronas difieren con respecto a los receptores que expresan.

Se realizó por lo tanto una hibridación *in situ* con una mezcla de cinco sondas de receptores (Figura 4F) o individualmente con cada una de las cinco sondas (Figura 4A-E). Se observó que el número de neuronas olfativas identificadas con la sonda mixta (aproximadamente 60 por antena) se aproxima a la suma de las neuronas positivas detectadas con las cinco sondas individuales. Estos resultados demuestran que receptores individuales se expresan en diferentes poblaciones no solapantes de neuronas olfativas.

Se ha realizado un experimento adicional usando hibridación *in situ* de ARN de dos colores para preguntarse si dos genes de receptores, DOR64 y DOR87, expresados en células intercaladas en la antena distal se expresan en diferentes neuronas. Se marcaron sondas de ARN antisentido para los dos genes con digoxigenina-UTP o FITC-UTP y se usaron en combinaciones de dos en dos en hibridación *in situ* con secciones por toda la antena de *Drosophila*. Aunque estos dos genes se expresan en dominios lateral-distal solapantes, la hibridación *in situ* de dos colores revela que las neuronas que expresan DOR64 no expresan DOR87, más bien cada gen se expresa en diferentes poblaciones celulares (Figura 5D y E). En su conjunto, estos datos sugieren que las neuronas sensoriales olfativas dentro de la antena son funcionalmente diferentes y expresan complementos diferentes de receptores de odorizantes. En última instancia, los experimentos concuerdan con un modelo en el que neuronas individuales expresan sólo un gen de receptor único.

El procedimiento de clonación diferencial identificó un gen adicional, A45, que comparte una identidad débil (24%) con la familia de genes DOR a lo largo de una región corta (93 aminoácidos). Este gen, sin embargo, no parece ser un miembro clásico de la familia DOR: es mucho más divergente y significativamente más grande que los otros miembros de la familia (486 aminoácidos). Este gen se expresa en todas las neuronas sensoriales olfativas (no se muestran los datos). Si A45 codifica un receptor de odorizantes divergente, entonces estaría presente en todas las neuronas sensoriales junto con diferentes complementos de los miembros más clásicos de la familia de genes DOR.

El tamaño y organización de la familia de genes de receptores de odorizantes

¿Cómo de grande es la familia de genes de receptores de odorizantes en *Drosophila*? A diferencia de los receptores de odorizantes en vertebrados, que comparten una identidad de secuencia de 40-98% a nivel de los aminoácidos, los receptores de mosca son extremadamente divergentes. El grado de similitud de secuencia entre subfamilias de receptores varía de 20-30%. El receptor del palpo maxilar DOR104 es el miembro relacionado más distantemente de la familia con una identidad de aproximadamente 17% con los otros genes de receptores. La inspección de las secuencias de los receptores sugiere que es poco probable que las hibridaciones de transferencia de Southern, incluso las realizadas a rigurosidad reducida, revelen múltiples miembros adicionales de una familia de genes. De acuerdo con esto, la hibridación en transferencia de Southern con sondas de receptores DOR24, 62 y 72, realizadas a rigurosidad elevada o reducida, revela sólo una banda de hibridación única después de la escisión del ADN genómico con tres endonucleasas de restricción diferentes (Figura 7C-E). Los dos agrupamientos ligados de receptores contienen genes

con un mayor grado de conservación de secuencia y definen pequeñas subfamilias de genes de receptores. Un grupo de tres receptores, DOR71, 72 y 73, se localiza en la posición del mapa 33B1-2. Los receptores de la antena DOR72 y 73 tienen una identidad de 55% y ambos presentan una identidad de aproximadamente 30% con el tercer gen en el locus, DOR71, que se expresa en el palpo maxilar. DOR67 y DOR53, miembros de una segunda subfamilia, residen con una distancia dentro de 1 kb entre sí en la posición de mapa 22A2-3 y presentan una identidad de secuencia de 76%. No es sorprendente que estos dos genes ligados presenten hibridación cruzada a rigurosidad reducida. Las transferencias de Southern sondadas con DOR67 o DOR53 revelan dos bandas de hibridación que se corresponden con los dos genes dentro de la subfamilia pero no detectan miembros adicionales de la subfamilia en el cromosoma (Figura 7A y B).

Los miembros de la familia de genes de receptores descritos en este documento están presentes en todos menos en el cuarto cromosoma pequeño. No se observa predisposición hacia regiones teloméricas o centroméricas. Las posiciones del mapa, según se determinan a partir de los clones P1 y cosmídico (Proyecto Genoma de *Drosophila* de Berkeley; Proyecto Genoma de *Drosophila* Europeo) se proporcionan en los Procedimientos Experimentales. Un número comparativamente grande de genes receptores se localizan en el cromosoma 2 porque el Proyecto Genoma de *Drosophila* de Berkeley ha concentrado sus esfuerzos en este cromosoma. A diferencia de la distribución de receptores de odorizantes en nematodos y mamíferos (Ben-Arie *et al.*, 1994; Troemel *et al.*, 1995; Robertson, 1998), sólo se han identificado pequeñas series ligadas y la mayoría de los miembros de la familia se aíslan en múltiples loci dispersos en el genoma de *Drosophila*.

El alto grado de divergencia entre miembros de la familia de genes de receptores de odorizantes de *Drosophila* recuerda más a la familia de quimiorreceptores en *C. elegans* que los receptores de odorizantes más conservados de vertebrados. Las estimaciones del tamaño de la familia de genes de receptores de *Drosophila*, por lo tanto, no pueden obtenerse ni por hibridación de transferencia de Southern ni por análisis de PCR de ADN genómico. En su lugar, las estimaciones de la familia de genes proceden de la estadística de pequeños números. Se detectaron 12 miembros de la familia de genes de receptores de odorizantes de una base de datos del genoma de *Drosophila* que incluye aproximadamente 10% del genoma. Reconociéndose una posible predisposición en la estimación de los presentes inventores, parece razonable en la actualidad estimar que la familia de receptores de odorizantes incluye probablemente de 100 a 200 genes. Esto está de acuerdo con estimaciones independientes a partir de los experimentos de hibridación *in situ* que demuestran que una sonda de receptor dada hibrida con 0,5-1,5% de las neuronas. Si se asume que una neurona dada expresa sólo un gen de receptor único, estas observaciones sugieren que la familia de genes incluiría de 100 a 200 miembros.

Discusión experimental

El tamaño y la divergencia de la familia de genes

Se ha identificado una nueva familia de proteínas de siete dominios transmembrana que probablemente codifican los receptores de odorizantes de *Drosophila*. El número de genes de receptores diferentes expresados en las neuronas de la antena y el palpo maxilar reflejará la diversidad y especificidad del reconocimiento del olor en la mosca de la fruta. ¿Cómo de grande es la familia de genes de receptores de odorizantes de *Drosophila*? Se han identificado 11 miembros de esta familia de genes divergente en la base de datos de ADN de *Drosophila*. A pesar de la posibilidad de predisposición, parece razonable asumir que como sólo se ha depositado 10% de la secuencia genómica, probablemente esta familia de genes contiene de 100 a 200 genes. Sin embargo, errores significativos en las estimaciones podrían ser el resultado de una predisposición en la naturaleza de las secuencias representadas en 10% del genoma de *Drosophila* analizado hasta la fecha. Los experimentos de hibridación *in situ* que demuestran que cada uno de los genes de receptores marca de 0,5 a 1,5% de las neuronas sensoriales olfativas están de acuerdo con la estimación de 100 a 200 genes de receptores.

Se han identificado varias familias de genes de receptores de odorizantes divergentes, codificando cada una proteínas de siete dominios transmembrana, en especies de vertebrados e invertebrados. En mamíferos, los olores volátiles se detectan por una familia de hasta 1.000 receptores expresados cada uno en el epitelio olfativo principal (Buck y Axel, 1991; Levy *et al.*, 1991; Parmentier *et al.*, 1992; Ben-Arie *et al.*, 1994). Esta familia de genes comparte características con los receptores serpentina de neurotransmisores y está conservada en todos los vertebrados examinados. Los vertebrados terrestres tienen un segundo sistema olfativo anatómicamente y funcionalmente diferente, el órgano vomeronasal, dedicado a la detección de feromonas. Las neuronas sensoriales vomeronasales expresan dos familias diferentes de receptores y se cree que cada una contiene de 100 a 200 genes: una nueva familia de receptores serpentina (Dulac y Axel, 1995) y una segunda familia relacionada con los receptores de neurotransmisores metabotrópicos (Herrada y Dulac, 1997; Matsunami y Buck, 1997; Ryba y Tirindelli, 1997).

En el invertebrado *C. elegans*, los receptores quimiosensoriales se organizan en cuatro familias de genes que comparten una similitud de secuencia de 20-40% dentro de una familia y esencialmente no tienen similitudes de secuencia entre familias (Troemel *et al.*, 1995; Sengupta *et al.*, 1996; Robertson, 1998). Las cuatro familias de genes en *C. elegans* en conjunto contienen aproximadamente 1.000 genes implicados en la detección de olores. Los receptores de nematodos no presentan conservación de secuencia con las tres familias diferentes de genes de receptores de odorizantes de vertebrados. Los estudios revelan que la *Drosophila* ha desarrollado una familia de genes divergente adicional de receptores serpentina compuesta por 100 a 200 genes. La observación de que se logra una función similar, la detección quimiosensorial, mediante al menos ocho familias de genes altamente divergentes que comparten poca o ninguna similitud de secuencia es muy inusual.

¿Por qué la necesidad evolutiva de receptores de odorizantes se satisface con frecuencia por el reclutamiento de nuevas familias de genes en lugar de aprovechar las familias de receptores de odorizantes preexistentes en genomas ancestrales?. El carácter de los odorizantes naturales junto con sus propiedades físicas (por ejemplo, acuosa o volátil) representa un agente de selección importante que dirige la evolución de familias de genes de receptores. El uso de conjuntos de odorizantes “antropomórficos” comunes en los análisis experimentales de especificidad olfativa ha conducido a que prevalezca el punto de vista de que existe un solapamiento significativo en el repertorio de olores percibidos entre diferentes especies. Los estudios de especificidad de odorizantes en diferentes especies con frecuencia emplean olores a concentraciones elevadas artificialmente y pueden presentar una imagen imprecisa del repertorio natural de odorizantes. Simplemente no se conoce la naturaleza de los olores que inicialmente condujeron a la selección ancestral de genes de receptores durante la evolución de las especies de nematodos, insectos o vertebrados. Claramente, propiedades enormemente diferentes en olores salientes pudieron imponer el reclutamiento de nuevas familias de genes para efectuar una función antigua, el olfato. El carácter del olor no es el único agente de selección evolutivo. Los receptores de odorizantes deben interactuar con otros componentes en la ruta de transducción de señales [Proteínas G (para una revisión, véase Buck, 1996; Bargmann y Kaplan, 1998) y quizás incluso RAMP (McLatchie *et al.*, 1998) y rho (Mitchell *et al.*, 1998)] que pueden dirigir la selección de una familia de receptores serpentina con respecto a otra. Además, los receptores de mamíferos no sólo reconocen odorizantes en el entorno, sino que probablemente reconocen indicaciones de guía que dirigen la formación de un mapa sensorial en el cerebro (Wang *et al.*, 1998). Por lo tanto, las múltiples propiedades que requieren los receptores de odorizantes pueden cambiar enormemente a lo largo del tiempo evolutivo y esto puede subyacer a los orígenes independientes de las múltiples familias de genes de receptores quimiosensoriales.

Establecimiento de un mapa topográfico en la antena y en el cerebro

Se observó que los genes de receptores individuales en la mosca se expresan en dominios topográficamente conservados dentro de la antena. Esta distribución espacial altamente ordenada de expresión de receptores difiere de la observada en el epitelio olfativo del mamífero. En mamíferos, un receptor dado puede expresarse en una de cuatro zonas amplias pero circunscritas en el epitelio olfativo principal (Ressler *et al.*, 1993; Vassar *et al.*, 1993). Una zona dada puede expresar hasta 250 receptores diferentes y neuronas que expresan un receptor dado dentro de una zona parecen estar dispersas aleatoriamente (Ressler *et al.*, 1993; Vassar *et al.*, 1993). El patrón altamente ordenado de expresión observado en la antena de *Drosophila* podría tener implicaciones importantes para la generación de patrones de las proyecciones hacia el lóbulo antenal. En los sistemas visual, somatosensorial y auditivo, la lámina de receptores periféricos está muy ordenada y se mantienen relaciones vecinas en la periferia en las proyecciones hacia el cerebro. Estas observaciones sugieren que la posición relativa de la neurona sensorial en la periferia determinará el patrón de proyecciones hacia el cerebro.

Los datos sobre la conservación espacial de la expresión de receptores en la antena sugiere que la superposición sobre la generación de patrones espaciales generales de sensilas olfativas (Venkatesh y Singh, 1984; Ray y Rodrigues, 1995; Reddy *et al.*, 1997) debe ser una información posicional más precisa que gobierne la elección de la expresión de receptores. Esta información espacial podría imponer el patrón topográfico fijo de expresión de receptores en la lámina periférica de receptores y al mismo tiempo gobernar las proyecciones sensoriales ordenadas hacia el cerebro. Esta relación entre la identidad posicional y el patrón de proyecciones neuronales se ha sugerido tanto para las neuronas sensoriales periféricas (Merriitt y Whittington, 1995; Grillenzoni *et al.*, 1998) como para las neuronas en el sistema nervioso central embrionario de *Drosophila* (Doe y Skeath, 1996).

Implicaciones para el procesamiento sensorial

En mamíferos, las neuronas olfativas expresan sólo uno de miles de genes de receptores de odorizantes. Las neuronas que expresan un receptor dado se proyectan con precisión hacia dos de los 1800 glomérulos en el bulbo olfatorio de ratón. Los odorizantes inducirán, por lo tanto, patrones definidos espacialmente de actividad glomerular de tal modo que la calidad de un estímulo olfativo esté codificada por la activación de una combinación específica de glomérulos (Stewart *et al.*, 1979; Lancet *et al.*, 1982; Kauer *et al.*, 1987; Imamura *et al.*, 1992; Mori *et al.*, 1992; Katoh *et al.*, 1993; Friedrich y Korsching, 1997). Además, la capacidad de un odorizante para activar una combinación de glomérulos permite la discriminación de una serie diversa de olores que supera con mucho el número de receptores y sus glomérulos asociados. En el nematodo, una familia igualmente grande de genes de receptores se expresa en 16 pares de células quimiosensoriales, respondiendo sólo tres de ellas a odorizantes volátiles (Bargmann y Horvitz, 1991; Bargmann *et al.*, 1993). Esto implica inmediatamente que una neurona quimiosensorial dada expresará múltiples receptores y que la diversidad de olores reconocidos por el nematodo puede aproximarse a la de los mamíferos, pero el poder discriminatorio se reduce necesariamente de manera drástica.

¿Qué aporta la naturaleza de la familia de genes que se ha identificado en *Drosophila* acerca de la lógica del procesamiento olfativo en este organismo?. Se estima que los receptores de odorizantes de *Drosophila* comprenden una familia de 100 a 200 genes. Además, el patrón de expresión de estos genes en el tercer segmento antenal sugiere que las neuronas sensoriales individuales expresan una dotación de receptores diferente y, en última instancia, los datos de los presentes inventores concuerdan con la sugerencia de que las neuronas individuales expresan uno o un pequeño número de receptores. Como en el caso de los mamíferos, el problema de la discriminación de olores se reduce por lo tanto a un problema de distinción por parte del de qué receptores se han activado por un odorizante dado. Si el número de tipos diferentes de neuronas supera el número de glomérulos (43) (Stocker, 1994; Laissue *et al.*, 1999) inmediatamente se deduce que un glomérulo dado debe recibir la entrada de más de una clase de neurona sensorial.

Esto implica que un solo glomérulo integrará múltiples estímulos olfativos. Una consecuencia posible de este modelo sería una pérdida de poder discriminatorio mientras se mantiene la capacidad de reconocer una amplia serie de olores. Como alternativa, podría producirse un procesamiento significativo de entradas sensoriales en el lóbulo antenal de la mosca para proporcionar una discriminación proporcionada con el gran número de receptores.

Este modelo de codificación olfativa contrasta enormemente con el sistema olfativo principal de vertebrados en los que las neuronas sensoriales expresan sólo un único receptor y convergen sólo en un único par de glomérulos espacialmente fijos en el bulbo olfatorio. Además, cada neurona de proyección en el bulbo de mamífero extiende sus dendritas hacia sólo un único glomérulo. Por lo tanto, la integración y decodificación de patrones espaciales de actividad glomerular, en vertebrados, debe producirse principalmente en la corteza olfativa. En la mosca de la fruta, la observación de que el número de receptores puede superar el número de glomérulos sugiere que los glomérulos individuales recibirán entradas de más de un tipo de neurona sensorial. Se proporciona un segundo nivel de integración en el lóbulo antenal por subconjuntos de neuronas de proyección que elaboran amplios árboles dendríticos que forman sinapsis con múltiples glomérulos. Por lo tanto, el sistema olfativo de *Drosophila* revela niveles de procesamiento e integración de entrada sensorial en el lóbulo antenal que probablemente están limitados a centros corticales superiores en el sistema olfativo principal de vertebrados.

Secuencias de proteína y de ácido nucleico de 55 genes de receptores de odorizantes de Drosophila

Lo siguiente incluye los genes identificados por primera vez en 1998-1999. Las secuencias proteicas usaron códigos de aminoácidos de una sola letra.

DOR10

```
MEKLRSYEDFI FMANMMFKTLGYDLFHTPKPWRYLLVRGYFVLCTI SNFYEASMTT
RIIEWESLAGSPSKIMRQGLHFFYMLSSQLKFITFMINRKRLQLSHRLKELYPHKEQ
NQRKYEVNKYLLSCSTRNVLYVYFVMVMALEPLVQSQFIVNVSLGTDLWMMCVSSQ
ISMHLGYLANMLASIRPSPETEQQDCDFLASIIKRHQLMIRLQKDVNYVFGLLLASNL
FTTSCLLCCMAYYTVVEGFNWEGISYMLFASVAAQFYVSSHGQMLIDLMTITYRF
FAVIRQTVEK
```

DOR10nt

```
ATGGAAAACTACGTTCCCTATGAGGATTTTCATCTTCATGGCCAACATGATGTTCAAGA
CCCTTGGCTACGATCTATTCCATACACCCAAACCCTGGTGGCGCTATCTGCTTGTGCG
AGGATACTTCGTTTTGTGCACGATCAGCAACTTTTACGAGGCTTCCATGGTGACGACA
AGGATAATTGAGTGGGAATCCTTGGCCGGAAGTCCCTCCAAAATAATGCGACAGGGTC
TGCACTTCTTTTACATGTTGAGTAGCCAATTGAAATTTATCACATTCATGATAAATCG
CAAACGCCTACTGCAGCTGAGCCATCGTTTGAAAGAGTTGTATCCTCATAAAGAGCAA
AATCAAAGGAAGTACGAGGTGAATAAAATACTACCTATCCTGTTCCACGCGCAATGTTT
TGACGTGTACTACTTTGTAATGGTCGTCATGGCACTGGAACCCCTCGTTTCAGTCCCA
GTTCATAGTGAATGTGAGCCTGGGCACAGATCTGTGGATGATGTGCGTCTCAAGCCAA
ATATCGATGCACTTGGGCTATCTGGCCAATATGTTGGCCTCCATTGACCAAGTCCAG
AAACGGAACAACAAGACTGTGACTTCTTGGCCAGCATTATAAAGAGACATCAACTAAT
GATCAGGCTTCAAAAGGACGTGAACTATGTTTTGGACTCTTATTGGCATCTAATCTG
TTTACCACATCCTGTTTACTTTGCTGCATGGCGTACTATACCGTCGTCGAAGGTTTCA
ATTGGGAGGGCATTTCCTATATGATGCTCTTTGCTAGTGTAGCTGCCAGTTCTACGT
TGTCAGCTCACACGGACAAATGTTAATAGATTTGTTGATGACCATCACATACAGATTT
TTCGCGGTTATACGACAACTGTAGAAAAG
```

DOR104

MASLQFHGNVDADIRYDISLDPARESNLFRLLMGLQLANGTKPSRPLPKWWPKRLEMI
 GKVLPKAYCSMVIFTSLHLGVLF TKTTLDVLP TGELQAITDALMTIIYFFTGYGTIY
 WCLRSRRLAYMEHMNREYRHSLAGVTFVSSHAAFRMSRNFTVWIMSCLLGVISWG
 VSPLMLGIRMLPLQCWYFPFDALGPGTYTAVYATQLFGQIMVGMTFGFGGSLFVTLSELL
 LLGQFDVLYCSLKNLDAHTKLLGGESVNLSSLQEELLLGDSKRELNQYVLLQEHPTD
 LLRLSAGRKCPDQGNAFHNALVECTRLHRFILHCSQELENLFSPLYCLVKSLOITFQLC
 LLVFVGVS GTREVL RIVNQLQYLGLTIFELLMFTYCGELLSRHSIRSGDAFWRGAWWK
 HAHFIRQDILIFLVNSRRVHV TAGKFYVMDVNRLRSVITQAFSFLTLLQKLAACKTE
 SEL

DOR104nt

GAATTCGGCACGAGCAGTCGATGGCCAGTCTTCAGTTCCACGGCAACGTCGATGCGGA
 CATCAGGTATGATATTAGCCTGGATCCGGCTAGGGAATCGAATCTCTCCGTCTGCTA
 ATGGGACTCCAGTTGGCGAATGGCACGAAGCCATCGCCGCGGTTACCCAAATGGTGGC
 CAAAGCGGCTGGAAATGATTGGTAAAGTGCTGCCCAAAGCCTATTGTTCCATGGTGAT
 TTTACCTCCCTGCATTTGGGTGTCCTGTTACGAAAACCACACTGGATGTCCTGCCG
 ACGGGGGAGCTGCAGGCCATAACGGATGCCCTCACCATGACCATAATATACTTTTTCA
 CGGGCTACGGCACCATCTACTGGTGCTGCGCTCCCGGCGCCTCTTGGCCTACATGGA
 GCACATGAACCGGGAGTATCGCCATCATTCGCTGGCCGGGGTGACCTTTGTGAGTAGC
 CATGCGGCCCTTTAGGATGTCCAGAACTTCACGGTGGTGTGGATAATGTCCTGCCTGC
 TGGGCGTGATTTCTGGGGCGTTTCGCCACTGATGCTGGGCATCCGGATGCTGCCGCT
 CCAATGTTGGTATCCCTTCGACGCCCTGGGTCCCGGCACATATACGGCGGTCTATGCT
 ACACAACTTTTCGGTCAGATCATGGTGGGCATGACCTTTGGATTCGGGGGATCACTGT
 TTGTACCCCTGAGCCTGCTACTCCTGGGACAATTCGATGTGCTCTACTGCAGCCTGAA
 GAACCTGGATGCCCATACCAAGTTGCTGGGCGGGAGTCTGTAAATGGCCTGAGTTCG
 CTGCAAGAGGAGTTGCTGCTGGGGGACTCGAAGAGGGAATTAAATCAGTACGTTTTGC
 TCCAGGAGCATCCGACGGATCTGCTGAGATTGTCCGCAGGACGAAAATGTCTTGACCA
 AGGAAATGCGTTTCACAACGCCTTGGTGGAATGCATTCGCTTGCATCGCTTCATTCTG
 CACTGCTCACAGGAGTTGGAGAATCTATTAGTCCATATTGTCTGGTCAAGTCACTGC
 AGATCACCTTTAGCTTTGCCTGCTGGTCTTTGTGGGCGTTTCGGGTACTCGAGAGGT
 CCTGCGGATTGTCAACCAGCTACAGTACTTGGGACTGACCATCTTCGAGCTCCTAATG
 TTCACCTATTGTGGCGAACTCCTCAGTCGGCATAGTATTTCGATCTGGCGACGCTTTT
 GGAGGGGTGCGTGGTGGAAGCACGCCCATTTTCATCCGCCAGGACATCCTCATCTTTCT
 GGTCAATAGTAGACGTGCAGTTCACGTGACTGCCGGCAAGTTTTATGTGATGGATGTG
 AATCGTCTAAGATCGGTTATAACGCAGGCGTTCAGCTTCTTGACTTTGCTGCAAAAGT
 TGGCTGCCAAGAAGACGGAATCGGAGCTCTAACTGGTACCACGCATCGATATTTATT
 TAGCGCATTAATAAAAAAGTCGAGTAAAGCAAAAAAAAAAAAAAAAAAAAAA

DOR105

MFEDIQLIYMNIKILRFWALLYDKNLRRYVCIGLASFHIFTQIVYMMSTNEGLTGIIR
 NSYMLVLWINTVLRAYLLLADHRYLALIQLTEAYYDLLNLNDSYISEILDQVNKVG
 KLMARGNLFFGMLTSMGFGLYPLSSSERVLFPFGSKIPLNEYESPYEMWYIFQMLIT
 PMGCCMYIPYTSLIVGLIMFGIVRCKALQHRRLRQVALKHPYGD RDPRELREEI IACIR
 YQQSII EYMDHINELTTMMFLFELMAFSALLCALLFMLIIVSGTSQLI IVCMYINMIL
 AQILALYWYANELREQNLAVATAAYETEWFTFDVPLRKNILFMMMRAQRPAAILLGNI
 RPITLELFQNLNNTTYTFFTVLKRVYG

DOR105nt

ATGTTTGAAGACATTCAGCTAATCTACATGAATATCAAGATATTGCGATTCTGGGCCC
 TGCTCTATGACAAAACTTGAGGCGTTATGTGTGCATTGGACTGGCCTCATTCCACAT
 CTTACCCCAAATCGTCTACATGATGAGTACCAATGAAGGACTAACCGGGATAATTCTG
 AACTCATATATGCTCGTCTTTGGATTAATACGGTGCTGCGAGCTTATCTCTTGCTGG
 CGGATCACGACAGATATTTGGCTTTGATCCAAAACTAACTGAGGCCTATTACGATTT
 ACTGAATCTGAACGATTCGTATATATCGGAAATATTGGACCAGGTGAACAAGGTGGGA
 AAGTTGATGGCTAGGGGCAATCTGTTCTTTGGCATGCTCACATCCATGGGATTCTGGTC
 TGTACCCATTGTCTCCAGCGAAAGAGTCCTGCCATTTGGCAGCAAAATTCCTGGTCT
 AAATGAGTACGAGAGTCCGTACTATGAGATGTGGTACATCTTTCAGATGCTCATCACC
 CCGATGGGCTGTTGCATGTACATTCCGTACACCAGTCTGATTGTGGGCTTGATAATGT
 TCGGCATTGTGAGGTGCAAGGCTTTGCAGCATCGCCTCCGCCAGGTGGCGCTTAAGCA
 TCCGTACGGAGATCGCGATCCCCGTGAACCTGAGGGAGGAGATCATAGCCTGCATACGT
 TACCAGCAGAGCATTATCGAGTACATGGATCACATAAACGAGCTGACCACCATGATGT
 TCCTATTGCAACTGATGGCCTTTTCGGCGCTGCTCTGTGCGCTGCTCTTTATGCTGAT
 TATCGTCAGCGGCACCAAGTCAGCTGATAATTGTTTGCATGTACATTAAACATGATTCTG
 GCCCCAAATACTGGCCCTCTATTGGTATGCAATGAGTTAAGGGAACAGAATCTGGCGG
 TGGCCACCGCAGCCTACGAAACGGAGTGTTTACCTTCGACGTTCCACTGCGCAAAAA
 CATCCTGTTTATGATGATGAGGGCACAGCGCCAGCTGCAATACTACTGGGCAATATA
 CGCCCCATCACTTTGGAAGTGTTCAAAACCTACTGAACACAACCTATACATTTTSTA
 CGGTTCTCAAGCGAGTCTACGGA

DOR107

MYPRFLSRNYPLAKHLFFVTRYSFGLLGLRFGKEQSWLHLLWLVFNFNLAHCCQAEF
 VFGWSHLRTSPVDAMDAFCPLACSFTTLFKLGWMWRRQEVADLMDRIRLLIGEQUEKR
 EDSRRKVAQRSYYLMVTRCGMLVFTLGSITTGAFVLRSLWEMWVRRHQEFKFDMPFRM
 LFHDFAHMPWFPVFYLYSTWSGQVTYAFAGTDGFFFGFTLYMAFLLQALRYDIQDA
 LKPIRDPSLRESKICQRLADIVDRHNEIEKIVKEFSGIMAAPTFFVHFVSASLVIATS
 VIDILLYSGYNIIRYVVYTFVSSAIFLYCYGGTEMSTESLSLGEAAAYSSAWYTDRE
 TRRRVFLIILRAQRPI TVRVPFFAPSLPVFTSVIKFTGSIVALAKTIL

DOR107nt

ES 2 315 227 T3

ATGTATCCGCGATTCCCTCAGCCGTAACATCCGCTGGCCAAGCATTTGTTCTTCGTCA
CCAGATACTCCTTTGGCCTGCTGGGCCTGAGATTTGGCAAAGAGCAATCGTGGCTTCA
5 CCTCTTGTGGCTGGTGTTC AATTTTCGTTAACCTGGCGCACTGCTGCCAGGCGGAGTTC
GTCTTCGGCTGGAGTCACTTGCGCACCAAGTCCCGTGGATGCCATGGACGCCTTTTGTC
CTCTGGCCTGCAGTTTCACCACGCTCTTCAAGCTGGGATGGATGTGGTGGCGTCGCCA
10 GGAAGTAGCTGATCTAATGGACCGCATCCGCTTGCTCATCGGGGAGCAGGAGAAGAGG
GAGGACTCCCGGAGAAAGGTGGCTCAAAGGAGCTACTATCTCATGGTCACCAGGTGCG
GTATGCTGGTCTTCACCCTGGGCAGCATTACCACTGGAGCCTTCGTTCTGCGTTCCCT
15 TTGGGAAATGTGGGTGCGTCGTCATCAGGAGTTCAAATTCGATATGCCCTTTCGCATG
CTGTTCCACGACTTTGCGCATCGCATGCCCTGGTTTCCAGTTTTCTATCTCTACTCCA
CATGGAGTGGCCAGGTCACTGTGTACGCCTTTGCTGGTACAGATGGTTTCTTCTTTGG
20 CTTTACCCTCTACATGGCCTTCTTGCTGCAGGCCTTAAGATACGATATCCAGGATGCC
CTCAAGCCAATAAGAGATCCCTCGCTTAGGGAATCCAAAATCTGCTGTCAGCGATTGG
CGGACATCGTGGATCGCCACAATGAGATAGAGAAGATAGTCAAGGAATTTTCTGGAAT
25 TATGGCTGCTCCAACTTTTGTTCACTTCGTATCAGCCAGCTTAGTGATAGCCACCAGC
GTCATTGATATACTATTGTATTCCGGCTATAACATCATCCGTTACGTGGTGTACACCT
TCACGGTTTTCTCGGCCATCTTCCTCTATTGCTACGGAGGCACAGAAATGTCAACTGA
30 GAGCCTTTCTTGGGAGAAGCAGCCTACAGCAGTGCCTGGTATACTTGGGATCGAGAG
ACCCGCAGGCGGGTCTTTCTCATTATCCTGCGTGCTCAACGACCCATTACGGTGAGGG
TGCCCTTTTTTGCACCATCGTTACCAGTCTTCACATCGGTCATCAAGTTTACAGGTTT
35 GATTGTGGCACTGGCTAAGACGATACTG

DOR108

MDKHKDRIESMRILQVMQLFGLWPWSLKSEEEWTF TG FVKRNYRFLHLHPITFTFIG
LMWLEAFISSNLEQAGQVLYMSITEMALVVKILSIWHYRTEAWRLMYELQHAPDYQLH
NQEEVDFWRREQRFFKWFFYIYILISLGVVYSGCTGVLFLLEGYELPFAYYVPFEWQNE
45 RRYWFAYGYDMAGMTLTCISNITLDTLGCYFLFHISLLYRLLGLRLRETKNMKNDTIF
GQQLRAIFIMHQRIRSLTLTCQRIVSPYILSQIILSALIICFSGYRLQHVGI RDNPGQ
FISMLQFVSV MILQIYLP CYYGNEITVYANQLTNEVYHTNWLECRPPIRKLLNAYMEH
50 LKKPVTIRAGNSFAVGLPIFVK TINNAYSFLALLLNVS N

DOR108nt

ES 2 315 227 T3

ATGGATAAACACAAGGATCGCATTGAATCCATGCGCCTAATTCTTCAGGTCATGCAAC
TATTTGGCCTCTGGCCGTGGTCCTTGAAATCGGAAGAGGAGTGGACTTTCACCGGTTT
5 TGTAAGCGCAACTATCGCTTCCTGCTCCATCTGCCCATTACCTTCACCTTTATTGGA
CTCATGTGGCTGGAGGCCTTCATCTCGAGCAATCTGGAGCAGGCTGGCCAGGTTCTGT
ACATGTCCATCACCGAGATGGCTTTGGTGGTGAAAATCCTGAGCATTGTCGCACTATCG
10 CACCGAAGCTTGGCGGCTGATGTACGAACTCCAACATGCTCCGGACTACCAACTCCAC
AACCAGGAGGAGGTAGACTTTTGGCGCCGGGAGCAACGATTCTTCAAGTGGTTCTTCT
ACATCTACATTCTGATTAGCTTGGGCGTGGTATATAGTGGCTGCACTGGAGTACTTTT
15 TCTGGAGGGCTACGAACTGCCCTTTGCCCTACTACGTGCCCTTCGAATGGCAGAACGAG
AGAAGGTACTGGTTTCGCCTATGGTTACGATATGGCGGGCATGACGCTGACCTGCATCT
CAAACATTACCCTGGACACCCTGGGTTGCTATTTCTGTTCATATCTCTCTTTTGTA
20 CCGACTGCTTGGTCTGCGATTGAGGGAAACGAAGAATATGAAGAATGATACCATTTTT
GGCCAGCAGTTGCGTGCCATCTTCATTATGCATCAGAGGATTAGAAGCCTAACCTGA
CCTGCCAGAGAATCGTATCTCCCTATATCCTATCTCAGATCATTTTGAGTGCCCTGAT
25 CATCTGCTTTAGTGGATACCGCTTGACGATGTGGGAATTCGCGATAATCCCGGCCAG
TTTATATCCATGTTGCAGTTTGTGAGTGTGATGATCCTGCAGATTTACTTGCCCTGCT
ACTATGGAAACGAGATAACCGTGTATGCCAATCAGCTGACCAACGAGGTTTACCATAC
30 CAATTGGCTGGAATGTGCGCCACCGATTGCAAAGTTACTCAATGCCTACATGGAGCAC
CTGAAGAAACCGGTGACCATCCGGGCTGGCAACTCCTTCGCCGTGGGACTACCAATTT
TTGTTAAGACCATCAACAACGCCTACAGTTTCTTGGCTTTATTACTAAATGTATCGAA
35 T

DOR109

MESTNRLSAIQTLQVIRWIGLLKWNENEGDGVLTWLKRIYPFVLHPLTFTYIALMW
YEAITSSDFEEAGQVLYMSITELALVTLLNIWYRRHEAASLIHELQHDPAFNLRNSE
EIKFWQQNQRFKRIFYWYIWGSLFVAVMGYISVFFQEDYELPFGYYVPFEWRTREY
45 FYAWGYNVVAMTLCCLSNILLDTLGCYFMFHIALSLFRLLGMRLEALKNAEERKARPEL
RRIFQLHTKVRRLTRECEVLVSPYVLSQVVFSAFIICFSAYRLVHMGFKQRPGLFVTT
VQFVAVMIVQIFLPCYYGNELTFHANALTNVFGTNWLEYSVGTRKLLNCYMEFLKRP
50 VKVRAGVFFEIGLPIFVKTIINNAYSFFALLLKISK

DOR109nt

ATGGAGTCTACAAATCGCCTAAGTGCCATCCAAACACTTTTAGTAATCCAACGTTGGA
TAGGACTTCTTAAATGGGAAAACGAGGGCGAGGATGGAGTATTAACCTGGCTAAAACG
60 AATATATCCTTTTGTACTGCACCTTCCACTGACCTTCACGTATATTGCCTTAATGTGG
TATGAAGCTATTACATCGTCAGATTTTGAGGAAGCTGGTCAAGTTCTGTACATGTCCA

ES 2 315 227 T3

TCACCGAACTGGCATTGGTCACTAACTGCTGAATATTTGGTATCGTCGTCATGAAGC
TGCTAGTCTAATCCACGAATTGCAACACGATCCCGCATTTAATCTGCGCAATTCGGAG
5 GAAATCAAATTCTGGCAGCAAAATCAGAGGAACTTTAAGAGAATATTTTACTGGTACA
TCTGGGGCAGCCTTTTCGTGGCTGTAATGGGTATATAAGCGTGTTCCTCAGGAGGA
TTACGAGCTGCCCTTTGGCTACTACGTGCCATTGAGTGGCGCACCAGGGAACGATAC
10 TTCTACGCTTGGGGCTATAATGTGGTGGCCATGACCCTGTGCTGTCTATCCAACATCC
TACTGGACACACTAGGCTGTTATTTTCATGTTCCACATCGCCTCGCTTTTCAGGCTTTT
GGGAATGCGACTGGAGGCCTTGAAAAATGCAGCCGAAGAGAAAGCCAGACCGGAGTTG
15 CGCCGCATTTTCCAACCTGCACACTAAAGTCCGCCGATTGACGAGGGAATGCGAAGTGT
TAGTTTCACCCTATGTTCTATCCCAAGTGGTCTTCAGTGCCTTCATCATCTGCTTCAG
TGCCTATCGACTGGTGCACATGGGCTTCAAGCAGCGACCTGGACTCTTCGTGACCACC
20 GTGCAATTCGTGGCCGTCATGATCGTCCAGATTTTCTTGCCCTGTTACTACGGCAATG
AGTTGACCTTTCATGCCAATGCACTCACTAATAGTGTCTTCGGTACCAATTGGCTGGA
GTACTCCGTGGGCACTCGCAAGCTGCTTAAGTGTACATGGAGTTCCTCAAGCGACCG
25 GTTAAAGTGCAGCTGGGGTGTCTTTGAAATAGGACTACCCATCTTTGTGAAGACCA
TCAACAATGCCTACAGTTTCTTCGCCCTGCTGCTAAAGATATCCAAG

DOR110

MLFNYLRKPNPTNLLTSPDSFRYFEYGMFCMGWHTPATHKIIYYITSCLIFAWCAVYL
PIGIIISFKTDINTFTPNELLTVMQLFFNSVGMFPKVLFFNLYISGFYKAKKLLSEMD
35 KRCTTLKERVEVHQGVVRCNKAYLIYQFIYTAYTISTFLSAALSGKLPWRIYNPFVDF
RESRSSFWKAALNETALMLFAVTQTLMSDIYPLLYGLILRVHLKLLRLRVESLCTDSG
KSDAENEQDLINYAAAIRPAVTRTIFVQFLLIGICLGLSMINLLFFADIWTGLATVAY
40 INGLMVQTFPFCFVCDLLKKDCELLVSAIFHSNWINSRSYKSSLRYFLKNAQKSIAP
TAGSIFPISTGSNIKVAKLAFSVVTFVNQLNIADRLTKN

DOR110nt

ATGTTGTTCAACTATCTGCGAAAGCCGAATCCCACAAACCTTTTGACTTCTCCGGACT
CATTTAGATACTTTGAGTATGGAATGTTTTGCATGGGATGGCACACACCAGCAACGCA
50 TAAGATAATCTACTATATAACATCCTGTTTGATTTTGGCTGGTGTGCCGTATACTTG
CCAATCGGAATCATCATTAGTTTCAAACCGGATATTAACACATTACACCCGAATGAAC
TGTTGACAGTTATGCAATTATTTTCAATTCAGTGGGAATGCCATTCAAGGTTCTGTT
55 CTTCAATTTGTATATTTCTGGATTTTACAAGGCCAAAAGCTCCTTAGCGAAATGGAC
AAACGTTGCACCACTTTGAAGGAGCGAGTGAAGTGCACCAAGGTGTGGTCCGTTGCA
ACAAGGCCTACCTCATTTACCAGTTCATTTATACCGCGTACACTATTTCAACATTTCT
ATCGGCGGCTCTTAGTGGAATTTGCCATGGCGCATCTATAATCCTTTTGTGGATTTT
60 CGAGAAAGTAGATCCAGTTTTTGGAAAGCTGCCCTCAACGAGACAGCACTTATGCTAT
TTGCTGTGACTCAAACCTAATGAGTGATATATCCACTGCTTTATGGTTTGATCCT
GAGAGTTCACCTCAAACCTTTTGGGACTAAGAGTGGAGAGCCTGTGCACAGATTCTGGA
65

ES 2 315 227 T3

AAAAGCGATGCTGAAAACGAGCAAGATTTGATTAAGTATGCTGCAGCAATACGACCAG
CGGTTACCCGCACAATTTTCGTTCAATTCCTCTTGATCGGAATTTGCCTTGGCCTTTT
5 AATGATCAATCTACTCTTCTTTGCCGACATCTGGACAGGATTGGCCACAGTGGCTTAC
ATCAATGGTCTAATGGTGCAGACATTTCCATTTTGCTTCGTTTGTGATCTACTCAAAA
AGGATTGTGAACTTCTTGTGTGCGGCCATATTTCAATCCAAGTGGATTAATTCAGCCG
10 CAGTTACAAGTCATCTTTGAGATATTTTCTGAAGAAGCCCGAGAAATCAATTGCTTTT
ACAGCCGGCTCTATTTTTCCCATTTCTACTGGCTCGAATATTAAGGTGGCTAAGCTGG
CATTTCGGTGGTTACTTTTGTCAATCAACTTAACATAGCTGACAGATTGACAAAGAA
15 C

DOR111

MLFRKRKPKSDDEVI TFDELTRFPMTFYKTIGEDLYSDRDPNVIRRYLLRFYLVLGFL
NFNAYVVGIEIAYFIVHIMSTTTTLEATAVAPCIGFSFMADFKQFGLTVNRKRLVRLLD
DLKEIFPLDLEAQRKYNVSFYRKHMNRVMTLFTILCMYTSFSFYPAIKSTIKYYLM
25 GSEIFERNYGFHILFPYDAETDLTVYWFWSYGLAHCAVAGVSYVCVDLLLIATITQL
TMHFNFIANDLEAYEGGDHTDEENIKYLNHLVVYHARALDINKKCTFQSSRIGHSAFN
QNWLPSTKYKRILQFI IARSQKPASIRPPTFPPI SFNTFMKVISMYSYQFFALLRTTY
30 YG

DOR111nt

ATGCTGTTCCGCAAACGTAAGCCAAAAAGTGACGATGAAGTCATCACCTTCGACGAAC
TTACCCGGTTTCCGATGACTTTCTACAAGACCATCGGCGAGGATCTGTACTCCGATAG
GGATCCGAATGTGATAAGGCGTTACCTGCTACGTTTTATCTGGTACTCGGTTTTCTC
40 AACTTCAATGCCTATGTGGTGGGCGAAATCGCGTACTTTATAGTCCATATAATGTGCA
CGACTACTCTTTTGGAGGCCACTGCAGTGGCACCGTGCATTGGCTTCAGCTTCATGGC
CGACTTTAAGCAGTTCGGTCTCACAGTGAATAGAAAGCGATTGGTCAGATTGCTGGAT
45 GATCTCAAGGAGATATTTCCCTTTAGATTTAGAAGCGCAGCGGAAGTATAACGTATCGT
TTTACCGGAAACACATGAACAGGGTCATGACCCTATTCACCATCCTCTGCATGACCTA
CACCTCGTCATTTAGCTTTTATCCAGCCATCAAGTCGACCATAAAGTATTACCTTATG
GGATCGGAAATCTTTGAGCGCAACTACGGATTTACATTTTGTTCCTACGACGCAG
50 AAACGGATCTGACGGTCTACTGGTTTTCTACTGGGGATTGGCTCATTGTGCCTATGT
GGCCGGAGTTTCTACGTCTGCGTGGATCTCCTGCTGATCGCGACCATAACCCAGCTG
ACCATGCACTTCAACTTTATAGCGAATGATTTGGAGGCCTACGAAGGAGGTGATCATA
55 CGGATGAAGAAAATATCAAATACCTGCACAACTTGGTCGTCTATCATGCCAGGGCGCT
GGATATTAACAAGAAATGTACATTTAGAGCTCTCGGATTGGCCATTTCGGCATTTAAT
CAGAACTGGTTGCCATGCAGCACCAAATACAAACGCATCCTGCAATTTATTATCGCGC
60 GCAGCCAGAAGCCCGCCTCTATAAGACCGCCTACCTTTCCACCCATATCTTTTAATAC
CTTTATGAAGTAATCAGCATGTCGTATCAGTTTTTTGCACTGCTCCGCACCACATAT
TATGGT
65

DOR114

ES 2 315 227 T3

MLTKKDTQSAKEQEKLKAIPLHSFLKYANVFYLSIGMMAYDHKYSQKWKEVLLHWTFI
AQMVLNNTVLISELIYVFLAIGKGSNFLEATMNLFIGFVIVGDFKIWNISRQRKRLT
5 QVVSRLLEELHPQGLAQQEPYNIGHHLSGYSRYSKFYFGMHMVLIWTYNLYWAVYYLVC
DFWLGMQRQFERMLPYWCWVPWDWSTGYSYFMYISQNI GGQACLSGQLAADMLMCALV
TLVVMHFIRLSAHIESHVAGIGSFQHDLEFLQATVAYHQSLIHLCDINEIFGVSLLS
10 NFVSSSFIIICFVGFMQMTIGSKIDNLVMLVLF LFCAMVQVFMIAHAQRLVDASEQIGQ
AVYNHDWFRADLRYRKMLILIIKRAQQPSRLKATMFLNISLVTVSDLLQLSYKFFALL
RTMYVN

DOR114nt

ATGTTGACTAAGAAGGATACTCAAAGTGCCAAGGAGCAGGAAAAGTTGAAGGCCATTC
20 CATTGCACAGCTTTCTGAAATATGCCAACGTGTTCTATTTATCGATTGGAATGATGGC
CTACGATCACAAGTACAGTCAAAGTGGAAGGAGGTCCTGCTGCACTGGACATTCATT
GCCCAGATGGTCAATCTGAATACAGTGCTCATCTCGGAAGTGAATTTACGTATTCCTGG
25 CGATCGGCAAAGGTAGCAATTTTCTGGAGGCCACCATGAATCTGTCTTTCATTGGATT
TGTCATCGTTGGTGACTTCAAATCTGGAACATTTTCGCGGCAGAGAAAAGAGACTCACC
CAAGTGCTCAGCCGATTGGAAGAACTGCATCCGCAAGGCTTGGCTCAACAAGAACCCT
30 ATAATATAGGGCATCATCTGAGCGGCTATAGCCGATATAGCAAATTTTACTTCGGCAT
GCACATGGTGCTGATATGGACGTACAACCTGTATTGGGCGGTTTACTATCTGGTCTGT
GATTTCTGGCTGGGAATGCGTCAATTTGAGAGGATGCTGCCCTACTACTGCTGGGTTC
35 CCTGGGATTGGAGTACCGGATATAGCTACTATTTTCATGTATATCTCACAGAATATCGG
CGGTACAGGCTTGTCTGTCCGGTCAGCTAGCAGCTGACATGTTAATGTGCGCCCTGGTC
ACTTTGGTGGTGATGCACTTCATCCGGCTTTCCGCTCACATCGAGAGTCATGTTGCGG
40 GCATTGGCTCATTCCAGCACGATTTGGAGTTCCTCCAAGCGACGGTGGCGTATCACCA
GAGCTTGATCCACCTCTGCCAGGATATCAATGAGATATTCGGTGTTTCACTGTTGTCC
AACTTTGTATCCTCGTCGTTTATCATCTGCTTCGTGGGTTTCCAGATGACCATCGGCA
45 GCAAGATCGACAACCTGGTAATGCTTGTGCTTTTCTGTTTTGTGCCATGGTTCAGGT
CTTCATGATTGCCACCCATGCTCAGAGGCTCGTTGATGCGAGTGAACAGATTGGTCAA
GCGGTCTATAATCACGACTGGTTCGGTGCTGATCTGCGGTATCGTAAAATGCTGATCC
50 TGATTATTAAGAGGGGCCCAACAGCCGAGTCAAGGCCACAATGTTCTGAAACAT
CTCACTGGTCACCGTGTCCGATCTCTTGCAACTCTCGTACAAATCTTTGCCCTTCTG
CGCACAAATGTACGTGAAT

DOR115

ES 2 315 227 T3

MEKLMKYASFFYTAVGIRPYTNGEESKMKNKLIFHIVFWSNVINLSFVGLFESIYVYSA
FMDNKFLEAVTALS YIGFVTVGMSKMFFIRWKKTAITELINELKEIYPNGLIREERYN
5 LPMYLGTC SRISLIYSLLYSVLIWTFNLFVMEYWVYDKWLNIRVVVKQLPYLMYIPW
KWQDNWSYYP LLFSQNFAGY TSAAGQISTDVLLCAVATQLVMHFDFLSNSMERHEL SG
DWKKDSRFLVDIVRYHERILRLSDAVNDIFGIPLLLNFMVSSFVICFVGFQMTVGVPP
10 DIVVKLFLFLVSSMSQVYLICHYQQLVADASYGFSVATYNQKWYKADVRYKRALV I I
ARSQKVTF LKATIFLDITRSTMTDVRNCVLSV

DOR115nt

ATGGAGAAGCTAATGAAGTACGCTAGCTTCTTCTACACAGCAGTGGGCATACGGCCAT
20 ATACCAATGGTGAAGAATCCAAAATGAACAACTTATATTTACATAGTTTTTTGGTC
CAATGTGATTAACCTCAGCTTCGTTGGATTATTTGAGAGCATTACGTTTACAGTGCC
TTCATGGATAATAAGTTCCTGGAAGCAGTCACTGCGTTGTCCTACATTGGCTTCGTAA
CCGTAGGCATGAGCAAGATGTTCTTCATCCGGTGAAGAAAACGGCTATAACTGAACT
25 GATTAATGAATTGAAGGAGATCTATCCGAATGGTTTGATCCGAGAGGAAAGATACAAT
CTGCCGATGTATCTGGGCACCTGCTCCAGAATCAGCCTTATATATTCCTTGCTCTACT
CTGTTCTCATCTGGACATTCAACTTGTTTTGTGTAATGGAGTATTGGGTCTATGACAA
30 GTGGCTCAACATTGAGTGGTGGGCAAACAGTTGCCGTACCTCATGTACATTCTTG
AAATGGCAGGATAACTGGTCGTA CTATCCACTGTTATTCTCCAGAATTTTGCAGGAT
ACACATCTGCAGCTGGTCAAATTTCAACCGATGTCTTGCTCTGCGCGGTGGCCACTCA
35 GTTGGAATGCACTTCGACTTTCTCTCAAATAGTATGGAACGCCACGAATTGAGTGGA
GATTGGAAGAAGGACTCCCGATTTCTGGTGGACATTGTTAGGTATCACGAACGTATAC
TCCGCCTTTTCAGATGCAGTGAACGATATATTTGGAATTCCACTACTACTCAACTTCAT
40 GGTATCCTCGTTCGTCATCTGCTTCGTGGGATTCCAGATGACTGTTGGAGTTCCGCCG
GATATAGTTGTGAAGCTCTTCCTCTTCCTTGCTCTCTTCGATGAGTCAGGTCTATTTGA
TTTGTCACTATGGTCAACTGGTGGCCGATGCTAGCTACGGATTTTCGGTTGCCACCTA
45 CAATCAGAAGTGGTATAAAGCCGATGTGCGCTATAAACGAGCCTTGTTATTATTATA
GCTAGATCGCAGAAGGTAAC TTTCTAAAGGCCACTATATTCTTGATATTACCAGGT
CCACTATGACAGATGTACGCAACTGTGTATTGT CAGTG

DOR116

MELLPLAMLMYDGTRVTAMQYLI PGLPLENNYCYVVTYMIQVTMLVQGVGFYSGDLF
55 VFLGLTQILTFADMLQVKVKELNDALEQKAEYRALVRVGASIDGAENRQRLLLDVIRW
HQLFTDYCRAINALYELIATQVLSMALAMMLSFCINLSSFHMPSAIFFVVSAYSMSI
YCILGTILEFAYDQVYESICNVTWYELSGEQRKLFGLLRESQYPHNIQILGVMSLSV
60 RTALQIVKLIYSVSMNNRA

DOR116nt

5 ATGGAACCTCCTGCCATTGGCCATGCTAATGTACGATGGAACCCGGGTACTGCGATGC
AGTATTTAATTCGGGTCTACCGCTTGAGAACAATTATTGCTACGTAGTCACGTACAT
GATTCAGACGGTGACAATGCTCGTGCAAGGAGTCGGATTCTACTCCGGTGATTTGTTT
10 GTATTTCTCGGCTTAACGCAGATCCTAACTTTGCGCGATATGCTGCAGGTGAAGGTGA
AAGAGCTAAACGATGCCCTGGAACAAAAGCGGAATACAGAGCTCTAGTCCGAGTTGG
AGCTTCTATTGATGGAGCGGAAAATCGTCAACGCCTTCTCTTGATGTTATAAGATGG
CATCAATTATTCACGGACTACTGTGCGGCCATAAATGCCCTCTACTACGAATTGATCG
15 CCACTCAGGTTCTTTGATGGCTTTGGCCATGATGCTCAGCTTCTGCATTAATTTGAG
CAGCTTTCACATGCCTTCGGCTATCTTTTTCTGCGTTTCTGCCTACAGCATGTCCATC
TATTGCATTCTGGGCACCATTTCTTGAGTTTGCATATGACCAGGTGTACGAGAGCATCT
20 GTAATGTGACCTGGTATGAGTTGAGTGGCGAACAGCGAAAGCTTTTTGGTTTTTTGTT
GCGGGAATCCCAGTATCCGCACAATATTCAGATACTTGGAGTTATGTCGCTTTCCGTG
AGAACGGCTCTGCAGATTGTTAACTAATTTATAGCGTATCCATGATGATGAATC
25 GGGCG

DOR117

30 MDLRRWFPTLYTQSKDSPVRSRDATLYLLRCVFLMGVRKPPAKFFVAYVLWSFALNFC
STFYQPIGFLTGYISHLSEFSPGEFLTSLQVAFNAWSCSTKVLIVWALVKRFDEANNL
35 LDEMRRITDPGERLQIHRAVSLSNRIFFFFMAVYMYATNTFLSAIFIGRPPYQNY
PFLDWRSSTLHLALQAGLEYFAMAGACFQDVCVDCYPVNFVLVLAHMSIFAERLRL
GTYPYESQEKYERLVQCIQDHKVLRFVDCLRPVISGTIFVQFLVVLGFTLINI
40 VLFANLGSIAALSFMAAVLLETPFCILCNYLTEDCYKLADALFQSNWIDEEKRYQK
TLMYFLQKLQPPITFMAMNVFPIISVGTNISVSRCL

DOR117nt

45 ATGGATCTGCGAAGGTGGTTTTCCGACCTTGACACCCAGTCGAAGGATTCGCCAGTTC
GCTCCCGAGACGCGACCCTGTACCTCCTACGCTGCGTCTTCTTAATGGGCGTCCGAA
50 GCCACCTGCCAAGTTTTTCGTGGCCTACGTGCTCTGGTCCTTCGCACTGAATTTCTGC
TCAACATTTTATCAGCCAATTGGCTTTCTCAGAGGTATATAAGCCATTTATCAGAGT
TCTCCCCGGGAGAGTTTCTAACTTCGCTGCAGGTGGCCTTTAATGCTTGGTCCTGCTC
55 TACAAAAGTCCTGATAGTGTGGGCACTAGTTAAGCGCTTTGACGAGGCTAATAACCTT
CTCGACGAGATGGATAGGCGTATCACAGACCCCGAGAGCGTCTTCAGATTCATCGCG
CTGTCTCCCTCAGTAACCGTATATTCTTTTTCATGGCAGTCTACATGGTTTATGC
60 CACTAATACGTTTCTGTGCGCGATCTTCATTGGAAGGCCACCGTACCAAATTAATAC
CCTTTTCTGGACTGGCGATCTAGCACTCTGCATCTAGCTCTGCAGGCCGGTCTGGAAT

65

ES 2 315 227 T3

ACTTCGCCATGGCTGGCGCCTGCTTCCAGGACGTTTGCGTTGATTGCTACCCAGTCAA
TTTCGTTTTTGGTCCTGCGTGCCACATGTCGATCTTCGCGGAGCGCCTTCGACGTTTG
5 GGAAC TTATCCTTATGAAAGCCAGGAGCAGAAATATGAACGATTGGTTCA GTGCATAC
AAGATCACAAAGTAATTTT GCGATTTGTTGACTGCCTGCGTCCTGTTATTTCTGGTAC
CATCTTCGTGCAATTCTTGGTTGTGGGGTTGGTGCTGGGCTTTACCC TAATTAACATT
10 GTCCTGTTCGCCAACTTGGGATCGGCCATCGCAGCGCTCTCGTTTATGGCCGCAGTGC
TTCTAGAGACGACTCCCTTCTGCATATTGTGCAATTATCTCACAGAAGACTGCTACAA
GCTGGCCGATGCCCTGTTTCAGTCAAAC TGGATTGATGAGGAGAAACGATACCAAAG
15 AACTCATGTACTTCCTACAGAAACTGCAGCAGCCTATAACCTTCATGGCTATGAACG
TGTTTCCAATATCTGTGGGAACTAACATCAGTGTAAGCAGATGTGCCCTT

DOR118

MKFIGWLPPKQGVLRVYVYLTWTLMTFVWCTTYLPLGFLGSYMTQIKSFSPGEFLTSLQ
VCINAYGSSVKVAITYSMLWRLIKAKNILDQLDLRCTAMEEREKIHLVVARSNHAF LI
25 FTFVYCGYAGSTYLSSVLSGRPPWQLYNPFIDWHDGTLKLWVASTLEYMVMMSGAVLQD
QLSDSYPLIYTLILRAHLDMLRERIRRLRSDENLSEAESYEELVKCVMDHKLILRYCA
IIKPVIQGTIFTQFL LIGLVLGFTLINVFFFSDIWTG IASFMFVITILLQTFPFCYTC
30 NLIMEDCESLTHAIFQSNWVDASRRYKTLLYFLQNVQQPIVFIAGGIFQISMSSNIS
VAKFAFSVITITKQMNIADKFKTD

DOR118nt

ES 2 315 227 T3

ATGAAGTTTATTGGATGGCTGCCCCCAAGCAGGGTGTGCTCCGGTATGTGTACCTCA
CCTGGACGCTAATGACGTTTCGTGTGGTGTACAACGTACCTGCCGCTTGGCTTCCTGG
5 TAGCTACATGACGCAGATCAAGTCCTTCTCCCCCTGGAGAGTTTCTCACTTCACTCCAG
GTGTGCATTAATGCCTACGGCTCATCGGTAAAAGTTGCAATCACATACTCCATGCTCT
GGCGCCTTATCAAGGCCAAGAACATTTTGGACCAGCTGGACCTGCGCTGCACCGCCAT
10 GGAGGAGCGCGAAAAGATCCACCTAGTGGTGGCCCGCAGCAACCATGCCTTTCTCATC
TTCACCTTTGTCTACTGCGGATATGCCGGCTCCACCTACCTGAGCTCGGTTCTCAGCG
GGCGTCCGCCCTGGCAGCTGTACAATCCCTTTATTGATTGGCATGACGGCACACTCAA
15 GCTCTGGGTGGCCTCCACGTTGGAGTACATGGTGTATGTCAGGCGCCGTTCTGCAGGAT
CAACTCTCGGACTCTTACCCATTGATCTATACCCTCATCCTTCGTGCTCACTTGGACA
TGCTAAGGGAGCGCATCCGACGCCTCCGTTCCGATGAGAACCTGAGCGAGGCCGAGAG
20 CTATGAAGAGCTGGTCAAATGTGTGATGGACCACAAGCTCATTCTAAGATACTGCGCG
ATTATTAAACCAGTAATCCAGGGGACCATCTTCACACAGTTTCTGCTGATCGGCCTGG
TTCTGGGCTTCACGCTGATCAACGTGTTTTTCTTCTCAGACATCTGGACGGGCATCGC
25 ATCATTATGTTTGTATAACCATTTTGCTGCAGACCTTCCCCTTCTGCTACACATGC
AACCTCATCATGGAGGACTGCGAGTCCTTGACCCATGCTATTTTCCAGTCCAACCTGGG
TGGATGCCAGTCGTCGCTACAAAACAACACTACTGTATTTTCTCCAAAACGTGCAGCA
30 GCCTATCGTTTTTCATTGCAGGCGGTATCTTTCAGATATCCATGAGCAGCAACATAAGT
GTGGCAAAGTTTGCTTTCTCCGTGATAACCATTACCAAGCAAATGAATATAGCTGACA
AATTTAAGACGGAC

DOR119

MAVFKLIKPAPLTEKVQSRQGNIIYLYRAMWLI GWIPPKEGVLRVYVYLFWTCVPFAFGV
 FYLPVGFIIISYVQEFKNFTPGEFLTSLQVCINVGASVKSTITYLFLWRLRKTEILLD
 5 SLDKRLANDSDRERIHNMVARCNYAFLIYSFIYCGYAGSTFLSYALSGRPPWSVYNPF
 IDWRDGMGSLWIQAI FEYITMSFAVLQDQLSDTYPLMFTIMFRAHMEVLKDHVRSLRM
 DPERSEADNYQDLVNCVLDHKTILKCCDMIRPMISRTIFVQFALIGSVLGLTLVNVFF
 10 FSNFWKGVASLLFVITILLQTFPFCYTCNMLIDDAQDLSNEIFQSNWDAEPYKATL
 VLFMHVQQPIIFIAGGIFPISMNSNITVAKFAFSIITIVRQMNLAEQFQATGGCGGT
 GTTCAAGCTAATCAAACCGGCTCCGTTGACCGAGAAGGTGCAGTCCCGCCAGGGGAAT
 15 ATATATCTGTACCGTGCCATGTGGCTCATCGGATGGATTCCGCCGAAGGAGGGAGTCC
 TGGCGCTACGTGTATCTCTTCTGGACCTGCGTGCCCTTCGCCTTCGGGGTGTTTTACCT
 GCCCGTGGGCTTCATCATCAGCTACGTGCAGGAGTTCAAGAACTTCACGCCGGGCGAG
 TTCCTTACCTCGCTGCAGGTGTGCATCAATGTGTATGGCGCCTCGGTGAAGTCCACCA
 20 TCACCTACCTCTTCCTCTGGCGACTGCGCAAGACGGAGATCCTTCTGGACTCCCTGGA
 CAAGAGGCTGGCGAACGACAGCGATCGCGAGAGGATCCACAATATGGTGGCGCGCTGC
 AACTACGCCTTTCTCATCTACAGCTTCATCTACTGCGGATACGCGGGTTCCACTTTCC
 25 TGTCTTACGCCCTCAGTGGTCTCCTCCGTGGTCCGTCTACAATCCCTTCATCGATTG
 GCGCGATGGCATGGGCAGCCTGTGGATCCAGGCCATATTGAGTACATCACCATGTCC
 TTCGCCGTGCTGCAGGACCAGCTATCCGACACGTATCCCCTGATGTTCCACCATTATGT
 30 TCCGGGCCCCACATGGAGGTCTCAAGGATCACGTGCGGAGCCTGCGCATGGATCCCGA
 GCGCAGTGAGGCAGACAACTATCAGGATCTGGTGAAGTGGTGCTGGACCACAAGACT
 ATACTGAAATGCTGTGACATGATTGCCCCATGATATCCCGCACCATCTTCGTGCAAT
 35 TCGCGCTGATTGGTTCGGTTTTGGGCCCTGACCCTGGTGAACGTGTTCTTCTCTCGAA
 CTTCTGGAAGGGCGTGGCCTCGCTCCTGTTCGTATCACCATCCTGCTGCAGACCTTC
 CCGTTCTGCTACACCTGCAACATGCTGATCGACGATGCCAGGATCTGTCCAACGAGA
 40 TTTTCCAGTCCAACCTGGGTGGACGCGGAGCCGCGCTACAAGGCGACGCTGGTGCTCTT
 CATGCACCATGTTTCAGCAGCCCATAATCTTCATTGCCGGAGGCATCTTTCCCATCTCT
 ATGAACAGCAACATAACCGTGGCCAAGTTCGCCTTCAGCATCATTACAATAGTGGCGAC
 45 AAATGAATCTGGCCGAGCAGTTCCAG

DOR120

MTKFFFKRLQTAPLDQEVSSLDASDYYYRIAFFLGWTPPKGALLRWIYSLWTLTTMWL
 GIVYLPLGLSLTYVKHFDRFTPTEFLTSLQVDINCIGNVIKSCVTYSQMWRFRRMNEL
 55 ISSLDKRCVTTTQRRIFHKMVARVNLIVILFLSTYLGFCFLTFTSVFAGKAPWQLYN
 PLVDWRKGHWQLWIASILEYCVVVSIGTMQELMSDTYAIVFISLFRCHLAILRDRIANL
 RQDPKLSEMEHYEQMVACIQDHRTIIQCSQIIRPILSITIFAQFMLVGIDLGLAAISI
 60 LFFPNTIWTIMANVSFIVAICTESFPCCMLCEHLIEDSVHVSNALFHSNWITADRSYK
 SAVLYFLHRAQQPIQFTAGSTFPISVQSNIAVAKFAFTIITIVNQMNLGKFFSDRSN
 GDINP

DOR120nt

5 ATGACCAAGTTCTTCTTCAAGCGCCTGCAAAGTCTCCACTTGATCAGGAGGTGAGTT
 CCCTTGATGCCAGCGACTACTACTACCGCATCGCATTTCCTGGGCTGGACCCCGCC
 CAAGGGGGCTCTGCTCCGATGGATCTACTCCCTGTGGACTCTGACCACGATGTGGCTG
 10 GGTATCGTGACCTGCCGCTCGGACTGAGCCTCACCTATGTGAAGCACTTCGATAGAT
 TCACGCGGACGGAGTTCCTGACCTCCCTGCAGGTGGATATCAACTGCATCGGGAACGT
 GATCAAGTCATGCGTAACTTATTCCCAGATGTGGCGTTTTTCGCCGGATGAATGAGCTT
 ATCTCGTCCCTGGACAAGAGATGTGTGACTACGACACAGCGTCGAATTTTCATAAGA
 15 TGGTGGCACGGGTTAATCTCATCGTGATTCTGTTCTTGTCCACGTACTTGGGCTTCTG
 CTTTCTAACTCTGTTCACTTCGGTTTTTCGCTGGCAAAGCTCCTTGGCAGCTGTACAAC
 CCACTGGTGGACTGGCGGAAAGGCCATTGGCAGCTATGGATTGCCTCCATCCTGGAGT
 20 ACTGTGTGGTCTCCATTGGCACCATGCAGGAGTTGATGTCCGACACCTACGCCATAGT
 GTTCATCTCCTTGTTCCGCTGCCACCTGGCTATTCTCAGAGATCGCATAGCTAATCTG
 CGGCAGGATCCGAAACTCAGTGAGATGGAACACTATGAGCAGATGGTGGCCTGCATTC
 25 AGGATCATCGAACCATCATAAGTGCTCCCAGATTATTCGACCCATCCTGTGATCAC
 TATCTTTGCCCAGTTCATGCTGGTTGGCATTGACTTGGGTCTGGCGGCCATCAGCATC
 CTCTTCTTTCCGAACACCATTGTTGGACGATCATGGCAAACGTGTCGTTTCATCGTGGCCA
 30 TCTGTACAGAGTCCTTTCCATGCTGCATGCTCTGCGAGCATCTGATCGAGGACTCCGT
 CCATGTGAGCAACGCCCTGTTCCACTCAAAGTGGATAACCGCGGACAGGAGCTACAAG
 TCGGCGGTTCTGTATTTCTGACCGGGCTCAGCAACCCATTCAATTCACGGCCGGCT
 35 CCATATTTCCCATTTTCGGTGCAGAGCAACATAGCCGTGGCCAAGTTCGCGTTACAAT
 CATCACAATCGTGAACCAAATGAATCTGGGCGAGAAGTTCTTCAGTGACAGGAGCAAT
 GGCGATATAAATCCT

DOR121

45 MLTDKFLRLQSAFLRLGLELLHEQDVGHRYPWRSICCILSVASFMPPLTIAFGLQNVQ
 NVEQLTDSLCSVLVDLLALCKIGLFLWLYKDFKFLIGQFYCVLQTETHHTAEMIVTR
 ESRRDQFISAMYAYCFITAGLSACLMSPLSMLISYHEQVNCNRNHFVPVCKKKYCLIS
 50 RILRYSFCRYPWDNMKLSNYIISYFWNVCAALGVALPTVCVDTLFCSLSHNLALFQI
 ARHKMMHFEGRNTKETHENLKHVFQLYALCLNLGHFLNEYFRPLICQFVAASLHLCVL
 CYQLSANILQPALLFYAAFTAAVVGQVSIYCFGSSIHSEQLFGQAIYESSWPHLLQ
 55 ENLQLVSSLKIAMMRSSLGCPIDGYFFEANRETLITVSKAFIKVSKKTPQVND

DOR121

60

65

ES 2 315 227 T3

ATGCTGACGGACAAGTTCCTCCGACTGCAGTCCGCTTTATTTTCGCTTCTCGGACTCG
AATTGTTGCACGAGCAGGATGTTGGCCATCGATATCCTTGGCGCAGCATCTGCTGCAT
TCTCTCGGTGGCCAGTTTCATGCCCCTGACCATTGCGTTTGGCCTGCAAAACGTCCAA
AATGTGGAGCAATTAACCGACTCACTCTGCTCGGTTCTCGTGGATTGCTGGCCCTGT
GCAAAATCGGGCTTTTCCTTTGGCTTTACAAGGACTTCAAGTTCCTAATAGGGCAGTT
CTATTGTGTTTTTGCAAACGGAAACCCACACCGCTGTCGCTGAAATGATAGTGACCAGG
GAAAGTCGTCGGGATCAGTTCATCAGTGCTATGTATGCCTACTGTTTCATTACGGCTG
GCCTTTCGGCCTGCCTGATGTCCCCTCTATCCATGCTGATTAGCTACCACGAACAGGT
GAATTGCAGCCGAAATTTCCATTTCCCAGTGTGTAAGAAAAAGTACTGCTTAATATCC
AGAATATTAAGATACAGTTTCTGCAGATATCCCTGGGACAATATGAAGCTGTCCAAC
ACATCATTTTCTATTTCTGGAATGTGTGTGCTGCATTGGGCGTGGCACTGCCCACCGT
TTGTGTGGACACACTGTTCTGTTCTCTGAGCCATAATCTCTGTGCCCTATTCCAGATT
GCCAGGCACAAAATGATGCACTTTGAGGGCAGAAATACCAAAGAGACTCATGAGAACT
TAAAGCACGTGTTTTCAACTATATGCGTTGTGTTTGAACCTGGGCCATTTCTTAAACGA
ATATTTTCAGACCGCTCATCTGCCAGTTTGTGGCAGCCTCACTGCACTTGTGTGTCTG
TGCTACCAACTGTCTGCCAATATCCTGCAGCCAGCGTTACTCTTCTATGCCGCATTTA
CGGCAGCAGTTGTTGGCCAGGTGTCTATATACTGCTTCTGCGGATCGAGCATCCATTC
GGAGTGTGAGCTATTTGGCCAGGCCATCTACGAGTCCAGCTGGCCCCATCTGCTGCAG
GAAAACCTGCAGCTTGTAAGCTCCTTAAAAATTGCCATGATGCGATCGAGTTTGGGAT
GTCCCATCGATGGTTACTTCTTCGAGGCCAATCGGGAGACGCTCATCACGGTGAGTAA
AGCGTTTATAAAAGTGTCCAAAAGACACCTCAAGTGAATGAT

DOR14

MDYDRIRPVRFLTGVLEKWWRLWPRKESVSTPDWTNWQAYALHVPFTFLFVLLLWLEAI
KSRDIQHTADVLLICLTITLGGKVINIWKYAHVAQGILSEWSTWDLFELRSKQEVDM
WRFEHRRFNRFVFMFYCLCSAGVIPFIVIQPLFDIPNRLPFWMWTPFDWQQPVLFWYAF
IYQATTIPIACACNVMTDAVNWYMLHLSLCLRMLGQRLSKLQHDDKDLREKFLELIH
LHQRLLKQALSIEIFISKSTFTQILVSSLIICFTIYSMQMDLPGFAAMMQYLVAMIMQ
VMLPTIYGNAVIDSANMLTDSMYNSDWPDMNCRMRLVLMFMVYLNRPVTLKAGGFFH
IGLPLFTKVVFSTLENPCISYLYFRP

DOR14nt

ES 2 315 227 T3

ATGGA CTACGATCGAATTCGACCGGTGCGATTTTGTACGGGAGTGCTGAAATGGTGGC
GTCTCTGGCCGAGGAAGGAATCGGTGTCCACACCGGACTGGACTAACTGGCAGGCATA
5 TGCCTTGCACGTTCCATTTACATTCTTGTGTTGTGCTTTTGTGGTTGGAGGCAATC
AAGAGCAGGGATATACAGCATACCGCCGATGTCCTTTTGATTGTCCTAACCACCACTG
CCTTGGGAGGTAAAGTTATCAATATCTGGAAGTATGCCCATGTGGCCCAAGGCATTTT
10 GTCCGAGTGGAGCACGTGGGATCTTTTCGAGCTGAGGAGCAAACAGGAAGTGGATATG
TGGCGATTTCGAGCATCGACGTTTCAATCGTGTGTTTTATGTTTTACTGTTTGTGCAGTG
CTGGTGTAATCCCATTTATTGTGATTCAACCGTTGTTTGATATCCCAAATCGATTGCC
15 CTTCTGGATGTGGACACCATTTCGATTGGCAGCAGCCTGTTCTCTTCTGGTATGCATTC
ATCTATCAGGCCACAACCATTCCTATTGCCTGTGCTTGCAACGTAACCATGGACGCTG
TTAATTGGTACTTGATGCTGCATCTGTCCTTGTGTTTGCGTATGTTGGGCCAGCGATT
20 GAGTAAGCTTCAGCATGATGACAAGGATCTGAGGGAGAAGTTCCTGGAAGTATCCAT
CTGCACCAGCGACTCAAGCAACAGGCCTTGAGCATTGAAATCTTTATTTTGAAGAGCA
CGTTCACCCAAATTCTGGTCAGTTCCTTATCATTTGCTTCACCATTTACAGCATGCA
25 GATGGACTTGCCAGGATTTGCCGCCATGATGCAGTACCTAGTGGCCATGATCATGCAG
GTCATGCTGCCCACCATATATGGTAACGCCGTCATCGATTCTGCAAATATGTTGACCG
ATTCCATGTACAATTCGGATTGGCCGGATATGAATTGCCGAATGCGTCGCCTAGTTTT
30 AATGTTTATGGTGTACTTAAATCGACCGGTGACCTTAAAAGCCGGTGGCTTTTTTCAT
ATTGTTTTACCTCTGTTTACCAAGGTGTATTTTCTACTCTGGAAAATCCTTGTATAA
GTTATCTTTATTTTCAGACCA

DOR16

MTDSGQPAIADHFYRIPRISGLIVGLWPQIRIRGGGGRPWHAHLLFVFAFAMVVVGAVG
40 EVSYGCVHLDNLVVALEAFCPGTTKAVCVLKLWVFFRSNRRWAELVQRLRAILWESRR
QEAQRMVLGLATTANRLSLLLLSSGTATNAAFTLQPLIMGLYRWIVQLPGQTELPFNI
ILPSFAVQPGVFPLTYVLLTASGACTVFASFVDGFFICSCLYICGAFRLVQQDIRRI
45 FADLHGDSVDVFTEEMNAEVRHRLAQVVERHNAIIDFCTDLTRQFTVIVLMHFLSAAF
VLCSTILDIMLVSPFSEAFWGGYPWVCRATGFSHRLHSAAVLKVFPCFHCLLFFPGF
SSRSVLIRFSRFVCLLCGCGCGSLRWQFISA

DOR16nt

ATGACTGACAGCGGGCAGCCTGCCATTGCCGACCACTTTTATCGGATTCCCCGCATCT
55 CCGGCCTCATTGTCCGCCTCTGGCCGCAAAGGATAAGGGGCGGGGCGGTCTGTCCTTG
GCACGCCCATCTGCTCTTCGTGTTTCGCCTTCGCCATGGTGGTGGTGGGTGCGGTGGGC
60 GAGGTGTCGTACGGCTGTGTCCACCTGGACAACCTGGTGGTGGCGCTGGAGGCCTTCT
GCCCCGGAACCACCAAGGCGGTCTGCGTTTTGAAGCTGTGGGTCTTCTTCCGCTCCAA

ES 2 315 227 T3

TCGCCGGTGGGCGGAGTTGGTCCAGCGCCTGCGGGCTATTTTGTGGGAATCGCGGCGG
CAGGAGGCCCAGAGGATGCTGGTCCGACTGGCCACCACGGCCAACAGGCTCAGCCTGT
TGTGCTCAGCTCTGGCACGGCGACAAATGCCGCCTTCACCTTGCAACCGCTGATTAT
GGGTCTCTACCGCTGGATTGTGCAGCTGCCAGGTCAAACCGAGCTGCCCTTTAATATC
ATACTGCCCTCGTTTGCCGTGCAGCCAGGAGTCTTCCGCTCACCTACGTGCTGCTGA
CCGCTTCCGGTGCCTGCACCGTTTTCGCCTTCAGCTTCGTGGACGGATTCTTCATTTG
CTCGTGCCTCTACATCTGCGGCGCTTTCGGCTGGTGCAGCAGGACATTTCGAGGATA
TTTGCCGATTTGCATGGCGACTCAGTGGATGTGTTACCGAGGAGATGAACGCGGAGG
TGCGGCACAGACTGGCCCAAGTTGTGAGCGGCACAATGCGATTATCGATTTCTGCAC
GGACCTAACACGCCAGTTACCGTTATCGTTTAAATGCATTTCCCTGTCCGCCGCCTTC
GTCCTCTGCTCGACCATCCTGGACATCATGTTGGTGAGCCCCTTTTCAGAGGCCTTCC
TTTGGGGCGGGTATCCTTGGGTTTGTGCGGCCACTGGCTTTTCGCATCGCCTGCATTC
GGCGGCTGTTTTAAAGTTTTCCCTGTTTTCACTGTTTGCTGTTTTCCCTGGCTTT
TCCAGCCGCTCCGTTCTGATTCGGTTTTCCCGATTTGTTTGTTGCTTTGTGGCTGCG
GCTGCGGCTCTCTCCGGTGGCAATTTATAAGCGCATGA

DOR19

MVTEDFYKYQVWYFQILGVWQLPTWAADHQRRFQSMRFGFILVILFIMLLLSFEMLN
NISQVREILKVFFMFATEISCMAKLLHLKLSRKLGLVDAMLSPEFGVKSEQEQML
ELDRVAVVRMRNSYGIMSLGAASLILIVPCFDNFGELPLAMLEVCSIEGWICYWSQYL
FHSICLLPTCVLNITYDSVAYSLLCFLKVQLQMLVLRLEKLGPIEPQDNEKIAMELR
ECAAYNRIVRFKDLVELFIKPGSVQLMCSVLVLVSNLYDMSTMSIANGDAIFMLKT
CIYQLVMLWQIFIICYASNEVTVQSSRLCHSIYSSQWTGWNRRNRIVLLMMQRFNSP
MLLSTFNPTFAFSLEAFGSVGQKFLYISFITGYALLLSDRQLLLQLLRTAEARQQLN
FETPQHLKIFKPIFKSTQNVMHVH

DOR19nt

ATGGTTACGGAGGACTTTTATAAGTACCAGGTGTGGTACTTCCAAATCCTTGGTGTTT
GGCAGCTCCCCACTTGGGCCGAGACCACCAGCGTCGTTTTTCAGTCCATGAGGTTTGG
CTTCATCCTGGTCATCCTGTTTCATCATGCTGCTGCTTTTCTCCTTCGAAATGTTGAAC
AACATTTCCCAAGTTAGGGAGATCCTAAAGGTATTCTTCATGTTGCCACGGAAATAT
CCTGCATGGCCAAATTATTGCATTTGAAGTTGAAGAGCCGCAAACCTCGCTGGCTTGGT
TGATGCGATGTTGTCCCCAGAGTTCGGCGTTAAAAGTGAACAGGAAATGCAGATGCTG
GAATTGGATAGAGTGGCGGTTGTCCGCATGAGGAACTCCTACGGCATCATGTCCCTGG
GCGCGGCTTCCCTGATCCTTATAGTTCCCTGTTTCGACAACTTTGGCGAGCTACCACT
GGCCATGTTGGAGGTATGCAGCATCGAGGGATGGATCTGCTATTGGTGCAGTACCTT
TTCCACTCGATTTGCCTGCTGCCCACTTGTGTGCTGAATATAACCTACGACTCGGTGG
CCTACTCGTTGCTCTGTTTCTTGAAGGTTTCAGCTACAAATGCTGGTCTGCGATTAGA

AAAGTTGGGTCTGTGATCGAACCCAGGATAATGAGAAAATCGCAATGGAACCTGCGT
 GAGTGTGCCGCTACTACAACAGGATTGTTTCGTTTCAAGGACCTGGTGGAGCTGTTCA
 5 TAAAGGGGCCAGGATCTGTGCAGCTCATGTGTTCTGTTCTGGTGCTGGTGTCCAACCT
 GTACGACATGTCCACCATGTCCATTGCAAACGGCGATGCCATCTTTATGCTCAAGACC
 TGTATCTATCAGCTGGTGATGCTCTGGCAGATCTTCATCATTTGCTACGCCTCCAACG
 10 AGGTAACGTCCAGAGCTCTAGGTTGTGTACAGCATCTACAGCTCCCAATGGACGGG
 ATGGAACAGGGCAAACCGCCGATTGTCTTCTCATGATGCAGCGCTTTAATTCCTCCG
 ATGCTCCTGAGCACCTTTAACCCACCTTTGCTTTCAGCTTGGAGGCCTTTGGTTCTG
 15 TAGGGCAGCAGAAATTCCTTTATATATCATTTATTACTGGTTATGCTCTTCTCCTTTC
 AGATCGTCAACTGCTCCTACAGCTACTTCGCACTGCTGAAGCGCGTCAACAGTTAAAT
 TTCGAAACACCGCAGCACCTAAAGATTTTCAAGCCGATTTTTAAAAGCACTCAAAACG
 20 TTATGCACGTACAT

DOR20

MSKGVEIFYKGQKAFLNILSLWPQIERRWRIIHQVNYVHVIVFWVLLFDLLLVLHVMA
 NLSYMSEVVKAIFILATSAGHTTKLLSIKANNVQMEELFRRLDNEEFRPRGANEELIF
 AAACERSRKLRFYFALSFAALSMILIPQFALDWSHLPLKTYNPLGENTGSPAYWLLY
 30 CYOCLALSVSCITNIGFDSLCSLFIPLKQDLILAVRLDKIGRLITTSGGTVEQQLK
 ENIRYHMTIVELSKTVERLLCKPISVQIFCSVLVLTANFYAIAVVSCEFATRRLSVCD
 LSGVHVDSDFYIVLLCRVGIPYPKCLPRPVMNFIVSEVTQRSLDLPHELYKTSWVDWD
 35 YRSRRIALLFMQRLHSTLRIRTLNPSLGFDLMLFSSVSSFRVLTFLCTVANFHNEAH

DOR20nt

ATGAGCAAAGGAGTAGAAATCTTTTACAAGGGCCAGAAGGCATTCTTGAACATCCTCT
 CGTTGTGGCCTCAGATAGAACGCCGGTGGAGAATCATCCACCAGGTGAACTATGTCCA
 45 CGTAATTGTGTTTTGGGTGCTGCTCTTTGATCTCCTCTTGGTGCTCCATGTGATGGCT
 AATTTGAGCTACATGTCCGAGGTTGTGAAAGCCATCTTTATCCTGGCCACCAGTGCAG
 GGCACACCACCAAGCTGCTGTCCATAAAGGCGAACAATGTGCAGATGGAGGAGCTCTT
 TAGGAGATTGGATAACGAAGAGTTCGCTCCTAGAGGCGCCAACGAAGAGTTGATCTTT
 50 GCAGCAGCCTGTGAAAGAAGTAGGAAGCTTCGGGACTTCTATGGAGCGCTTTCGTTTG
 CCGCCTTGAGCATGATTCTCATACCCAGTTTCGCCTTGGACTGGTCCCACCTTCCGCT
 CAAAACATACAATCCGCTTGGCGAGAATACCGGCTCACCTGCTTATTGGCTCCTCTAC
 55 TGCTATCAGTGTCTGGCCTTGTCGTATCCTGCATCACCAACATAGGATTCGACTCAC
 TCTGCTCCTCACTGTTTCATCTTCCTCAAGTGCCAGCTGGACATTCTGGCCGTGCGACT
 GGACAAGATCGGTCGGTTAATCACTACTTCTGGTGGCACTGTGGAACAGCAACTTAAG
 60 GAAAATATCCGCTATCACATGACCATCGTTGAACTGTCGAAAACCGTGGAGCGTCTAC
 TTTGCAAGCCGATTTCCGGTGCAGATCTTCTGCTCGGTTTTGGTGCTGACTGCCAATTT
 CTATGCCATTGCTGTGGTGAGCTGTGAATTCGCAACAAGAAGACTATCAGTATGTGAC
 65 CTATCAGGCGTGCATGTTGATTCAGATTTTTATATTGTGCTACTATGCCGGGTGGGTA

ES 2 315 227 T3

TTCCATATCCGAAATGCCTCCCCAGGCCAGTAATGAATTTTCATCGTCAGTGAGGTAAC
CCAGCGCAGCCTGGACCTTCCGCACGAGCTGTACAAGACCTCCTGGGTGGACTGGGAC
5 TACAGGAGCCGAAGGATTGCGCTCCTCTTTATGCAACGCCTTCACTCGACCTTGAGGA
TTAGGACACTTAATCCAAGTCTTGGTTTTGACTTAATGCTCTTCAGCTCGGTGAGTTC
TTTCCGTGTTTTGACTTTTTTGTGCACTGTAGCCAATTTCCATAATGAGGCTCAT

DOR2 4

MDSFLQVQKSTIALLGFDLFSENREMWKRPYRAMNVFSIAAIFPFILAAVLHNWKNVL
LLADAMVALLITILGLFKFSMILYLRRDFKRLIDKFRLMSNEAEQGEYAEILNAAN
KQDQRMCTLFRTCFLLAWALNSVLPVLRMGLSYWLAGHAEPELPFPCLPWNIH I RN
15 YVLSFIWSAFASTGVVLPVAVSLDTIFCSFTSNLCAFFKIAQYKVVRFKGGSCLKESQAT
LNKVFALYQTSLDMCNDLNQCYQPIICAQFFISSLQLCMLGYLFSITFAQTEGVYYAS
20 FIATII IQAYIYCYCGENLKTESASFEWAIYDSPWHESLGAGGASTSICRSLI SMMR
AHRGFRITGYFFEANMEAFSSIVRTAMS YITMLRSFS

DOR24nt

GGCAGGAGCCTTGTCGACATGGACAGTTTTCTGCAAGTACAGAAGAGCACCATTGCTC
TTCTGGGCTTTGATCTCTTTAGTGAAAATCGAGAAATGTGGAAACGCCCTATAGAGC
AATGAATGTGTTTAGCATAGCTGCCATTTTTCCCTTTATCCTGGCAGCTGTGCTCCAT
35 AATTGGAAGAATGTATTGCTGCTGGCCGATGCCATGGTGGCCCTACTAATAACCATTCT
TGGGCCATTCAAGTTTAGCATGATACTTTACTTACGTCGCGATTTCAAGCGACTGAT
TGACAAATTTGTTTTGCTCATGTGCAATGAGGCGGAACAGGGCGAGGAATACGCCGAG
40 ATTCTCAACGCAGCAAACAAGCAGGATCAACGAATGTGCACTCTGTTTAGGACTTGTT
TCCTCCTCGCCTGGGCCTTGAATAGTGTTCTGCCCCCTCGTGAGAATGGGTCTCAGCTA
TTGGTTAGCAGGTCATGCAGAGCCCGAGTTGCCTTTTCCCTGTCTTTTCCCTGGAAT
45 ATCCACATCATTCGCAATTATGTTTTGAGCTTCATCTGGAGCGCTTTCGCCTCGACAG
GTGTGGTTTTACCTGCTGTCAGCTTGGATACCATATTCTGTTCCCTCACCAGCAACCT
GTGCGCCTTCTTCAAAATTGCGCAGTACAAGGTGGTTAGATTTAAGGGCGGATCCCTT
50 AAAGAATCACAGGCCACATTGAACAAAGTCTTTGCCCTGTACCAGACCAGCTTGATA
TGTGCAACGATCTGAATCAGTGCTACCAACCGATTATCTGCGCCCAGTTCTTCATTTCT
ATCTCTGCAACTCTGCATGCTGGGATATCTGTTCTCCATTACTTTTGCCAGACAGAG
55 GGCGTGACTATGCCTCTTTCATAGCCACCATCATTATACAAGCCTATATCTACTGCT
ACTGCGGGGAGAACCTGAAGACGGAGAGTGCCAGCTTCGAGTGGGCCATCTACGACAG
TCCGTGGCACGAGAGTTTGGGTGCTGGTGGAGCCTCTACCTCGATCTGCCGATCCTTG
60 CTGATCAGCATGATGCGGGCTCATCGGGGATTCCGCATTACGGGATACTTCTTCGAGG
CAAACATGGAGGCCTTCTCATCGATTGTTTCGCACGGCTATGTCCTACATCACAATGCT
GAGATCATCTCCTAAATGTGGTTTGACCACAAGGCTTTGGATTGATTTTTGTGCAAT

TTTTGTTTTATTGCTGAGCATGCGTTGCCGTACGACATTTAACAATCGATCTTACGTA
 ATTTACATATGATAATCTCACATATTGTTTCGTTAAGCACTAAGTAGAATGTAGAATGT
 GAATTGGCTGTAGAAATGCACAGATGAAGCACGAAAAAAAAAAAAAAAAAAAAAAAAA

DOR25

MNDSGYQSNLSLLRVFLDEFRSVLRQESPLIPRLAFYYVRAFLSLPLYRWINLFIMC
 NVMTIFWTFMVALPESKNVIEMGDDLWISGMALVFTKIFYMHLRCDEIDELISDFEY
 YNREL RPHNIDEEVLGWQRLCYVIESGLYINCFCLVNFFSAAIFLQPLLGEGLKLPFHS
 VYPFQWHRLDLHPYTFWFLYIWQSLTSQHNLM SILMVD MVGISTFLQTALNLKLLCIE
 IRKLGDM EVSDKRFHEEF CRVVRFHQHI I KLVGKANRAFNGAFNAQLMASFSLISIST
 FETMAAAVDPKMAAKFVLLMLVAFIQLSLWCVSGTLVYTQSVEVAQA AFDINDWHTK
 SPGIQRDISFVILRAQKPLMYVAEPFLPFTLGTYMLVLKNCYRLLALMQESM

DOR25nt

ATGAACGACTCGGGTTATCAATCAAATCTCAGCCTTCTGCGGGTTTTTCTCGACGAGT
 TCCGATCGGTTCTGCGGCAGGAAAGTCCCGGTCTCATCCACGCCTGGCTTTTTACTA
 TGTTGCGCCTTTCTGAGCTTGCCCTGTACCGATGGATCAACTTGTTTCATCATGTGC
 AATGTGATGACCATTTTCTGGACCATGTTTCGTGGCCCTGCCCGAGTCGAAGAACGTGA
 TCGAAATGGGCGACGACTTGTTTGGATTTCGGGGATGGCACTGGTGTTACCAAGAT
 CTTTTACATGCATTTGCGTTGCGACGAGATCGATGAACTTATTTGCGATTTTGAATAC
 TACAACCGGGAGCTGAGACCCATAATATCGATGAGGAGGTGTTGGGTTGGCAGAGAC
 TGTGCTACGTGATAGAATCGGGTCTATATATCAACTGCTTTGCTGGTCAACTTCTT
 CAGTGCCGCTATTTTCTGCAACCTCTGTTGGGCGAGGGAAAGCTGCCCTTCCACAGC
 GTCTATCCGTTTCAATGGCATCGCTTGGATCTGCATCCCTACACGTTCTGGTTCCTCT
 ACATCTGGCAGAGTCTGACCTCGCAGCACAACTAATGAGCATTCTAATGGTGGATAT
 GGTAGGCATTTCCACGTTCTCCAGACGGCGCTCAATCTCAAGTTGCTTTGCATCGAG
 ATAAGGAAACTGGGGGACATGGAGGTCAGTGATAAGAGGTCCACGAGGAGTTTGTGTC
 GTGTGGTTGCTTCCACCAGCACATTATCAAGTTGGTGGGAAAGCCAATAGAGCTTT
 CAATGGCGCCTTCAATGCACAATTAATGGCCAGTTTCTCCCTGATTTCCATATCCACT
 TTCGAGACCATGGCTGCAGCGGCTGTGGATCCCAAATGGCCGCCAAGTTCGTGCTTC
 TCATGCTGGTGGCATTCAATTCAACTGTCGCTTTGGTGGTCTCTGGAACCTTGTTT
 TACTCAGTCAGTGGAGGTGGCTCAGGCTGCTTTTGATATCAACGATTGGCACACCAA
 TCGCCAGGCATCCAGAGGGATATATCCTTTGTGATACTACGAGCCCAGAAACCCCTGA
 TGTATGTGGCCGAACATTTCTGCCCTTACCCTGGGAACCTATATGCTTGTACTGAA
 GAACTGCTATCGTTTGCTGGCCCTGATGCAAGAATCGATGTAG

DOR28

5 MYSPEEAAELKRRNYRSIREMIRLSYTVGFNLLDPSRCGQVLRIWTIVLSVSSLASLY
 GHWQMLARYIHDIPRIGETAGTALQFLTSLAKMWYFLFAHRQIYELLRKARCHELLQK
 CELFERMSDLPVIKEIRQQVESTMNRYWASTRRQILIYLYSCICITTNFYINSFVINL
 10 YRYFTKPKGSYDMLPLPSLYPAWEHKGLEFPYYHIQMYLETCSLYICGMCAVSFDGV
 FIVLCLHSVGLMRSLNQMVQATSELVPPDRRVEYLRCCLYQYQVRVANFATEVNNCFR
 HITFTQFLLSLFNWGLALFQMSVGLGNNSSITMIRMTMYLVAAGYQIVVYCYNGQRF
 15 TASEEIANAFYQVRWYGESREFRHLIRMLMRTNRGFRLDVSWFMQMSLPTLMAVSSG
 AEQSRGPAGPAGPAGPPRVPSSYSQFHLIDSQMVRTSGQYFLLLQNVNQK

DOR28nt

20 ATGTACTCACCGGAAGAGGCGGCCGAAGAGGCGCAACTATCGCAGCATCAGG
 GAGATGATCCGACTCTCCTATACGGTGGGCTTCAACCTGTTGGATCCTTCCCGATGCG
 GACAGGTGCTCAGAATCTGGACAATTGTCCTTAGCGTGAGTAGCTTGGCATCGCTTAA
 25 TGGGCACTGGCAAATGTTAGCCAGGTACATTCATGATATTCCACGCATTGGAGAGACC
 GCTGGAAGTGGCCTGCAGTTCCTAACATCGATAGCAAAGATGTGGTACTTTCTGTTTG
 CCCATAGACAGATATACGAATTGCTACGAAAGGCGCGCTGCCATGAATTACTCCAAAA
 30 GTGTGAGCTCTTTGAAAGGATGTCAGATCTACCTGTTATCAAAGAGATTGCCAGCAG
 GTTGAGTCCACGATGAATCGGTACTGGGCCAGCACTCGTCGGCAAATTCTTATCTATT
 TGTACAGCTGTATTTGTATTACTACAACTACTTTATCAACTCCTTCGTAATCAACCT
 35 CTATCGCTATTTCACTAAACCGAAAGGATCCTACGACATAATGTTACCTCTGCCATCT
 CTGTATCCCGCCTGGGAGCACAAGGGATTAGAGTTTCCCTACTATCATATACAGATGT
 ACCTGGAAACCTGTTCTCTGTATATCTGCGGCATGTGTGCCGTTAGCTTTGATGGAGT
 40 CTTTATTGTCCTGTGCCTTCATAGCGTGGGACTTATGAGGTCACCTAACCAGGTTGAGT
 GAACAAGCCACATCTGAGTTGGTTCCTCCAGATCGCAGGGTTGAATACTTGCGATGCT
 GTATTTATCAGTACCAACGAGTGGCGAACTTTGCAACCGAGGTTAACAACCTGCTTTCC
 45 GCACATCACTTTACGCGAGTTCCTGCTTAGCCTTTTCAACTGGGGCCTGGCCTTGTTT
 CAAATGAGCGTCGGATTGGGCAACAACAGCAGCATCACCATGATCCGGATGACCATGT
 ACCTGGTGGCAGCCGGCTATCAGATAGTTGTGTACTGCTACAATGGCCAGCGATTGTC
 GACTGCTAGCGAGGAGATTGCCAACGCCTTTTACCAGGTGCGATGGTACGGAGAGTCC
 50 AGGGAGTTCCGCCACCTCATCCGCATGATGCTGATGCGCACGAACCGGGGATTACGGC
 TGGACGTGTCCTGGTTCATGCAAATGTCCTTGCCCACTCATGGCGGTGAGTAGCGG
 AGCAGAGCAGAGCAGGGTCCCTGCAGGTCCCTGCAGGTCCCTGCAGGTCCACCCCAAGG
 55 GTCCCTCCTACAGCCAGTTCCACTTGATTGATTGCGAGATGGTCCGGACAAGTGGAC
 AGTACTTCCTGCTGCTGCAGAACGTCAACCAGAAA

DOR30

MAVSTRVATKQEVPESTRRAFRNLFNCFYALGMQAPDGSRPPTSSTWQRIYACFSVVMY
 VWQLLLVPTFFVISYRYMGGMEITQVL TSAQVAIDAVILPAKIVALAWNLP LLRRAEH
 HLAALDARCREQE EFQLILDAVRFCNYLVWFYQICYAIYSSSTFVCAFLLGQPPYALY
 LPGLDWQRSQMQFCIQAWIEFLIMNWTCLHQASDDVYAVIYLYVVRIQVQLLARRVEK
 LGTDDSGQVEIYPDERRQEEHCAELQRCIVDHQTM LQLLDCISPVISRTIFVQFLITA
 AIMGTTMINIFIFANTNTKIASIIYLLAVTLQTAPCCYQATSLMLDNERLALAI FQCQ
 WLGQSARFRKMLLYLHRAQQPITLTAMKLF PINLATYFSIAKFSFSLYTLIKGMNLG
 ERFNRTN

DOR30nt

ATGGCGGTGAGCACTCGTGTGGCCACAAAGCAGGAAGTGCCCGAATCCCGGCGAGCGT
 TTAGGAATCTCTTCAATTGCTTCTATGCCCTTGGCATGCAGGCACCGGATGGCAGTCG
 ACCGACCACGAGCAGCACATGGCAACGCATCTACGCCTGCTTCTCGGTGGTCATGTAC
 GTGTGGCAACTGCTGCTGGTGCCACATTCTTTGTGATCAGCTATCGGTACATGGGCG
 GCATGGAGATTACCCAGGTGCTGACCTCCGCCCAGGTGGCCATCGATGCGGTCAATTCT
 GCCGCCAAGATTGTGGCACTGGCGTGGAATTTGCCATTGCTGCGCAGAGCAGAGCAT
 CATCTGGCCGCCTTGGATGCGCGGTGCAGGGAACAGGAGGAGTTCCAATTGATCCTCG
 ATGCGGTGAGGTTTTGCAACTATCTGGTATGGTTCTACCAGATCTGCTATGCCATCTA
 CTCCTCGTCGACATTTGTGTGCGCCTTCCTGCTGGGCCAACCGCCATATGCCCTCTAT
 TTGCTTGGCCTCGATTGCGCAGCGTTCCAGATGCAGTTCTGCATCCAGGCCTGGATTG
 AGTTCTTATCATGAACTGGACGTGCCTGCACCAAGCTAGCGATGATGTGTACGCCGT
 TATCTATCTGTATGTGGTCCGGATTCAAGTGCAATTGCTGGCCAGGCGGGTGGAGAAG
 CTGGGCACGGATGATAGTGGCCAGGTGGAGATCTATCCCGATGAGCGGCGGCAGGAGG
 AGCATTGCGCGGAACTGCAGCGCTGCATTGTAGATCACCAGACGATGCTGCAGCTGCT
 CGACTGCATTAGTCCCGTCATCTCGCGTACCATATTCGTTCAAGTTCTGATCACC GCC
 GCCATCATGGGCACCAACCATGATCAACATTTTCATTTTCGCCAATACGAACACGAAGA
 TCGCATCGATCATTTACCTGCTGGCGGTGACCCTGCAGACGGCTCCATGTTGCTATCA
 GGCCACCTCGCTGATGTTGGACAACGAGAGGCTGGCCCTGGCCATCTTCCAGTGCCAG
 TGGCTGGGCCAGAGTGCCCGGTTCCGTAAGATGCTGCTCTACTATCTTCATCGCGCCC
 AGCAGCCCATCACGCTGACCGCCATGAAGCTGTTTCCCATCAATCTGGCCACGTACTT
 CAGTATAGCCAAGTTCTCGTTTTTCGCTCTACACGCTCATCAAGGGGATGAATCTCGGC
 GAGCGATTCAACAGGACAAAT

DOR31

MIFKYIQEPVLGSLFRSRDSLIIYLNRSIDQMGWRLPPRTKPYWWLYYIWTLVVIVLVF
 IFIPYGLIMTGIKEFKNFTTTDLFTYVQVPVNTNASIMKGIIVLFMRRRFSRAQKMMMD

ES 2 315 227 T3

AMDIRCTKMEEKVQVHRAAALCNRVVVIYHCIYFGYLSMALTGALVIGKTPFCLYNPL
VNPDDHFYLATAIESVTMAGIILANLILDVYPIIYVVVLRIHMELLSERIKTLRTDVE
5 KGDDQHYAELVECVKDHKLIVEYGNTLRPMISATMFIQLLSVGLLLGLAAVSMQFYNT
VMERVVSGVYTIAILSQTFFCYVCEQLSSDCESLTNTLFHRSKWIGAERRYRTTMLYF
IHNQQSILFTAGGIFPICLNTNIKMAKFAFSVVTIVNEMDLAEKLRRE

DOR31nt

ATGATTTTTAAGTACATTCAAGAGCCAGTCCTTGGATCCTTATTTTCGATCCCGGGATT
15 CGCTGATCTACTTAAACAGATCCATAGATCAAATGGGATGGAGACTGCCGCCACGAAC
TAAGCCGTACTGGTGGCTCTATTACATTGGACATTGGTGGTCATAGTACTCGTCTTT
ATCTTTATACCCTATGGACTGATAATGACTGGAATAAAGGAGTTCAAGAACTTCACGA
20 CCACGGATCTGTTTACGTATGTCCAGGTGCCGGTTAACACCAATGCTTCGATCATGAA
GGGCATTATAGTGTGTGTTTATGCCGCCGGCGATTTTCAAGGGCTCAGAAGATGATGGAC
GCCATGGACATTTCGATGCACCAAGATGGAGGAGAAAGTCCAGGTGCACCGAGCAGCAG
25 CCTTATGCAATCGTGTGTTGTGATTACCATTGCATATACTTCGGCTATCTATCCAT
GGCCTTAACCGGAGCTCTGGTGATTGGGAAGACTCCATTCTGTTTGTACAATCCACTG
GTTAACCCCGACGATCATTTCTATCTGGCCACTGCCATTGAATCGGTACCATGGCTG
GCATTATTCTGGCCAATCTCATTTTGGACGTATATCCCATCATATATGTGGTCGTTCT
30 GCGGATCCACATGGAGCTCTTGAGTGAGCGAATCAAGACGCTGCGTACTGATGTGGAA
AAAGGCGACGATCAACATTATGCCGAGCTGGTGGAGTGTGTAAAGGATCACAAGCTAA
TTGTTCGAATATGGAAACACTCTGCGTCCCATGATATCCGCCACGATGTTTCATCCAAC
35 ACTATCCGTTGGCTTACTTTTGGGTCTGGCAGCGGTGTCCATGCAGTTCTATAACACC
GTAATGGAGCGTGTGTTCTCCGGGGTCTACACCATAGCCATTCTATCCCAGACCTTTC
CATTTTGCTATGTCTGTGAGCAGCTGAGCAGCGATTGCGAATCCCTGACCAACACACT
40 GTTCATTCCAAGTGGATTGGAGCTGAGCGACGATACAGAACCACGATGTTGTACTTC
ATTACAATGTTTCAGCAGTCGATTTTGTTCAGTGGCGGGAATTTTCCCATATGTC
TAAACACCAATATAAAGATGGCCAAGTTCGCTTTCTCAGTGGTGACCATTGTAAATGA
45 GATGGACTTGGCCGAGAAATTGAGAAGGGAG

DOR32

MEPVQYSYEDFARLPTTVFWIMGYDMLGVPKTRSRRIYWIYRFLCLASHGVCVGVMV
FRMVEAKTIDNVSLIMRYATLVTYIINSDTKFATVLQRSQSLNSKLAEVLPKTTLD
55 RIYHRVNDHYWTKSFVYLVIIYIGSSIMVVIGPIITSIIAYFTHNVFTYMHCPYFLY
DPEKDPVWIYISIIYALEWLHSTQMVISNIGADIWLLYFQVQINLHFRGIIRSLADHKP
SVKHDQEDRKFIKIVDKQVHLVSLQNDLNGIFGKSLLSLLTTAAVICTVAVYTLIQ
60 GPTLEGFTYVIFIGTSVMQVYLVCIYQGVLDLSGEVAHAVYNHDFHDASIAYKRYLL
IIIIIRAQQPVELNAMGYLSISLDTFKQLMSVSYRVITMLMQMIQ

DOR32nt

ATGGAACCTGTGCAGTACAGCTACGAGGATTTCTGCTCGATTGCCCACGACGGTGTCT
 GGATCATGGGCTACGACATGCTGGGCGTTCCGAAGACCCGCTCTCGCAGGATACTATA
 CTGGATATATCGTTTCCTCTGTCTCGCCAGCCATGGGGTCTGTGTAGGAGTCATGGTA
 TTTCGTATGGTGGAGGCAAAGACCATTGACAATGTTTCGCTGATCATGCGGTATGCCA
 CTCTGGTCACTATATCATCAACTCGGATACGAAATTCGCAACTGTCTTACAAAGGAG
 TGCAATTCAAAGTCTAAACTCAAACTGGCCGAATATATCCGAAGACCACGCTGGAC
 AGGATCTATCACCGGGTGAATGATCACTATTGGACCAAGTCATTTGTATATTTGGTTA
 TTATCTACATTGGTTTCGTCGATTATGGTTGTTATTGGACCGATTATTACGTCGATTAT
 AGCTTACTTCACGCACAACGTTTTACCTACATGCACTGCTATCCGTACTTTTGTAT
 GATCCTGAGAAGGATCCGGTTTGGATCTACATCAGCATCTATGCTCTGGAATGGTTGC
 ACAGCACACAGATGGTCATTTTGAACATTGGCGCGGATATCTGGCTGCTGTACTTTCA
 GGTGCAGATAAATCTCCACTTCAGGGGCATTATACGATCACTGGCGGATCACAAGCCC
 AGTGTGAAGCACGACCAGGAGGACAGGAAATTCATTGCGAAAATTGTCGACAAGCAGG
 TGCACCTGGTCAGTTTGCAAAACGATCTGAATGGTATCTTTGGAAAATCGCTGCTTCT
 AAGCCTGCTGACCACCGCAGCGGTTATCTGCACGGTGGCGGTGTACACTCTGATTAG
 GGTTCCACCTTGGAGGGCTTACCTATGTGATCTTCATCGGGACTTCTGTGATGCAGG
 TCTACCTGGTGTGCTATTACGGTCAGCAAGTTCTCGACTTGAGCGGCGAGGTGGCCCA
 CGCCGTGTACAATCATGATTTTACGATGCTTCTATAGCGTACAAGAGGTACCTGCTC
 ATAATCATTATCAGGGCGCAGCAGCCCGTGGAACTTAATGCCATGGGCTACCTGTCCA
 TTTCGCTGGACACCTTTAAACAGCTGATGAGCGTCTCCTACCGGGTTATAACCATGCT
 CATGCAGATGATTAG

DOR37

****La secuencia proteica está incompleta y está en progreso****

KVDSTRALVNHWRIFRIMGIHPPGKRTFWGRHYTAYSMVWNVTFHICIWVSFSVNLLQ
 SNSLETFCESLCVTMPHTLYMLKLINVRMRGQMISHWLLRLLDKRLGCDDERQIIM
 AGIERAEFIFRTIFRGLACTVVLGIIYISASSEPTLMYPTWI PWNWRDSTSAYLATAM
 LHTTALMANATLVNLSSYPGTYLILVSVHTKALALRVSKLGYGAPLPAVRMQAILVG
 YIHDHQIILR*VSGNLISQCKNF*SISGVLTFIERRYTHFGVPNIFIVIEDYIILFL
 NYSLFKSLERSLSMTCFLQFFSTACAQCTICYFLLFGNVGIMRFMNMLFLLVILTTET
 LLLCYTAELPCKEGESLLTAVYSCNWSQSVNFRLLLLMLARCQIPMILVSGVIVPI
 SMKTF

DOR37nt

****La información de la secuencia nucleotídica está en progreso ****

DOR38

ES 2 315 227 T3

MRLIKISYSALNEVCVWLKLNKSWPLTESSRPWRSQSLLATAYIVWAWYVIASVGITI
SYQTAFLLNLSDIITTTENCCTTFMGVLNFVRLIHLRLNQKFRQLIENFSYEIWIP
5 NSSKNVAAECRRRMVTFSIMTSLLACLIIMYCVLPLVEIFFGPAFDAQNKPFPYKMI
FPYDAQSSWIRYVMTYIFTSYAGICVVTTLFAEDTILGFFITYTCGQFHLLHQRIAGL
FAGSNAELAESIQLERLKRIVEKHNNIISANSV

DOR38nt

ATGCGTTTGATCAAAATTTTCATATTCGGCACTTAATGAGGTGTGCGTTTGGCTGAAAC
15 TGAATGGTTCTTGCCATTAACCGAATCATCGAGGCCATGGAGGAGCCAATCCTTATT
GGCCACCGCCTACATCGTGTGGGCGTGGTACGTCAATGCATCTGTGGGCATAACAATC
AGCTATCAGACGGCCTTTTTTGCTGAACAACCTTTCGGACATTATTATCACCACGGAAA
20 ATTGTTGCACCACCTTTATGGGTGTCCTGAACTTTGTCCGACTCATCCATCTTCGCCT
CAATCAGAGGAAATTCGCCAGCTTATTGAGAACTTTTCCTACGAAATTTGGATACCT
AATTCTTCCAAAAACAATGTTGCCGCCGAGTGTGCGAGACGCATGGTTACCTTCAGCA
25 TAATGACATCCTTGCTAGCGTGCCTGATCATAATGTATTGTGTCCTGCCGCTGGTGGA
GATCTTCTTTGGACCCGCCTTCGATGCACAGAACAAGCCGTTTCCCTACAAGATGATC
TTTCCGTACGATGCCCAGAGCAGTTGGATCCGATATGTGATGACCTACATCTTCACCT
30 CCTACGCGGGAATCTGTGTGGTCACCACCTTGTTTGAGAGGACACCATCTTGCTT
CTTCATAACCTACACTTGTGGCCAAATTTCAATTTGCTACACCAACGAATCGCAGGTTTA
TTTGCGGGTTCCAATGCGGAATTGGCCGAGAGCATTCAGCTGGAGCGACTCAAACGTA
35 TTGTGGAAAAACACAACAATATTATCAGCGCAAATTCTGTA

DOR44

MKSTFKEERIKDDSKRRDLFVFVRQTMCIAMYPFGYYVNGSGVLAVLVRFCDLTYEL
FNYFVSVHIAGLYICTIYINYGQDLDFFVNCLIQTIIYLWTIAMKLYFRFRPGLLN
TILSNINDEYETRSVAGFSFVTMAGSYRMSKLWIKTYVYCCYIGTIFWLALPIAYRDR
45 SLPLACWYPFDYTQPGVYEVVFLQAMGQIQVAASFASSSGLHMLCVLISGQYDVLF
CSLKNVLASSYVLMGANMTELNLQAEQSAADVEPGQYAYSVEEETPLQELLKVGSSM
DFSSAFRLSFVRICIQHRYIVAALKKIESFYSPIWFKIGEVTFMLCLVAFVSTKSTA
50 ANSFMRMVSLGQYLLLVLVELFIICYFADIVFQNSQRCGEALWRSPWQRHLKDVRSDY
MFFMLNSRRQFQLTAGKISNLNVDRFRGVGILT

DOR44nt

ATGAAGAGCACATTCAAGGAAGAAAGGATTAAGGACGACTCCAAGCGTCGCGACCTGT
60 TTGTATTCGTGAGGCAAACCATGTGTATAGCGGCCATGTATCCCTTCGGTTACTACGT
GAATGGATCTGGAGTCCTGGCCGTTCTGGTGCGATTCTGTGACTTGACCTACGAGCTC

ES 2 315 227 T3

TTTAACTACTTCGTTTCGGTACACATAGCTGGCCTGTACATCTGCACCATCTACATCA
ACTATGGGCAAGGCGATTTGGACTTCTTCGTGAACTGTTTGATACAAACCATTATTTA
5 TCTGTGGACAATAGCGATGAAACTCTACTTTCGGAGGTTTCAGACCTGGTTTGTGTAAT
ACCATTTCTGTCCAACATCAATGATGAGTACGAGACACGTTTCGGCTGTGGGATTTCAGTT
TCGTCACAATGGCGGGATCCTATCGGATGTCCAAGCTATGGATCAAAACCTATGTGTA
10 TTGCTGCTACATAGGCACCATTTTCTGGCTGGCTCTTCCCATTGCCTACCGGGATAGG
AGTCTTCCTCTTGCCTGCTGGTATCCCTTTGACTATACACAACCCGGTGTCTATGAGG
TAGTGTTCCTTCTCCAGGCGATGGGACAGATCCAAGTGGCCGCATCCTTTGCCTCCTC
15 CAGTGGCCTGCATATGGTGCTTTGTGTGCTGATATCAGGGCAGTACGATGTCTCTTT
TGCAGTCTCAAGAATGTATTAGCCAGCAGCTATGTCTTATGGGAGCCAATATGACGG
AACTGAATCAATTGCAGGCTGAGCAATCTGCGGCCGATGTGAGCCAGGTCAGTATGC
20 TTA CTCCGTGGAGGAGGAGACACCTTTGCAAGAACTTCTAAAAGTTGGGAGCTCAATG
GACTTCTCCTCCGCATTCAGGCTGTCTTTTGTGCGGTGCATTCAGCACCATCGATACA
TAGTGGCGGCACTGAAGAAAATTGAGAGTTTCTACAGTCCCATATGGTTTCGTGAAGAT
25 TGGCGAAGTCACCTTTCTTATGTGCCTGGTAGCCTTCGTCTCCACGAAGAGCACCGCG
GCCAACTCATTATGCGAATGGTCTCCTTGGGCCAGTACCTGCTCTTAGTTCTCTACG
AGCTGTTTCATCATCTGCTACTTCGCGGACATCGTTTTTCAGAACAGCCAGCGGTGCGG
30 TGAAGCCCTCTGGCGAAGTCCTTGGCAGCGACATTTGAAGGATGTTTCGCAGTGATTAC
ATGTTCTTTATGCTGAATTCCTCGCAGGCAGTTCCAACCTACGGCCGGAAAAATAAGCA
ATCTAAACGTGGATCGTTTCAGAGGGGTGGGTATCCTTACT

DOR46

MAEVRVDSLEFFKSHWTAWRYLGVAHFRVENWKNLYVFYSIVSNLLVTLCPVHLGIS
40 LFRNRTITEDIILNLTTFATCTACSVKCLLYAYNIKDVLEMERLLRLDERVVGPEQRS
IYGQVRVQLRNVLYVFIGIYMPCALFAELSFLFKEERGLMYPWFPPDWLHSTRNYYI
ANAYQIVGISFQLLQNYVSDCFPAVVLCLISSHIKMLYNRFEEVGLDPARDAEKDLEA
45 CITDCHKHILELFRRIEAFISLPMLIQFTVTALNVCIGLAALVFFVSEPMARMYFIFY
LAMPLQIFPSCFFGTDNEYWFGRLHYAAFSCNWHQTQNRSFKRKMMLFVEQSLKKSTAV
AGGMMRIHLDTFSTLKGAYSLFTIIRMRK

DOR46nt

ATGGCAGAGGTCAGAGTGGACAGTCTGGAGTTTTTCAAGAGCCATTGGACCGCCTGGC
55 GGTACTTGGGAGTGGCTCATTTTCGGGTGCGAACTGGAAGAACCTTTACGTGTTTTA
CAGCATTGTGTGCAATCTTCTCGTGACCCTGTGCTACCCCGTTACCTGGGAATATCC
60 CTCTTTTCGCAACCGCACCATCACCGAGGACATCCTCAACCTGACCACCTTTGCGACCT
GCACAGCCTGTTTCGGTGAAGTGCCTGCTCTACGCCTACAACATCAAGGATGTGCTGGA
GATGGAGCGGCTGTTGAGGCTTTTGGATGAACGCGTCGTGGGTCCGGAGCAACGCAGC
65 ATCTACGGACAAGTGAGGGTCCAGCTGCGAAATGTGCTATACGTGTTCATCGGCATCT

ES 2 315 227 T3

ACATGCCGTGTGCCCTGTTGCGCGAGCTATCCTTTCTGTTCAAGGAGGAGCGCGGTCT
GATGTATCCCGCCTGGTTTTCCCTTCGACTGGCTGCACTCCACCAGGAACTATTACATA
5 GCGAACGCCTATCAGATAGTGGGCATCTCGTTTCAGCTGCTGCAAACTATGTTAGCG
ACTGCTTTCGGCGGTGGTGTGTGCCTGATCTCATCCCACATCAAAATGTTGTACAA
CAGATTGAGGAGGTGGGCCTGGATCCAGCCAGAGATGCGGAGAAGGACCTGGAGGCC
10 TGCATCACCGATACAAGCATATTCTAGAGTGGGCAGGCGGCTCATTGGTTCGTGTTC
TATTTCACTTTCCAACTTTTTTCCAGACTATTCGACGCATCGAGGCCTTCATTTCCCT
GCCCATGCTAATTCAGTTCACAGTGACCGCCTTGAATGTGTGCATCGGTTTAGCAGCC
15 CTGGTGTTTTTCGTCAGCGAGCCCATGGCACGGATGTACTTCATCTTCTACTCCCTGG
CCATGCCGCTGCAGATCTTTCCGTCTGCTTTTTTCGGCACCGACAACGAGTACTGGTT
CGGACGCCTCCACTACGCGGCCTTCAGTTGCAATTGGCACACACAGAACAGGAGCTTT
20 AAGCGGAAAATGATGCTGTTTCGTTGAGCAATCGTTGAAGAAGAGCACCGCTGTGGCTG
GCGGAATGATGCGTATCCACCTGGACACGTTCTTTTCCACCCTAAAGGGGCCTACTC
CCTCTTTACCATCATTATTTCGGATGAGAAAG

DOR48

MERHYFMVPKFALSLIGFYPEQKRTVLVKLWSFFNFILTYGCYAEAYYGIHYIPINI
30 ATALDALCPVASSILSLVKMVAIWYQDELRLSLIERRFYTLATQLTFLLLCCGFCTST
SYSVRHLIDNILRRTGKDWIYETPFKMMFPDLLLLRLPLYPITYILVHWHGYITVVCF
VGADGFFLGLFCLYFTVLLCLQDDVCDLLEVENIEKSPSEAEEARIVREMEKLVDRHN
35 EVAELTERLSGVMVEITLAHFVTSSLIIGTSVVDILLFSGLGIIVYVYTCAVGVEIF
LYCLGGSHIMEACSNLARSTFSSHWHGHSVRVQKMTLLMVARAQRVLTIKIPFFSPSL
ETLTSILRFTGSLIALAKSVI

DOR48nt

ATGGAGCGCCATTATTTTCATGGTGCCAAAGTTTGCATTATCGCTGATTGGTTTTTATC
45 CCGAACAGAAGCGAACGGTTTTGGTGAACTTTGGAGTTTCTTCAACTTTTTTCATCCT
CACCTACGGCTGTTATGCAGAGGCTTACTATGGCATACTATATACCGATTAACATA
GCCACTGCATTGGATGCCCTTTGTCTGTGGCCTCCAGCATTTTGTCTGCTGGTGAAAA
50 TGGTCGCCATTTGGTGGTATCAAGATGAATTAAGGAGTTTGATAGAGCGGGTAAGATT
TTTAACAGAGCAACAGAAGTCCAAGAGGAACTGGGCTATAAGAAGAGGTTCTATACA
CTGGCAACGCAACTAACATTCCTGCTACTATGCTGTGGATTTTGCACCAGTACTTCCT
55 ATTCCGTCAGACATTTGATTGATAATATCCTGAGACGCACCCATGGCAAGGACTGGAT
CTACGAGACTCCGTTCAAGATGATGTAAGGAAAGGGAAGAATGGTTTATATATACTTT
TGGAACGAAATAATGATGTGATCTAAACAAGATGCACTTTTTTTTAGGTTCCCCGATC
60 TTCTCCTGCGTTTGCCACTCTATCCCATCACCTATATACTCGTGCATTGGCATGGCTA
CATTACTGTGGTTTGTGTTTGTCTGGCGCGGATGGTTTCTCCTGGGGTTCTGTTGTAC
65 TTCACTGTTTTGCTGCTCTGTCTGCAGGACGATGTTTGTGATTACTAGAGGTTGAAA

ACATCGAGAAGAGTCCCTCCGAAGCGGAGGAAGCTCGCATAGTTCGGGAAATGGAAAA
 ACTGGTGGACCGGCATAACGAGGTGGCCGAGCTGACAGAAAGATTGTCGGGTGTTATG
 5 GTGGAAATAACACTGGCCCACTTTGTTACTTCGAGTTTGATAATCGGAACCAGCGTGG
 TGGATATTTTATTAGTGGGTATTTACATTTGATTAGATCCTTTTCGATATATGTTCTTA
 AATTCTAGTTTTCGGCCTGGGAATCATTGTGTATGTGGTCTACACTTGTGCCGTAGG
 10 TGTGGAAATATTTCTATACTGTTTAGGAGGATCTCATATTATGGAAGCGGTATATTCA
 TAAGAACTACTATAAAGTTACTTTTAAATTCATTGCATTTCTTAGTGTTCCAATCTA
 GCGCGCTCCACATTTTCAGCCACTGGTATGGCCACAGTGTTCCGGTCCAAAAGATGA
 15 CCCTTTTGATGGTAGCTCGTGCTCAACGAGTTCTCACAATTAAAATTCCTTTCTTTTC
 CCCATCATTAGAGACTCTAACTTCGGTAAGCTTATGCGAAAATGTTATGGTACACACA
 AGTCTACATTTCTATGAGGTCTTGTAGATTTTGCGCTTCACTGGATCTCTGATTGCCC
 20 TGGCAAAGTCGGTTATA

DOR53

25 MLSKFFPHIKEKPLSERVKSRAFIYLDVMWSFGWTEPENKRWILPYKLWLAFVNI
 MLILLPISISIEYLHRFKTFSAGEFLSSLEIGVNMYGSSFKCAFTLIGFKKRQEAKVL
 LDQLDKRCLSDKERSTVHRYVAMGNFFDILYHIFYSTFVVMNFPYFLLERRHAWRMF
 30 PYIDSDEQFYISSIAECFLMTEAIYMDLCTDVCPLISMLMARCHISLLKQRLRNLRSK
 PGRTEDEYLEELTECIRDHRLLLDYVDALRPVFSGTIFVQFLLIGTVLGLSMINLMFF
 STFWTGVATCLFMFDVSMETFPFCYLCNMIIDDCQEMSNCLFQSDWTSADRRYKSTLV
 35 YFLHNLQQPITLTAGGVFPI SMQTNLAMVKLAFSVVTVIKQFNLAERFQ

DOR53nt

40 TCAAACAAAGCCACGGACAAGATGTTAAGCAAGTTTTTTCCCCACATAAAAGAAAAGC
 CATTGAGCGAGCGGGTTAAGTCCCGAGATGCCTTCATTTACTTGGATCGGGTGATGTG
 45 GTCCTTTGGCTGGACAGAGCCTGAAAACAAAAGGTGGATCCTTCCTTATAAACTGTGG
 TTAGCGTTCGTGAACATAGTAATGCTCATCCTTCTGCCGATCTCGATAAGCATCGAGT
 ACCTCCACCGATTTAAACCTTCTCGGCGGGGGAGTTCCTTAGTTCCTCGAGATTGG
 50 AGTCAACATGTACGGAAGCTCTTTTAAGTGCGCCTTCACCTTGATTGGATTCAAGAAA
 AGACAGGAAGCTAAGGTTTACTGGATCAGCTGGACAAGAGATGCCTTAGCGATAAGG
 AGAGGTCCACTGTTTCATCGCTATGTCGCCATGGGAACTTTTTCGATATTTTGTATCA
 55 CATTTTTTACTCCACCTTCGTGGTAATGAACTTCCCGTATTTTCTGCTTGAGAGACGC
 CATGCTTGGCGCATGTACTTTCCATATATCGATTCCGACGAACAGTTTTACATCTCCA
 GCATCGCCGAGTGTTTTCTGATGACGGAGGCCATCTACATGGATCTCTGTACGGACGT
 60 GTGTCCCTTGATCTCCATGCTTATGGCTCGATGCCACATCAGCCTCCTGAAACAGCGA
 CTGAGAAATCTCCGATCGAAGCCAGGAAGGACCGAAGATGAGTACTTGAGAGAGCTCA
 CCGAGTGCATTCCGGATCATCGATTGCTATTGGACTATGTTGACGCATTGCGACCCGT
 65 CTTTTCGGGAACCATTTTTGTGCAGTTCCTCCTGATCGGTACTGTACTGGGTCTCTCA

ES 2 315 227 T3

ATGATAAATCTAATGTTCTTCTCGACATTTTGGACTGGTGTGCCACTTGCCTTTTTTA
TGTTTCGACGTGTCCATGGAGACGTTCCCCTTTTGCTATTTGTGCAACATGATTATCGA
5 TGACTGCCAGGAAATGTCCAATTGCCTCTTTCAATCGGACTGGACCTCTGCCGATCGT
CGCTACAAATCCACTTTGGTATACTTTCTTCACAATCTTCAGCAACCCATTACTCTCA
CGGCTGGTGGAGTGTTTCCTATTTCCATGCAAAACAAATTTGGCTATGGTGAAGCTGGC
10 ATTTTCTGTGGTTACGGTAATTAAGCAATTTAACTTGGCCGAAAGGTTTCAATAAGTT
GAGAGGGACGAGCTCTGCTACTATTATATTATATATTATATTATATTATATATATT
ATTTTATATTATATATTGCTGTACCCTAATAAATATTTAGTAATAAAAAAAAAAAAAA
15 AAAA

DOR56

MDPVEMPIFGSTLKLKMFWSYLFVHNWRRYVAMTPYIIINCTQYVDIYLSTESLDFII
RNVYLAVLFTNTVVRGVLLCVQRFSEYERFINILKSFYIELLVSTERLSQKCILHKWAV
LPYGYMLPTIDEYKYASPYEIFFVIQAIMAPMGCCMYIPYTNMVVTFTLFAILMCRV
25 LQHKLRSLKLEKNEQVRGEIAQTIAQTVIVIAVMVIFANSVVLYYVANELYFQSFDI
AIAAYESNWMDFDVTQKTLKFLIMRSQKPLASLVGGTYPMNLKMLQSLLNAIYSFFT
LLRRVYG

DOR56nt

ATGGATCCGGTGGAGATGCCCATTTTGGTAGCACTCTGAAGCTAATGAAGTTCTGGT
CATATCTGTTTGTTCACAACTGGCGCCGCTATGTCGCAATGACTCCGTACATCATTAT
CAACTGTACTCAGTATGTGGATATATATCTGAGCACCGAATCCTTGGACTTTATCATC
40 AGAAATGTATACCTGGCTGTATTGTTTACCAACACGGTGGTCAGAGGTGTATTGTTAT
GCGTACAGCGGTTTAGCTACGAGCGTTTCATTAATATTTTGAAAAGCTTTTACATTGA
GTTGTTGGTGAGTACCGAAAGATTATCTCAAAAATGCATATTGCATAAATGGGCAGTT
CTGCCATATGGCATGTATTTGCCCACTATTGATGAATACAAATACGCATCACCTTACT
45 ACGAGATTTTCTTTGTGATTCAAGCCATTATGGCTCCAATGGGGTGTTCATGTACAT
ACCATACACAAACATGGTAGTGACATTTACCCCTTTTCGCCATTCTCATGTGTCGAGTG
TTGCAACATAAGTTGAGAAGCCTAGAAAAGCTGAAAAATGAACAAGTACGTGGTGAAA
50 TCGCTCAAACAATTGCTCAGACCGTCATAGTCATCGCATACATGGTAATGATATTTGC
CAACAGTGTAGTCCTTTACTACGTGGCCAATGAGCTATACTTTCAAAGCTTTGATATT
GCCATTGCTGCCTATGAGAGCAATTGGATGGACTTTGATGTGGACACACAAAAGACTT
55 TGAAGTTCCTCATCATGCGCTCGCAAAGCCCTTGGCGAGTCTGGTGGGTGGCACATA
TCCCATGAACTTGAAAATGCTTCAGTCACTACTAAATGCCATTTACTCCTTCTTCACC
CTTCTGCGTCGCGTTTACGGC

DOR58

ES 2 315 227 T3

MDASYFAVQRRRALEIVGFDPSTPQLSLKHPIWAGILILSLISHNWPVYALQDLSDL
TRLTDNFAVFMQGSQSTFKFLVMMAKRRRIGSLIHRLHKLNQAAATPNHLEKIEREN
5 QLDYVARSFRNAAAYGVICASAIAPMLLGLWGYVETGVFTPTTPEFNFWLDERKPHF
YWPIYVWGVLGVAAAAWLAIATDTLFSWLTHNVVIQFQLELVLEEKDLNGGDSRLTG
FVSRHRIALDLAKELSSIFGEIVFVKYMLSYLQCLMAFRFSRSGWSAQVPFRATFLV
10 AII IQLSSYCYGGEYIKQQLAIAQAVYGQINWPEMTPKKRRLWQMVIMRAQRPAKIF
GFMFVVDLPLLLWVIRTAGSFLAMLRTFER

DOR58nt

ATGGACGCCAGCTACTTTGCCGTCCAGAGAAGAGCTCTGGAAATAGTTGGATTTCGATC
20 CCAGTACTCCGCAACTGAGTCTGAAACATCCCATCTGGGCCGGGATTCTCATCCTGTC
CTTGATCTCTCACAACCTGGCCCATGGTAGTCTATGCCCTGCAGGATCTCTCCGACTTG
ACCCGTCTGACGGACAACTTTGCCGTGTTTATGCAAGGATCACAGAGCACCTTCAAGT
25 TCCTGGTCATGATGGCGAAACGAAGGCGCATTGGATCGTTGATTCACCGTTTGCATAA
GCTAAACCAGGCGGCCAGTGCCACGCCCAATCACCTGGAGAAGATCGAGAGGGAAAAC
CAACTGGATAGGTATGTCGCCAGGTCCTTTAGAAATGCCGCTACGGAGTGATTTGTG
CCTCGGCCCATAGCGCCCATGTTGCTTGGCCTGTGGGGATATGTGGAGACGGGTGTATT
30 TACCCCCACCACACCCATGGAGTTCAACTTCTGGCTGGACGAGCGAAAGCCTCACTTT
TATTGGCCCATCTACGTTTGGGGCGTACTGGGCGTGGCAGCTGCCGCCTGGTTGGCCA
TTGCAACGGACACCCGTGTTCTCCTGGCTGACTCACAATGTGGTGATTTCAGTTCCAAC
35 ACTGGAGCTTGTTCTCGAAGAGAAGGATCTGAATGGCGGAGACTCTCGCCTGACCGGG
TTTGTTAGTCGTCATCGTATAGCTCTGGATTTGGCCAAGGAACTAAGTTTCGATTTTCG
GGGAGATCGTCTTTGTGAAATACATGCTCAGTTACCTGCAACTCTGCATGTTGGCCTT
40 TCGCTTCAGCCGAGTGGCTGGAGTGCCAGGTGCCATTTAGAGCCACCTTCCTAGTG
GCCATCATCATCCAACCTGAGTTTCGTATTGCTATGGAGGCGAGTATATAAAGCAGCAAA
GTTTGGCCATCGCACAAGCCGTTTATGGTCAAATCAATTGGCCAGAAATGACGCCAAA
45 GAAAAGAAGACTCTGGCAAATGGTGATCATGAGGGCGCAGCGACCGGCTAAGATTTTT
GGATTCATGTTTCGTTGTGGACTTGCCACTGCTGCTTTGGGTCATCAGAACTGCGGGCT
CATTTCTGGCCATGCTTAGGACTTTTCGAGCGT

DOR59

MHEADNREMELLVATQAYTRTITLLIWIPSVIAGLMAYSDCIYRSLFLPKSVFNVPV
55 RRGEHPILLFQLFPFGELCDNFVVGYLGPWYALGLGITAIPLWHTFITCLMKYVNLK
LQILNKRVEEMDITRLNSKLVIGRLTASELTFWQMQLFKEFVKEQLRIRKFVQELQYL
ICVPVMADFIIFSVLICFLFFALTVGHDELSLAYFSCGWYNFEMPLQKMLVFMMHHAQ
60 RPMKMRALLVDLNLRTFIDIGRGAYSFYNNLRSSHL

DOR59nt

ES 2 315 227 T3

ATGCACGAAGCAGATAATCGGGAGATGGAACCTTTTGGTCGCCACTCAGGCTTATACAC
GAACCATTACCCTGTTGATCTGGATAACCATCGGTTATTGCTGGCCTAATGGCCTATTC
5 AGACTGCATCTACAGGAGTCTGTTTCTGCCGAAATCGGTTTTCAATGTGCCAGCTGTG
CGACGTGGTGAGGAGCATCCCATTTCTGCTATTTTCAGCTGTTTCCCTTCGGAGAACTTT
GCGATAACTTCGTTGTTGGATACTTGGGACCTTGGTATGCTCTGGGCCTGGGAATCAC
10 GGCTATCCCATTTGTGGCACACCTTTATCACTTGCCCTCATGAAGTACGTAAATCTCAAG
CTGCAAATACTCAACAAGCGAGTGGAGGAGATGGATATTACCCGACTTAATTCCAAAT
TGGTAATTGGTCGCCTAACTGCCAGTGAGTTAACCTTCTGGCAAATGCAACTCTTCAA
15 GGAATTTGTAAAGGAACAGCTGAGGATTCGAAAATTTGTCCAGGAACTACAGTATCTG
ATTTGCGTGCCTGTGATGGCAGATTTTATTATCTTCTCGGTTCTCATTGCTTTCTCT
TTTTTGCTTGACAGTTGGCCACGATGAACTGAGCCTTGCTTACTTTTCTTGCGGATG
20 GTACAACTTCGAAATGCCTTTGCAGAAAATGCTGGTTTTTATGATGATGCATGCCCAA
AGGCCGATGAAGATGCGCGCCCTGCTGGTCGATTTGAATCTGAGGACCTTCATAGACA
TTGGCCGTGGAGCCTACAGCTACTTCAATTTGCTGCGTAGCTCCCACTTGTAT

DOR61

MGHKDDMDSTDSTALSLKHISLIFVISAQYPLISYVAYNRNDMEKVTACLSVVFTNM
30 LTVIKISTFLANRKDFWEMIHRFRKMHEQCKYREGLDYVAEANKLASFLGRAYCVSCG
LTGLYFMLGPVIGVCRWHGTTCDKELPMPMKFPFNDLESPGYEVCFLYTVLVTVVV
VAYASAVDGLFISFAINLRAHFQTLQRQIENWEFPSSEPDTQIRLKSIVEYHVLLLSL
35 SRKLRSIYPTVMGQFVITSLQVGVI IYQLVTNMDSVMDLLLYASFFGSIMLQLFIYC
YGGEI IKAESLQVDTAVRLSNWHLASPKTRTSLSLI ILSQKEVLIRAGFFVASLANF
PYRLITLIKSIDSIC

DOR61nt

La información de la secuencia nucleotídica está en progreso

DOR62

MEKQEDFKLNTHSAVYYHWRVWELTGLMRPPGVSSLLYVVYSITVNLVVTVLFPPLSL
45 ARLLFTTNMAGLCENLTITITDIVANLKFANVYMVRKQLHEIRSLRLMDARARLVGD
PEEISALRKEVNIAQGTFRTFASIFVFGTTLSCVRVVVRPDRELLYPAWFGVDWMHST
RNYVLINIYQLFGLIVQAIQNCASDSYPPAFLCLLTGHMRALELRVRRIGCRTEKSNK
50 GQTYEAWREEVYQELIECIRDLARVHRLREI IQRVLSVPCMAQFVCSAAVQCTVAMHF
LYVADDHDHTAMI IISIVFFSAVTLEVFVICYFGDRMRTQSEALCDAFYDCNWIQLPK
55 FKRELLFTLARTQRPSLIYAGNYIALSLETFEQVMRFTYSVFTLLLRK

DOR62nt

ES 2 315 227 T3

ATGGAGAAGCAAGAGGATTTCAAACCTGAACACCCACAGTGCTGTGTACTACCACTGGC
GCGTTTGGGAGCTCACTGGCCTGATGCGTCCTCCGGGCGTTTCAAGCCTGCTTTACGT
5 GGTATACTCCATTACGGTCAACTTGGTGGTCACCGTGCTGTTTCCCTTGAGCTTGCTG
GCCAGGCTGCTGTTACCACCAACATGGCCGGATTGTGCGAGAACCTGACCATAACTA
TTACCGATATTGTGGCCAATTTGAAGTTTGCGAATGTGTACATGGTGAGGAAGCAGCT
10 CCATGAGATTGCTCTCTCCTAAGGCTCATGGACGCTAGAGCCCGGCTGGTGGGCGAT
CCCGAGGAGATTTCTGCCTTGAGGAAGGAAGTGAATATCGCACAGGGCACTTTCCGCA
CCTTTGCCAGTATTTTCGTATTTGGCACTACTTTGAGTTGCGTCCGCGTGCTCGTTTCG
15 CCCGGATCGAGAGCTCCTGTATCCGGCCTGGTTCCGGCGTTGACTGGATGCACTCCACC
AGAAACTATGTGCTCATCAATATCTACCAGCTCTTCGGCTTGATAGTGCAGGCTATAC
AGAACTGCGCTAGTGACTCCTATCCGCCTGCGTTTCTCTGCCTGCTCACGGGTCATAT
20 GCGTGCTTTGGAGCTGAGGGTGCGGCGGATTGGCTGCAGGACGGAAAAGTCCAATAAA
GGGCAGACATATGAAGCCTGGCGGGAGGAGGTGTACCAGGAACTCATCGAGTGCATCC
GCGATCTGGCGCGGGTCCATCGGCTGAGGGAGATCATTACGCGGGTCCTTTCAGTGCC
CTGCATGGCCAGTTTCGTCTGCTCCGCCGCCGTCCAGTGACCGTCGCCATGCACTTC
25 CTGTACGTAGCGGATGACCACGACCACACCGCCATGATCATCTCGATTGTATTTTCT
CGGCCGTACCTTGGAGGTGTTTGTAACTCTGCTATTTTGGGGACAGGATGCGGACACA
GAGCGAGGCGCTGTGCGATGCCTTCTACGATTGCAACTGGATAGAACAGCTGCCCAAG
30 TTCAAGCGCGAACTGCTCTTACCCTGGCCAGGACGCAGCGGCCTTCTCTTATTTACG
CAGGCAACTACATCGCACTCTCGCTGGAGACCTTCGAGCAGGTGATGAGGTTACATA
CTCTGTTTTCACACTCTTGCTGAGGGCCAAGTAAGAACTTTATAATCTTTTTTGGGG
35 AGAAAAATTTTAAAGCACAATAGCAGAAAAATATATCAGATAATATAACAAAAAAA
AAAAAAA

DOR64

MKLSETLKDIDYFRVQLNAWRICGALDLSEGRYWSWSMLLCILVYLPTPMLLRGVYSFE
DPVENNFSLSLTVTSLSNLMKFCMYVAQLTKMVEVQSLIGQLDARVSGESQSERHRNM
45 TEHLLRMSKLFQITYAVVFI IAAVPFVFETELSLPMPMWFPPDWKNSMVAYIGALVFQ
EIGYVFQIMQCFAADSFPPLVLYLISEQCQLLILRISEIGYGYKTLEENEQDLVNCIR
DQNALYRLLDVTKSLVSYPMMVQFMVIGINIAITLFLIFYVETLYDRIYYLCFLLGI
50 TVQTYPLCYYGTMVQESFAELHYAVFCSNWVDQSASRGHMLILAERTKRMQLLLAGN
LVPIHLSTYVACWKGAYSFFTLMADRDGLGS

DOR64nt

GGCACGAGCCAAGAATTCAAAATGAAACTCAGCGAAACCCTAAAAATCGACTATTTTC
60 GAGTCCAGTTGAATGCCTGGCGAATTTGTGGTGCTTGGATCTCAGCGAGGGTAGGTA

ES 2 315 227 T3

CTGGAGTTGGTCGATGCTATTGTGCATCTTGGTGTACCTGCCGACACCCATGCTACTG
AGAGGAGTATACAGTTTCGAGGATCCGGTGGAAAATAATTTAGCTTGAGCCTGACGG
5 TCACATCGCTGTCCAATCTCATGAAGTTCTGCATGTACGTGGCCCACTAACAAAGAT
GGTCGAGGTCCAGAGTCTTATTGGTCAGCTGGATGCCCGGGTTTCTGGCGAGAGCCAG
TCTGAGCGTCATAGAAATATGACCGAGCACCTGCTAAGGATGTCCAAGCTGTTCCAGA
10 TCACCTACGCTGTAGTCTTCATCATTGCTGCAGTTCCTTCGTTTTTCGAAACTGAGCT
AAGCTTACCCATGCCCATGTGGTTTTCCCTTCGACTGGAAGAACTCGATGGTGGCCTAC
ATCGGAGCTCTGGTTTTCCAGGAGATTGGCTATGTCTTTCAAATTATGCAATGCTTTG
15 CAGCTGACTCGTTTTCCCCCGCTCGTACTGTACCTGATCTCCGAGCAATGTCAATTGCT
GATCCTGAGAATCTCTGAAATCGGATATGGTTACAAGACTCTGGAGGAGAACGAACAG
GATCTGGTCAACTGCATCAGGGATCAAAACGCGCTGTATAGATTACTCGATGTGACCA
20 AGAGTCTCGTTTTCGTATCCCATGATGGTGCAGTTTATGGTTATTGGCATCAACATCGC
CATCACCTATTTGTCCTGATATTTTACGTGGAGACCTTGTACGATCGCATCTATTAT
CTTTGCTTTCTCTTGGGCATCACCGTGCAGACATATCCATTGTGCTACTATGGAACCA
25 TGGTGCAGGAGAGTTTTGCTGAGCTTCACTATGCGGTATTCTGCAGCAACTGGGTGGA
TCAAAGTGCCAGCTATCGTGGGCACATGCTCATCCTGGCGGAGCGCACTAAGCGGATG
CAGCTTCTCCTCGCCGGCAACCTGGTGGCCATCCACCTGAGCACCTACGTGGCCTGTT
30 GGAAGGGAGCCTACTCCTTCTTCACCCTGATGGCCGATCGAGATGGCCTGGGTCTTA
GTAGCCCAGTCATTTCACTCACATTCTACATCAAGTAGTACTACCACTGAACACGAAC
ACGAATATTTCAAAAGTAAACACATAATATTACAATAGTGTATCACTTTAATAAAAT
35 TTTTGGTTACCATGAAAAAAAAAAAAAAAAAAAA

DOR67

40 MLSQFFPHIKEKPLSERVKS RDAFVYLDRVMWSFGWTVPENKRWDLHYKLWSTFVTLV
IFILLPISVSVEYIQRKTF SAGEFLSSIQIGVNMYGSSFKS YLTMGYKKRQEAKMS
LDELDRKRCVCDEERTIVHRHVALGNFCYIFYHIA YTSFLISNFLSFIMKRIHAWRMYF
45 PYVDPEKQFYISSIAE VILRGWAVFMDLCTDVCPLISMVIARCHITLLKQRLRLNRSE
PGRTEDEYLKELADCVRD HRLILDYVDALRSVFSGTIFVQFL LIGIVLGLS MINIMFF
STLSTGVAVVLFMSCVSMQTFPFCYLCNMIMDDCQEMADSLFQSDWTSADRRYKSTLV
50 YFLHNLQQPIILTAGGVFPISMQTNLMNVKLAFTVVTIVKQFNLA EK FQ

DOR67nt

55 GGCACGAGGAAATGTTAAGCCAGTTCCTTTCCCCACATTAAAGAAAAGCCATTGAGCGA
GCGGGTTAAGTCCCGAGATGCCTTCGTTTACTTAGATCGGGTGATGTGGTCCTTTGGC
60 TGGACAGTGCCTGAAAACAAAAGGTGGGATCTACATTACAAACTGTGGTCAACTTTG
TGACATTGGTGATATTTATCCTTCTGCCGATATCGGTAAGCGTTGAGTATATTCAGCG
GTTCAAGACCTTCTCGGCGGGTGAGTTTCTTAGCTCAATCCAGATTGGCGTTAACATG
65 TACGGAAGCAGCTTTAAAGTTATTTGACCATGATGGGATATAAGAAGAGACAGGAGG

CTAAGATGTCACCTGGATGAGCTGGACAAGAGATGCGTTTGTGATGAGGAGAGGACCAT
 TGTACATCGACATGTCGCCCTGGGAACTTTTGTATATTTTCTATCACATTGCGTAC
 5 ACTAGCTTTTGTATTCAAACCTTTTGTCAATTTATAATGAAGAGAATCCATGCCTGGC
 GCATGTACTTTCCCTACGTGACCCCGAAAAGCAATTTTACATCTCTAGCATCGCCGA
 AGTCATTCTTAGGGGGTGGGCCGTCTTCATGGATCTCTGCACGGATGTGTGTCCTTTG
 10 ATCTCCATGGTAATAGCACGATGCCACATCACCTTCTGAAACAGCGCCTGCGAAATC
 TACGATCGGAACCAGGAAGGACGGAAGATGAGTACTTGAAGGAGCTCGCCGACTGCGT
 TCGAGATCACCGCTTGATATTGGACTATGTGACGCATTGCGATCCGTCTTTTCGGGG
 15 ACAATTTTGTGAGTTCTCTTGATCGGTATTGTACTGGGTCTGTCAATGATAAATA
 TAATGTTTTTCTCAACACTTTTCGACTGGTGTGCGCGTTGTCCTTTTTATGTCCTGCGT
 ATCTATGCAGACGTTCCCTTTTGTCTATTGTGTAAACATGATTATGGATGACTGCCAA
 20 GAGATGGCCGACTCCCTTTTCAATCGGACTGGACATCTGCCGATCGTCGCTACAAAT
 CCACTTTGGTATACTTTCTTCACAATCTTCAGCAGCCCATTTCTTACGGCTGGTGG
 AGTCTTTCCTATTTCCATGCAAACAAATTTAAATATGGTGAAGCTGGCCTTTACTGTG
 25 GTTACAATAGTGAAACAATTTAACTTGGCAGAAAAGTTTCAATAAGTTAAGATATGCA
 AGCTCTGCTATTATAAACCTACACTCGAGAAAATATTTCTTCACATTAATAAACCTTC
 AGTACTTACTGCTTGTGGCGCCCCCGGAAAAAAAAAAAAAAAAAAAAA

DOR68

MSKLIIEVFLGNLWTQRFTFARMGLDLQPDKKGNVLRSPLLYICIMCLTTSFELCTVCAF
 35 MVQNRNQIVLCSEALMHGLQMVSSLLKMAIFLAKSHDLVDLIQQIQSPFTEEDLVGTE
 WRSQNQRGQLMAAIYFMMCAGTSVSFLLMPVALTMLKYHSTGEFAPVSSFRVLLPYDV
 TQPHVYAMDCCLMVFLSFFCCSTTGVDTLYGWCALGVSLQYRRLGQQLKRIPSCFNP
 40 SRSDFGLSGIFVEHARLLKIVQHFNYSFMEIAFVEVVIICGLYCSVICQYIMPHTNQ
 FAFLGFFSLVVTTLQCIYLFGAEQVRLEAERFSRLLYEVIPWQNLPPKHRKLFLFPIE
 RAQRETVLGAYFFELGRPLLWVVSIFLFIVLLF

DOR68nt

ATGTCAAAGCTAATCGAGGTGTTTCTGGGTAATCTGTGGACCCAGCGTTTTACCTTCG
 50 CCCGAATGGGTTTGGATTTGCAGCCCGATAAAAAGGGCAATGTTTTCGATCTCCGCT
 TCTTTATTGTATTATGTGTCTGACAACAAGCTTTGAGCTCTGCACCGTGTGCGCCTTT
 55 ATGGTCCAAAATCGCAACCAAATCGTGCTTTGTTCGAGGCCCTGATGCACGGACTAC
 AGATGGTCTCCTCGCTACTGAAGATGGCTATATTCTTGGCCAAATCTCACGACCTGGT
 GGACCTAATTCAACAGATTGAGTCGCCTTTTACAGAGGAGGATCTTGTAGGTACAGAG
 60 TGGAGATCCCAAATCAAAGGGGACAATAATGGCTGCCATTTACTTTATGATGTGTG
 CCGGTACGAGTGTGTCATTTCTGTTGATGCCAGTGGCTTTGACCATGCTTAAGTACCA
 TTCCACTGGGGAATTCGCGCCTGTCAGCTCGTTCCGGGTTCTGCTTCCATACGATGTG
 65 ACACAACCGCATGTTTATGCCATGGACTGCTGCTTGATGGTATTTGTGTTAAGTTTTT

ES 2 315 227 T3

TTTGCTGCTCCACCACCGGAGTGGATACCTTATATGGATGGTGTGCTTTAGGCGTGAG
TTTACAATACCGTCGCCTCGGTCAACAACTTAAAAGGATACCCTCCTGTTTCAATCCA
5 TCTCGGTCTGACTTTGGATTAAGTGGGATTTTTGTGGAGCATGCTCGTCTGCTTAAAA
TAGTCCAACATTTTAATTATAGTTTTATGGAGATCGCATTTGTGGAGGTTGTTATAAT
CTGTGGACTCTATTGCTCAGTAATTTGTCAGTATATAATGCCACACACCAACCAAAAC
10 TTCGCCTTTCTGGGTTTCTTTTCATTGGTAGTTACCACACAGCTGTGCATCTATCTTT
TCGGTGCCGAACAGGTCCGTTTGGAGGCTGAGCGATTTTCCCGGCTGCTATACGAAGT
AATTCCTTGCCAAAACCTTCCTCCTAAACACCGGAACTTTTCTTTTCCAATTGAG
15 CGCGCCCAACGAGAACTGTTCTCGGTGCTTATTTCTTCGAACTAGGCAGACCTCTTC
TTGTTTGGGTAAGCATATTCCTTTTTATTGTATTATTATT

DOR71

MVI IDSLSFYRPFWICMRLLVPTFFKDSSRPVQLYVLLHILVTLWFPLHLLLHLLLL
PSTAEFFKNLTMSLTCVACSLKHVAHLYHLPQIVEIESLIEQLDTFIASEQEHRYRD
25 HVHCHARRFTRCLYISFGMIYALFLGVFVQVISGNWELLYPAYFPFDLESNRFLGAV
ALGYQVFSMLVEGFQGLGNDTYTPLTLCLLAGHVHLWSIRMGQLGYFDDETUVNHQRL
LDYIEQHKL LVRFHNLVSRTISEVQLVQLGGCGATLCIIVSYMLFFVGDTISLVYYLV
30 FFGVVCVQLFPSCYFASEVAEELERLPYAIFSSRWYDQSRDHRFDLLIFTQLTLGNRG
WIIKAGGLIELNLNAFFATLKMAYSLFAVVHRETGNPLQREH

DOR71nt

ATGGTCATTATCGACAGTCTTAGTTTTTATCGTCCATTCTGGATCTGCATGCGATTGC
TGGTACCGACTTTCTTCAAGGATTCCTCACGTCTGTCCAGCTGTACGTGGTGTTGCT
40 GCACATCCTGGTCACCTTGTGGTTTCCACTGCATCTGCTGCTGCATCTTCTGCTACTT
CCATCTACCGCTGAGTTCTTTAAGAACCTGACCATGTCTCTGACTTGTGTGGCCTGCA
GTCTGAAGCATGTGGCCCACTTGTATCACTTGCCGCAGATTGTGGAAATCGAATCACT
45 GATCGAGCAATTAGACACATTTATTGCCAGCGAACAGGAGCATCGTTACTATCGGGAT
CACGTACATTGCCATGCTAGGCGCTTTACAAGATGTCTCTATATTAGCTTTGGCATGA
TCTATGCGCTTTTCTGTTCTGGCGTCTTCGTTCAAGTTATTAGCGGAAATTGGGAACT
50 TCTCTATCCAGCCTATTTCCCATTCGACTTGAGAGCAATCGCTTTCTCGGCGCAGTA
GCCTTGGGCTATCAGGTATTCAGCATGTTAGTTGAAGGCTTCCAGGGGCTGGGCAACG
ATACCTATACCCCACTGACCCTATGCCTTCTGGCCGGACATGTCCATTTGTGGTCCAT
55 ACGAATGGGTCAACTGGGATACTTCGATGACGAGACGGTGGTGAATCATCAGCGTTTG
CTGGATTACATTGAGCAGCATAAACTCTTGGTGCGGTTCCACAACCTGGTGAGCCGGA
CCATCAGCGAAGTGCAACTGGTGACAGCTGGGCGGATGTGGAGCCACTCTGTGCATCAT
60 TGTCTCCTACATGCTCTTCTTTGTGGGCGACACAATCTCGCTGGTCTACTACTTGGTG
TTCTTTGGAGTGGTCTGCGTGCAGCTCTTCCCAGCTGCTATTTTGCCAGCGAAGTAG
CCGAGGAGTTGGAACGGCTGCCATATGCGATCTTCTCCAGCAGATGGTACGATCAATC
65

ES 2 315 227 T3

GCGGGATCATCGATTGATTTGCTCATCTTTACACAATTAACACTGGGAAACCGGGGG
TGGATCATCAAGGCAGGAGGTCTTATCGAGCTGAATTTGAATGCCTTTTTTCGCCACCC
5 TGAAGATGGCCTATTCCTTTTTTGCAGTTGTGGTGCGGGCAAAGGGTATA

DOR72

MDLKPRVIRSEDIYRTYWLYWHLLGLESNFFLNRLLDLVITIFVTIWYPIHLILGLFM
ERSLGDVCKGLPITAACFFASFKFICFRFKLSEIKEIEILFKELDQRALSREECEFFN
QNTREANFIWKSFIVAYGLSNISAIASVLFGGGHKLLYPAWFPYDVQATELI FWLSV
15 TYQIAGVSLAILQNLANDSYPPMTFCVVAGHVRLLAMRLSRIGQGPEETIYLTGKQLI
ESIEDHRKLMKIVELLRSTMNISQLGQFISSGVNISITLVNILFFADNNFAITYYGVY
FLSMVLELFPCCYYGT LISVEMNQLTYAIYSSNWMSMNRYSRILLIFMQLTLAEVQI
20 KAGGMIGIGMNAFFATVRLAYSFFTLAMSLR

DOR72nt

ATGGACTTAAAACCGCGAGTCATTGGAAGTGAAGATATCTACAGAACCTATTGGTTAT
ATTGGCATCTTTTGGGCCTGGAAAGCAATTTCTTTCTGAATCGCTTGTTGGATTGTTGGT
GATTACAATTTTCGTAACCATTGTTGGTATCCAATTCACCTGATTCTGGGACTGTTTATG
30 GAAAGATCTTTGGGGGATGTCTGCAAGGGTCTACCAATTACGGCAGCATGCTTTTTTCG
CCAGCTTTAAATTTATTTGTTTTTCGCTTCAAGCTATCTGAAATTAAGAAATCGAAAT
ATTATTTAAAGAGCTGGATCAGCGAGCTTTAAGTCGAGAGGAATGCGAGTTTTTCAAT
35 CAAAATACGAGACGTGAGGCGAATTTTCATTTGGAAAAGTTTCATTGTGGCCTATGGAC
TGTCGAATATCTCGGCTATTGCATCAGTTCTTTTCGGCGGTGGACATAAGCTATTATA
TCCCGCCTGGTTTCCATACGATGTGCAGGCCACGGAATAATTTTTGGCTAAGTGTA
ACATACCAAATTGCCGGAGTAAGTTTGGCCATACTTCAGAATTTGGCCAATGATTCCT
40 ATCCACCGATGACATTTTGCCTGGTTGCCGGTCATGTAAGACTTTTGGCGATGCGCTT
GAGTAGAATTGGCCAAGGTCCAGAGGAAACAATATACTTAACCGGAAAGCAATTAATC
GAAAGCATCGAGGATCACCGAAAATAATGAAGATAGTGAATTACTGCGCAGCACCA
45 TGAATATTTTCGAGCTCGGCCAGTTTATTTCAAGTGGTGTTAATATTTCCATAACACT
AGTCAACATTCTCTTCTTTGCGGATAATAATTTGCTATAACCTACTACGGAGTGATC
TTCTATCGATGGTGTTGGAATTATTTCCCGTGCTGCTATTACGGCACCTGATATCCG
50 TGGAGATGAACCAGCTGACCTATGCGATTTACTCAAGTAACTGGATGAGTATGAATCG
GAGCTACAGCCGCATCCTACTGATCTTCATGCAACTCACCTGGCGGAAGTGCAGATC
AAGGCCGGTGGGATGATTGGCATCGGAATGAACGCCTTCTTTGCCACCGTGCGATTGG
55 CCTACTCCTTCTTCACTTTGGCCATGTCGCTGCGT

DOR73

ES 2 315 227 T3

MDSRRKVRSENLKYTYWLYWRLLGVEGDYPFRRRLVDFTITSFITILFPVHLILGMYKK
PQIQVFRSLHFTSECLFCSYKFFCFRWKLKEIKTIEGLLQDLDSRVESEEERNYFNQN
5 PSRVARMLSKSYLVAAISAIITATVAGLFSTGRNLMYLGWFPYDFQATAAIYWISFSY
QAIGSSLLILENLANDSYPPITFCVVSGHVRLIMRLSRIGHDVKLSSSENTRKLI EG
IQDHRKLMKIIRLLRSTLHLSQLGQFLSSGINISITLINILFFAENNFAMLYYAVFFA
10 AMLIELFPSCYYGILMTMEFDKLPYAI FSSNWLKMDKRYNRSIIILMQLTLVPVNIKA
GGIVGIDMSAFFATVRMAYSFYTLALSFRV

DOR73nt

ATGGATTCAAGAAGGAAAGTCCGAAGTGAAAATCTTTACAAAACCTATTGGCTTTACT
GGCGACTTCTGGGAGTCGAGGGCGATTATCCTTTTCGACGGCTAGTGGATTTTACAAT
20 CACGTCTTTTATTACGATTTTATTTCCCGTGCATCTTATACTGGGAATGTATAAAAAG
CCCCAGATTCAAGTCTTCAGGAGTCTGCATTTTACATCGGAATGCCTTTTCTGCAGCT
ATAAGTTTTTCTGTTTTCTGTTGGAACTTAAAGAAATAAGACCATCGAAGGATTGCT
25 CCAGGATCTCGATAGTCGAGTTGAAAGTGAAGAAGAACGCAACTACTTTAATCAAAAT
CCAAGTCGTGTGGCTCGAATGCTTTGAAAAGTTACTTGGTAGCTGCTATATCGGCCA
TAATCACTGCAACTGTAGCTGGTTTATTTAGTACTGGTCGAAATTTAATGTATCTGGG
30 TTGGTTTCCCTACGATTTTCAAGCAACCGCCGCAATCTATTGGATTAGTTTTTCTAT
CAGGCGATTGGCTCTAGTCTGTTGATTCTGGAAAATCTGGCCAACGATTCATATCCGC
CGATTACATTTTGTGTGGTCTCTGGACATGTGAGACTATTGATAATGCGTTTAAAGTCG
35 AATTGGTCACGATGTAAAATTATCAAGTTCGGAAAATACCAGAAAACCTCATCGAAGGT
ATCCAGGATCACAGGAACTAATGAAGATAATACGCCTACTTCGCAGCACTTTACATC
TTAGCCAACTGGGCCAGTTCCTTTCTAGTGAATCAACATTTCCATAACACTCATCAA
40 CATCCTGTTCTTTGCGGAAAACAACTTTGCAATGCTTTATTATGCGGTGTTCTTTGCT
GCAATGTTAATAGAACTATTTCCAAGTTGTTACTATGGAATTCTGATGACAATGGAGT
TTGATAAGCTACCATATGCCATCTTCTCCAGCAACTGGCTTAAATGGATAAAAAGATA
45 CAATCGATCCTTGATAATTCTGATGCAACTAACACTGGTTCAGTGAATATAAAAGCA
GGTGGTATTGTTGGCATCGATATGAGTGCATTTTTTGGCCACAGTTCGGATGGCATATT
CCTTTTACACTTTAGCCTTGTCATTTTCGAGTA

DOR77

MELMRVPVQFYRTIGEDIYHRSTNPLKSLLFKIYLYAGFINFNLLVIGELVFFYNSI
QDFETIRLAIAVAPCIGFSLVADFKQAAMIRGKKTLMMLDDLENMHPKTLAKQMEYK
LPDFEKTMKRVINIFTFLCLAYTTTTFSFYPAIKASVKFNFLGYDTFDRNFGFLIWFPF
60 DATRNNLIYWIMYWDIAHGAYLAAFQVTESTVEVIIICYIFLMTSMVQVFMVCYYGDT
LIAASLKVGDAAYNQKWFQCSKSYCTMLKLLIMRSQKPASIRPPTFPPI SLVTYMKNP
FNNLPKHSSSLQINANRYI

DOR77nt

5 ATGGAATTGATGCGAGTGCCAGTACAGTTTTACAGAACGATTGGAGAGGATATCTACG
 CCCATCGATCCACGAATCCCCTAAAATCGCTTCTCTTCAAGATCTATCTATATGCGGG
 ATTCATAAATTTTAATCTGTTGGTAATCGGTGAACTGGTGTCTTCTACAACCTCAATT
 CAGGACTTTGAAACCATTTCGATTGGCCATCGCGGTGGCTCCATGTATCGGATTTTCTC
 10 TGGTTGCTGATTTTAAACAAGCTGCCATGATTAGAGGCAAGAAAACACTAATTATGCT
 ACTCGATGATTTGGAGAACATGCATCCGAAAACCCTGGCAAAGCAAATGGAATACAAA
 TTGCCGGACTTTGAAAAGACCATGAAACGTGTGATCAATATATTCACCTTTCTCTGCT
 15 TGGCCTATACGACTACGTTCTCCTTTTATCCGGCCATCAAGGCATCCGTGAAATTTAA
 TTTCTTGGGCTACGACACCTTTGATCGAAATTTTGGTTTCCTCATCTGGTTTCCCTTC
 GATGCAACAAGGAATAATTTGATATACTGGATCATGTACTGGGACATAGCCCATGGGG
 20 CCTATCTAGCGGCCTTTTCAAGTCACCGAATCAACAGTGGAAGTGATTATTATTTACTG
 CATTTTTTTTGATGACCTCGATGGTTCAGGTATTTATGGTGTGCTACTATGGGGATACT
 TTAATTGCCGCGAGCTTGAAAGTGGGCGATGCCGCTTACAACCAAAGTGGTTTCAGT
 25 GCAGCAAATCCTATTGCACCATGTTGAAGTTGCTAATCATGAGGAGTCAGAAACCAGC
 TTCAATAAGACCGCCGACTTTTCCCCCATATCCTTGGTTACCTATATGAAGAATCCC
 TTCAACAATCTACCCAAACACAGCTCTTCCCTGCAAATCAACGCCAATCGCTATATC
 30

DOR78

35 MKFMKYAVFFYTSVGIEPYTIDSRSKKASLWSHLLFWANVINLSVIVFGEILYLGVAY
 SDGKFIDAVTVLSYIGFVIVGMSKMFFIWWKKTDLSDLVKELEHIYPNGKAEEMYRL
 DRYLRSCSRISITYALLYSVLIWTFNLFSIMQFLVYEKLLKIRVVQQLPYLMYFPWN
 40 WHENWTTYVLLFCQNFAGHTSASGQISTDLLCAVATQVVMHFDYLARVVEKQVLDRD
 WSENSRFLAKTVQYHQRI LR LMDVLNDIFGIPLLLNFMVSTFVICFVGFQMTVGVPD
 IMIKLFLFLFSSLSQVYLICHYGQLIADAVRDFRSSSLISAYKQNWQNADIRYRRAL
 45 VFFIARPQRTTYLKATIFMNITRATMTDVRYNLKCH

DOR78nt

50 ATGAAGTTCATGAAGTACGCAGTTTTCTTTTACACATCGGTGGGCATTGAGCCGTATA
 CGATTGACTCGCGGTCCAAAAAGCGAGCCTATGGTCACATCTTCTCTTCTGGGCCAA
 TGTGATCAATTTAAGTGTCATTGTTTTCGGAGAGATCCTCTATCTGGGAGTGGCCTAT
 55 TCCGATGGAAAGTTCATTGATGCCGTCACTGTACTGTATATATCGGATTCGTAATCG
 TGGGCATGAGCAAGATGTTCTTCATATGGTGGAGAAGACCGATCTAAGCGATTTGGT
 TAAGGAATTGGAGCACATCTATCCAAATGGCAAAGCTGAGGAGGAGATGTATCGGTTG
 60 GATAGGTATCTGCGATCTTGTTACGAATTAGCATTACCTATGCACTACTCTACTCCG
 65

ES 2 315 227 T3

TACTCATCTGGACCTTCAATCTGTTCAAGTATCATGCAATTCCTTGTCTATGAAAAGTT
 GCTTAAAATCCGAGTGGTCCGCCAAACGCTGCCATATTTGATGTACTTTCCCTGGAAC
 5 TGGCATGAAAACCTGGACGTATTATGTGCTGCTGTTCTGTCAAACTTCGCAGGACATA
 CTTCGGCATCGGGACAGATCTCTACGGATCTTTTGCTTTGTGCTGTTGCTACCCAGGT
 GGTAATGCACTTCGATTACTTGGCCAGAGTGGTGGAAAAACAAGTGTAGATCGCGAT
 10 TGGAGCGAAAACCTCCAGATTTTTGGCAAAAACCTGTACAATATCATCAGCGCATTCTTC
 GGCTAATGGACGTTCTCAACGATATATTCGGGATACCGCTACTGCTTAACTTTATGGT
 CTCCACATTTGTCTATCTGCTTTGTGGGATTCCAAATGACCGTGGGTGTCCCGCCGGAC
 15 ATCATGATTAAGCTCTTCTTGTTCCTGTTCTCGTCCTTGTGCGAAGTGTACTTGATAT
 GCCACTACGGCCAGCTGATTGCCGATGCGGTAAGAGACTTTCGAAGCTCTAGCTTATC
 GATTTCTGCATATAAGCAGAATTGGCAAAATGCTGACATTGCTATCGTCGGGCTCTG
 20 GTATTCTTTATAGCTCGACCTCAGAGGACAACTTATCTAAAAGCTACAATTTTCATGA
 ATATAACAAGGGCCACCATGACGGACGTAAGATACAATTTGAAATGTCAT

DOR81

MMETLRNSGLNLKNDGIGRKIWRVFSFTYNMVILPVSFPINYVIHLAEFPPELLLQS
 LQLCLNTWCFALKFFTLIVYTHRLELANKHFDELDKYCVKPAEKRKVRDMVATITRLY
 30 LTFVVVYVLYATSTLLDGLLHHRVPYNTYYPFINWRVDRTOQMYIQSFLEYFTVGYAIY
 VATATDSYPVIYVAALRTHILLKDRIIYLGDPSESSDPSYMFKSLVDCIKAHRTM
 LNFCDAIQPIISGTIFAQFIICGSILGIIMINMVLFADQSTRFGIVYVMAVLLQTFP
 35 LCFYCNAIVDDCKELAHALFHSWWVDKRYQRTVIQFLQKLQQPMTFTAMNIFNINL
 ATNINVSPLLSVRTGKEAKSELQSLQVAKFAFTVYAIASGMNLDQKLSIKE

DOR81nt

ATGATGGAGACGCTGCGAAATTCGGGCTTGAATTTGAAGAACGATTTCCGGTATAGGCC
 45 GCAAGATTTGGAGGGTGTTCCTGTTACCTACAATATGGTGATACTTCCCGTAAGTTT
 CCCAATCAACTATGTGATACATCTGGCGGAGTTCCCGCCGGAGCTGCTGCTGCAATCC
 CTGCAACTGTGCCTCAACACTTGGTGCTTCGCTCTGAAGTTCTTCACTCTGATCGTCT
 50 ATACGCACCGCTTGGAGCTGGCCAACAAGCACTTTGACGAATTGGATAAGTACTGCGT
 GAAGCCGGCGGAGAAGCGCAAGGTTGCGGACATGGTGGCCACTATTACAAGACTGTAC
 CTGACCTTCGTGCTGGTCTACGTCCTCTACGCCACCTCCACGCTACTGGACGGACTAC
 55 TGCACCACCGTGTTCCCTACAATACGTAATCCTGTTTCATAAACTGGCGAGTCGATCG
 GACCCAGATGTACATCCAGAGTTTTCTGGAGTACTTCACCGTGGGTTATGCCATATAT
 GTGGCCACCGCCACCGATTCTACCTGTGATTTACGTGGCAGCCCTGCGAACTCATA
 60 TTCTCTTGCTCAAGGACCGTATCATTTACTTGGGCGATCCAGCAACGAGGGTAGCAG
 CGACCCGAGCTACATGTTTAAATCGTTGGTGGATTGTATCAAGGCACACAGAACCATG
 CTAAAGTGCAGTTTTTGTGATGCCATTCAACCAATCATCTCTGGCACGATATTTGCCC
 65 AATTCATCATATGCGGATCGATCCTGGGCATAATTATGATCAACATGGTATTGTTTCG

ES 2 315 227 T3

TGATCAATCGACCCGATTCGGCATAGTCATCTACGTTATGGCCGTCCTTCTGCAGACT
TTTCCGCTTTGCTTCTACTGCAACGCCATCGTGGACGACTGCAAAGAACTGGCCCACG
5 CACTTTTCCATTCCGCCTGGTGGGTGCAGGACAAGCGATACCAGCGGACTGTCATCCA
GTTCTGTCAGAACTGCAGCAGCCCATGACCTTCACCGCCATGAACATATTTAACATT
AATTTGGCCACTAACATCAATGTAAGTCCACTGCTCTCGGTAGAACGGGGAAGGAAG
10 CAAAGTCCGAACTTCAATCCTTGCAGGTAGCCAAGTTCGCCTTCACCGTGTACGCCAT
CGCGAGCGGTATGAACCTGGACCAAAAGTTAAGCATTAAGGAA

DOR82

MACIPRYQWKGRPTERQFYASEQRIVFLLGTICQIFQITGVLIYWYCNGRLATETGT
VAQLSEMCSSFCLTFVGFCNVYAISTNRNQIETLLEELHQIYPRYRKNHYRCQHYFDM
20 AMTIMRIEFLFYMILYVYNSAPLWVLLWEHLHEEYDLSFKTQTNWFPWKVHGSALG
FGMAVLSITVGSFVGVGFSIVTQNLICLLTFQLKLHYDGISSQLVSLDCRRPGAHKEL
SILIAHHSRILQLGDQVNDIMNFVFGSSLVGATIAICMSSVSIMLLDLASAFKYASGL
25 VAFVLYNFVICYMGTEVTLAVKIGSYMDGRRWIPKDSLLRSQRLQVLVAVGFFNICVL
SNRRPKIEILLRYYYHIMFYSFKLYFSLRKGSLWKILSSFTLLRI

DOR82nt

ATGGCATGCATACCAAGATATCAATGGAAAGGACGCCCTACTGAAAGACAGTTCTACG
CTTCGGAGCAAAGGATAGTGTTCCTTCTTGGAAACCATTGCGCAGATATTCCAGATTAC
35 TGGAGTGCTTATCTATTGGTATTGCAATGGCCGTCCTTGCCACGGAAACGGGCACCTTT
GTGGCACAATTATCTGAAATGTGCAGTTCTTTTTGTCTAACATTTGTGGGATTCTGTA
ACGTTTATGCGATCTCTACAAACCGCAATCAAATTGAAACATTACTCGAGGAGCTTCA
40 TCAGATATATCCGAGATACAGGAAAAATCACTATCGCTGCCAGCATTATTTTGACATG
GCCATGACAATAATGAGAATTGAGTTTCTTTTCTATATGATCTTGTACGTGTACTACA
ATAGTGCACCATTTATGGGTGCTTCTTTGGGAACACTTGCACGAGGAATATGATCTTAG
45 CTTCAAGACGCAGACCAACACTTGGTTTCCATGGAAAGTCCATGGGTGGCAGTTGGA
TTTGGTATGGCTGTACTAAGCATAACCGTGGGATCCTTTGTGGGCGTAGGTTTCAGTA
TTGTCACCCAGAATCTTATCTGTTTGTAACTTCCAATAAGTTGCACTACGATGG
50 AATATCCAGTCAGTTAGTATCTCTCGATTGCCGTCGTCCTGGAGCTCATAAGGAGTTG
AGCATCCTCATCGCCCACCACAGCCGAATCCTTCAGCTGGGCGACCAAGTCAATGACA
TAATGAACTTTGTATTTCGGCTCTAGCCTAGTAGGTGCCACTATTGCCATTTGTATGTC
55 AAGTGTCTTCTATAATGCTACTGGACTTAGCATCTGCCTTCAAATATGCCAGTGGTCTA
GTGGCATTTCGTCTCTACAACTTTGTTCATCTGCTACATGGGAACCGAGGTCACTTTAG
CTGTGAAGATTGGTTCATATATGGACGGAAGGCGGTGGATACCCAAAGATTGCTTGCT
60 GAGATCTCAGAGGCTACAGGTGCTCGTCGCAGTTGGATTTTTTAATATATGTGTCTC
TCGAATCGTCGTCTTAAATTTGAAATTTTGCTTAGATATTATTACCATATTATGTTTT
ATTCATTTAAATTATATTTTTCTTTAAGGAAAGGTAGCCTTTGGAAAATCTTGTCTTC
65 TTTCACCTTATTGAGGATC

DOR83

MQLEDFMRYPDLVCQAAQLPRYTWNRRSLEVKRNLAKR IIFWLGA VNLVYHNIGCVM
 YGYFGDGR TKDPIAYLAELASVASMLGFTIVGTLNLWKMLSLKTHFENLLNEFEELFQ
 LIKHAYRIHHYQEKYTRHIRNTFIFHTSAVVYNSLPILLMIREHFSNSQQLGYRIQ
 SNTWYPWQVQGSIPGFFAAVACQIFSCQTNMCVNMFIQFLINFFGIQLEIHFDGLARQ
 LETIDARNPHAKDQLKYLIVYHTKLLNLADRVNRSFNFTFLISLSVSMISNCFLAFSM
 TMFDFGTS LKHL LGLLLFITYNFSMCRSGTHLILTSGKVLPAAFYNNWYEGDLVYRRM
 LLILMMRATKP YMWKTYKLAPVSITTYMAECKTKEAHEQRHFR RHERQKPRVARI

DOR83nt

ATGCAGTTGGAGGACTTTATGCGGTACCCGGACCTCGTGTGTCAAGCGGCCCAACTTC
 CCAGATACACGTGGAATGGCAGACGATCCTTGGGAAGTTAAACGCAACTTGGCAAAACG
 CATTATCTTCTGGCTTGGAGCAGTAAATTTGGTTTATCACAATATTGGCTGCGTCATG
 TATGGCTATTTCCGTGATGGAAGAACAAAGGATCCAATTGCGTATTTAGCTGAATTGG
 CATCTGTGGCCAGCATGCTTGGTTTACCATTGTGGGCACCCTCAACTTGTGGAAGAT
 GCTGAGCCTTAAGACCCATTTTGAGAACCTACTAAATGAATTCGAGGAATTATTTCAA
 CTAATCAAGCACAGGGCGTATCGCATACACCCTATCAAGAAAAGTATACGCGTCATA
 TACGAAATACATTTATTTTCCATACCTCTGCCGTTGTCTACTACAACCTACTACCAAT
 TCTTCTAATGATTCGGGAACATTTCTCGAACTCACAGCAGTTGGGCTATAGAATTCAG
 AGTAATACCTGGTATCCCTGGCAGGTTCAAGGATCAATTCCTGGATTTTTTGCTGCAG
 TCGCCTGTCAAATCTTTTCGTGCCAAACCAATATGTGCGTCAATATGTTTATCCAGTT
 TCTGATCAACTTTTTTGGTATCCAGCTAGAAATACACTTCGATGGTTTGGCCAGGCAG
 CTGGAGACCATCGATGCCCGCAATCCCCTATGCCAAGGATCAATTGAAGTATCTGATTG
 TATATCACACAAAATTGCTTAATCTAGCCGACAGAGTTAATCGATCGTTTAACTTTAC
 GTTCTCTATAAGTCTGTCCGTATCCATGATATCCAACCTGTTTCTGGCATTTCATG
 ACCATGTTTCGACTTTGGCACCTCTCTAAAACATTTACTCGGACTTTTGCTATTTCATCA
 CATATAATTTTTCAATGTGCCGCAGTGGTACGCACTTGATTTTAACGAGTGGCAAAGT
 ATTGCCAGCGGCCTTTTATAACAATTGGTATGAAGGCGATCTTGTTTATCGAAGGATG
 CTCCTCATCCTGATGATGCGTGCTACGAAACCTTATATGTGGAAAACCTACAAGCTGG
 CACCTGTATCCATAACTACATATATGGCAGAATGCAAAAACAAAAGAAGCCCATGAACA
 ACGCCATTTTAGACGCCATGAAAGACAAAACCTCGGGTTGCACGAATA

DOR84

MVFSFYAEVATLVDR LRDNENFLESCILLSYVSFVVMGLSKIGAVMKKKPKMTALVRQ
 LETCFPSPSAKVQEEYAVKSWLKRCHIYTKGFGGLFMIMYFAHALIPLFIYFIQRVLL
 HYPDAKQIMPFYQLEPWEFRDSWLFYPSYFHQSSAGYTATCGSIAGDLMIFAVVLQVI
 MHYERLAKVLREFKIQAHNAPNGAKEDIRKLQSLVANHIDILRLTDLMNVEVFGIPLLL
 NFIASALLVCLVGVQLTIALSPEYFCKQMLFLISVLLLEVYLLCSFSQRLIDAVC

DOR84nt

ES 2 315 227 T3

ATGGTGTTTAGTTTTATGCCGAGGTAGCGACTCTGGTGGACAGGTTACGCGATAATG
AAAATTTTCTCGAGAGCTGCATCTTACTGAGCTACGTGTCCTTTGTGGTCATGGGCCT
5 CTCCAAGATAGGTGCTGTAATGAAAAAAGCCAAAAATGACAGCTTTGGTCAGGCAA
TTGGAGACCTGCTTTCCGTCGCCAAGTGCAAAGGTTCAAGAGGAATATGCTGTGAAGT
CCTGGCTGAAACGCTGCCATATATACACAAAGGGATTTGGTGGTCTCTTCATGATCAT
10 GTATTTTCGCTCACGCTCTGATTCCTTATTCATATACTTCATTCAAAGAGTGCTGCTC
CACTATCCGGATGCCAAGCAGATTATGCCGTTTTACCAACTCGAACCTTGGGAATTC
GCGACTCCTGGTTGTTTTATCCAAGCTATTTTACCAGTCGTCCGCCGATATACGGC
15 TACATGTGGATCCATTGCCGGTGACCTAATGATCTTCGCTGTGGTCCTGCAGGTCATC
ATGCACTACGAAAGACTGGCCAAGGTTCTTAGGGAGTTTAAGATTCAAGCCCATAACG
CACCCAATGGAGCTAAGGAGGATATAAGGAAGTTGCAGTCCCTAGTCGCCAATCACAT
20 TGATATACTTCGACTCACTGATCTGATGAACGAGGTCTTTGGAATTCCTTGTGCTA
AACTTTATTGCATCTGCGCTGCTGGTCTGCCTGGTGGGAGTTCAATTAACCATCGCTT
TAAGTCCAGAGTATTTTTGCAAGCAGATGCTATTTCTGATTTCCGTACTGCTTGAGGT
25 CTATCTCCTTTGCTCCTTCAGCCAGAGGTTAATAGATGCTGTATGT

DOR87

MTIEDIGLVGINVRMWRHLAVLYPTPGSSWRKFAFVLPVTAMNLMQFVYLLRMWGDLP
AFILNMFFFSAILNALMRTWLVIKRRQFEEFLGQLATLFHSILDSTDEWGRGILRRA
EREARNLAILNLSASFLDIVGALVSPLFREERAHPFGVALPGVSMTSSPVYEVYLAQ
35 LPTPLLLSMMYMPFVSLFAGLAIFGKAMLQILVHRLGQIGGEEQSEERFQRLASCIA
YHTQVMRYVWQLNKLIVANIVAVEAIFGSIICSLFLCLNIITSPTQVISIVMYILTML
YVLFYYNRANEICLENNRVAEAVYNVPWYEAGTRFRKTLIFLMQTQHPMEIRVGNV
40 YPMTLAMFQSLNLSYFTMLRGVTGK

DOR87nt

GGCACGAGGCTTATAGAAAGTGCCGAGCAATGACAATCGAGGATATCGGCCTGGTGGG
CATCAACGTGCGGATGTGGCGACACTTGGCCGTGCTGTACCCCACTCCGGGCTCCAGC
50 TGGCGCAAGTTCGCCTTCGTGCTGCCGGTGAAGTCTGATGCAGTTCGTCT
ACCTGCTGCGGATGTGGGGCGACCTGCCCCGCTTCATTCTGAACATGTTCTTCTCTC
GGCCATTTTCAACGCCCTGATGCGCACGTGGCTGGTCATAATCAAGCGCGCCAGTTC
55 GAGGAGTTTCTCGGCCAACTGGCCACTCTGTTCCATTCTCGACTCCACCGACG

ES 2 315 227 T3

AGTGGGGGCGTGGCATCCTGCGGAGGGCGGAACGGGAGGCTCGGAACCTGGCCATCCT
TAATTTGAGTGCCTCCTTCCTGGACATTGTCTGGTGCTCTGGTATCGCCGCTTTTCAGG
5 GAGGAGAGAGCTCATCCCTTCGGCGTAGCTCTACCAGGAGTGAGCATGACCAGTTCAC
CCGTCTACGAGGTTATCTACTTGGCCCAACTGCCTACGCCCTGCTGCTGTCCATGAT
GTACATGCCTTTTCGTCAGCCTTTTGGCCGGCCTGGCCATCTTTGGGAAGGCCATGCTG
10 CAGATCCTGGTACACAGGCTGGGCCAGATTGGCGGAGAAGAGCAGTCGGAGGAGGAGC
GCTTCCAAAGGCTGGCCTCCTGCATTGCGTACCACACGCAGGTGATGCGCTATGTGTG
GCAGCTCAACAACTGGTGGCCAAACATTGTGGCGGTGGAAGCAATTATTTTTGGCTCG
15 ATAATCTGCTCACTGCTCTTCTGTCTGAATATTATAACCTCACCCACCCAGGTGATCT
CGATAGTGATGTACATTCTGACCATGCTGTACGTTCTCTTCACCTACTACAATCGGGC
CAATGAAATATGCCTCGAGAACAACCGGGTGGCGGAGGCTGTTTACAATGTGCCCTGG
20 TACGAGGCAGGAACCTCGGTTTCGCAAAACCCCTCCTGATCTTCTTGATGCAAAACACAAC
ACCCGATGGAGATAAGAGTCGGCAACGTTTACCCCATGACATTGGCCATGTTCCAGAG
TCTGTTGAATGCGTCTACTCCTACTTTACCATGCTGCGTGGCGTCACCGGCAAATGA
25 GCTGAAAGACCGAAAAACCGGAGTATCCCCTTCCATATTCCCCCTGCTCCTTTATTT
TCCTTTTCCTTTTCCCTTTCCGTTTCCCATTGCGTTTTCAGCAATCCGGGTAATGCA
AAAAGTTGTTGCTGGCTGTGGTCCTGGCTGCTTGTGTTGGCATTGTCATATGCTTGTGCG
30 TTTGAAAGGATTTAATCGGACTGCTGGCACGGAGTCGGCATCCTGGCTCCTGGATCCT
GGCATGCAAATAGTTGGCTTCTTAGATTGTTACACAAAATAGATTGTAGATTGCAGCT
GAATGTTGTGCTTGAATAAAGTCAAAGGATGTGGAGTCGGCCCAAGGCTCTGCCCA
35 TTCTGTTTGCTCGGGATGCCCCGAAAGTATGAAAAAAAAAAAAAAAAAAAAA

DOR91

MVRYVPRFADGQKVKLAWPLAVFRLNHI FWPLDPSTGKWGRYLDKVLAVAMSLVFMQH
40 NDAELRYLRFEASNRNLDAFLTGMPTYLILVEAQFRSLHILLHFEKLQKFLEIFYANI
YIDPRKEPEMFRKVDGKMI INRLVSAMYGAVISLYLIAPVFSI INQSKDFLYSMIFPF
DSDPLYIFVPLLLTNVWVGIVIDTMMFGETNLLCELIVHLNGSYMLLKRDQLAIEKI
45 LVARDRPHMAKQLKVLITKTLRKNVALNQFGQQLAQYTVRVFIMFAFAAGLLCALSF
KAYTTDSLSTMYYLTHWEQILQYSTNPSENLRLKLINLAIEMNSKPFYVTGLKYFRV
SLQAGLKRQKFLRSASSSTLSTADVLAFAFAFTRWLL
50

DOR91nt

ATGGTTCGTTACGTGCCCCGGTTCGCTGATGGTCAGAAAGTAAAGTTGGCTTGGCCCT
55 TGGCGGTTTTTTCGGTTAAATCACATATTCTGGCCATTGGATCCGAGCACAGGGAAATG
GGGCCGATATCTGGACAAGGTTCTAGCTGTTGCGATGTCCTTGGTTTTTATGCAACAC
60 AACGATGCAGAGCTGAGGTACTTGGCTTCGAGGCAAGTAATCGGAATTTGGATGCCT
TTCTCACAGGAATGCCAACGTATTTAATCCTCGTGGAGGCTCAATTTAGAAGTCTTCA
CATTCTACTGCACTTCGAGAAGCTTCAGAAGTTTTTAGAAATATTCTACGCAAATATT
65

ES 2 315 227 T3

TATATTGATCCCCGTAAGGAACCCGAAATGTTTCGAAAAGTGGATGGAAAGATGATAA
TTAACAGATTAGTTTCGGCCATGTACGGTGCAGTTATCTCTCTGTATCTAATCGCACC
5 CGTTTTTTCCATCATTAAACCAAAGCAAAGATTTTCTATACTCTATGATCTTTCCGTTT
GATTCCGATCCCTTGTACATATTTGTGCCACTGCTTTTGACAAACGTATGGGTTGGCA
TTGTAATAGATACCATGATGTTTCGGGGAGACGAATTTGTTGTGTGAACATAATTGTCCA
10 CCTAAATGGTAGTTATATGTTGCTCAAGAGGGACTTGCGAGTTGGCCATTGAAAAGATA
TTAGTTGCAAGGGACCGTCCGCATATGGCCAAACAGCTAAAGGTTTTAATTACAAAAA
CTCTCCGAAAGAATGTGGCTCTAAATCAGTTTGGCCAGCAGCTGGAGGCTCAGTATAC
15 TGTGCGGGTTTTTATTATGTTTGCATTCGCTGCGGGCCTTTTATGTGCTCTTTCTTTT
AAGGCTTATACGACGGATTCCCTCAGCACAATGTACTACCTTACCCATTGGGAGCAAA
TCCTGCAGTACTCTACAAATCCCAGCGAAAATCTGCGATTACTAAAGCTCATTAAGTT
20 GGCCATTGAGATGAACAGCAAGCCCTTCTATGTGACAGGGCTAAAATATTTTCGCGTT
AGTCTGCAGGCTGGCTTAAAACGTCAAAAGTTTCTGCGGTCTGCCAGCTCATCCACCC
TTAGCACCGCTGATGTGTTGGCATTGCTTTTGCTTTTACTCGCTGGCTGCTT

DOR92

MSEWLRFLKRDQQLDVYFFAVPRLSLDIMGYWPGKTGDTWPWRS LIHFAILAIGVATE
30 LHAGMCF LDRQQT LALET LCPAGTSAV TLLKMFLMLRFRQDLSIMWNRLRGLLFDPN
WERPEQRDIRLKHSAMAARINFWPLSAGFFTCTTYNLKPILIAMILYLQNR YEDFVWF
TPFNMTMPKVLNYPFFPLTYIFIA YTYGYVTIFMFGGCDGFYFEFCAHLSALFEVLQA
35 EIESMFRPYTDHLELSPVQLYILEQKMRSVIRHNAIIDLTRFFRDRYTIITLAHFVS
AAMVIGFSMVNLLTLGNNGLGAMLYVAYTVAALSQLLVYCYGGTLVAESSTGLCRAMF
SCPWQLFKPKQRRVLQLLILRSQRPVSMVAPFFSPSLATFAAILQTS GSIIALVKS FQ

DOR92nt

ATGTCCGAGTGGTTACGCTTTCTGAAACGCGATCAACAGCTGGATGTGTACTTTTTTG
CAGTGCCCCGCTTGAGTTTAGACATAATGGGCTATTGGCCGGGCAAACTGGTGATAC
ATGGCCCTGGAGATCCCTGATTCACTTCGCAATCCTGGCCATTGGCGTGGCCACCGAA
50 CTGCATGCTGGCATGTGTTTTCTAGACCGACAGCAGATTACCTTGGCACTGGAGACCC
TCTGTCCAGCTGGCACATCGGCGGTACGCTGCTCAAGATGTTCTTAATGCTGCGCTT
TCGTCAGGATCTCTCCATTATGTGGAACCGCCTGAGGGGCCTGCTCTTCGATCCCAAC
TGGGAGCGACCCGAGCAGCGGGACATCCGGCTAAAGCACTCGGCCATGGCGGCTCGCA
55 TCAATTTCTGGCCCCCTGTCAGCCGGATTCTTCACATGCACCACCTACAACCTAAAGCC
GATACTGATCGCAATGATATTGTATCTCCAGAATCGTTACGAGGACTTCGTTTGGTTT
ACACCCTTCAATATGACTATGCCCAAAGTTCTGCTAAACTATCCATTTTTTCCCCTGA
60 CCTACATATTTATTGCCTATACGGGCTATGTGACCATCTTTATGTTTCGGCGGCTGTGA
TGGTTTTTATTTTCGAGTTCTGTGCCACCTATCAGCTCTTTTCGAAGTGCTCCAGGCG
GAGATAGAATCAATGTTTAGACCCTACACTGATCACTTGGAAGTGTGCCAGTGCAGC
65

ES 2 315 227 T3

TTTACATTTT TAGAGCAAAGATGCGATCAGTAATCATTAGGCACAATGCCATCATCGA
TTTGACCAGATTTTTTCGTGATCGCTATAACCATTATTACCCTGGCCCATTTTGTGTCC
5 GCCGCCATGGTGATTGGATT CAGCATGGTTAATCTCCTGACATTGGGCAATAATGGTC
TGGGCGCAATGCTCTATGTGGCCTACACGGTTGCCGCTTTGAGCCAACCTGCTGGTTTA
TTGCTATGGCGGAACCTCTGGTGGCCGAAAGTAGCACTGGTCTGTGCCGAGCCATGTTT
10 TCCTGTCCGTGGCAGCTTTTTTAAGCCTAAACAACGTCGACTCGTTTCAGCTTTTGATT
TCAGATCGCAGCGTCCTGTTTCCATGGCAGTGCCATTCTTTTCGCCATCGTTGGCTAC
CTTTGCTGCGATTCTTCAAACCTCGGGTTCATAATTGCGCTGGTTAAGTCCTTTTCAG

DOR95

MSDKVKGKKQEEKDQSLRVQILVYRCMGIDLWSPTMANDRPWLTFVTMGPLFLFMVPM
20 FLAAHEYITQVSLSDTLGSTFASMLTLVKFLLFCYHRKEFVGLIYHIRAILAKEIEV
WPDAREIIIEVENQSDQMLSLTYTRCFGLAGIFAALKPFVGIILSSIRGDEIHLELPHN
GVYPYDLQVVMFYVPTYLWNVMASYSAVTMALCVDLSLFFFTYNVCAIFKIAKHRMIH
25 LPAVGGKEELEGLVQVLLHQLQIADHIADKYRPLIFLQFFLSALQICFIGFQVAD
LFPNPQSLFYIAFVGSLLIALFIYSKGENIKSASLDFGNGLYETNWTDFSPTKRAL
LIAAMRAQRPCQMKGYFFEASMATFSTIVRSVSYIMMLRSFNA

DOR95nt

ATGAGCGACAAGGTGAAGGGAAAAAAGCAGGAGGAAAAGGATCAATCCTTGCGGGTG CAAATTC
35 CCAGCTATAGTGCTGTAACCATGGCACTCTGCGTGGACTCGCTGCTCTTCTTTTTCAC
CTACAACGTGTGCGCCATTTTCAAGATCGCCAAGCACCGGATGATCCATCTGCCGGCG
GTGGGCGGAAAGGAGGAGCTGGAGGGGCTCGTCCAGGTGCTGCTGCTGCACCAGAAGG
40 GCCTCCAGATCGCCGATCACATTGCGGACAAGTACCGGCCGCTGATCTTTTTCAGTT
CTTTCTGTCCGCCTTGAGATCTGCTTCATTGGATTCCAGGTGGCTGATCTGTTTCCC
AATCCGAGAGTCTCTACTTTATCGCCTTTGTGGGCTCGCTGCTCATCGCACTGTTCA
TCTACTCGAAGTGCGGCGAAAATATCAAGAGTGCCAGCCTGGATTCGGAACGGGCT
45 GTACGAGACCAACTGGACCGACTTCTCGCCACCCACTAAAAGAGCCCTCCTCATTGCC
GCCATGCGCGCCAGCGACCTTGCCAGATGAAGGGCTACTTTTTCGAGGCCAGCATGG
CCACCTTCTCGACGATTGTTGCTCTGCCGTGTCGTACATCATGATGTTGCGCTCCTT
TAATGCC

DOR99

MEEFLRPQMFQEVAQMVHFQWRRNPVDNSMVNASMVPFCLSAFLNVLFFGCNGWDIIG
55 HFWLGH PANQNPPVLSITIYFSIRGLMLYLKRKEIVEFVNDLDRECPRDLVSQ LDMQM
DETYRNFWRQRYRFIRIYSHLGGPMFCVVPLALFLLTHEGKDT PVAQHEQLLGGWLPCG
60 VRKDPNFYLLVWSFDLMCTTCGVSFVTFDNLFNVMQGHLMHGLHGLARQFS AIDPRQ
SLTDEKRFFVDLRLLVQRQQLNLGLCRKYNDIFKVAFLVSNFVGAGSLCFYLFMLSET
SDVLI I AQYI LPTLVLVGFTFEICLRGTQLEKASEGLESSLRSQEWYLGSRRYRK FYL
65 LWTQYCQRTQQLGAFGLIQVNMVHFTEIMQLAYRLFTFLKSH

DOR99nt

5 ATGGAGGAGTTTCTGCGTCCGCAGATGTTCCAGGAGGTGGCTCAGATGGTGCATTTCC
 AGTGGCGGAGAAATCCGGTGGACAACAGCATGGTGAACGCATCCATGGTCCCCTTCTG
 CTTGTGCGCGTTTCTTAATGTCCTGTTTTTTCGGCTGCAATGGTTGGGACATCATAGGA
 CATTTCCTGGCTGGGACATCCTGCCAACCAGAATCCGCCCCGTGCTTAGCATCACCATT
 10 ACTTCTCGATCAGGGGATTGATGCTATACCTGAAACGAAAGGAAATCGTTGAGTTTGT
 TAACGACTTGGATCGGGAGTGTCCGCGGGACTTGGTCAGCCAGTTGGACATGCAAATG
 GATGAGACGTACCGAAACTTTTGGCAGCGCTATCGCTTCATCCGTATCTACTCCCATT
 15 TGGGTGGTCCGATGTTCTGCGTTGTGCCATTAGCTCTATTCCCTCCTGACCCACGAGGG
 TAAAGATACTCCTGTTGCCCAGCAGCAGCTCCTTGGAGGATGGCTGCCATGCGGT
 GTGCGAAAGGACCCAAATTTCTACCTTTTAGTCTGGTCCTTCGACCTGATGTGCACCA
 20 CTTGCGGCGTCTCCTTTTTTCGTTACCTTCGACAACCTATTCAATGTGATGCAGGGACA
 TTTGGTCATGCATTTGGGCCATCTTGCTCGCCAGTTTTTCGGCCATCGATCCTCGACAG
 AGTTTGACCGATGAGAAGCGATTCTTTGTGGATCTTAGGTTATTAGTTCAGAGGCAGC
 25 AGCTTCTTAATGGATTGTGCAGAAAATACAACGACATCTTTAAAGTGGCCTTCCTGGT
 GAGCAATTTTGTAGGCGCCGGTTCCTCTGCTTCTACCTCTTTATGCTCTCGGAGACA
 TCAGATGTCCTTATCATCGCCCAGTATATATTACCCACTTTGGTCCTGGTGGGCTTCA
 30 CATTTGAGATTTGTCTACGGGGAACCCAACTGGAAAAGGCGTCGGAGGGACTGGAATC
 GTCGTTGCGAAGCCAGGAATGGTATTTGGGAAGTAGGCGGTACCGGAAGTTCTATTTG
 CTCTGGACGCAATATTGCCAGCGAACACAGCAACTGGGCGCCTTTGGGCTAATCCAAG
 35 TCAATATGGTGCACCTTCACTGAAATAATGCAGCTGGCCTATAGACTCTTCACTTTTCT
 CAAATCTCAT

DORA45

40 MTTSMQPSKYTGLVADLMPNIRAMKYSGLFMHNFTGGSAFMKKVYSSVHLVFLLMQFT
 FILVNMALNAEEVNELSGNTITTLFFTHCITKFIYLAVNQKNFYRTLNIWNQVNTHTPL
 45 FAESDARYHSIALAKMRKLFVLVMLTTVASATAWTTITFFGDSVKMVVDHETNSSIPV
 EIPRLPIKSFYPWNASHGMFYMISFAFQIYYVLFSMIHSNLCDVMFCSWLI FACEQLQ
 50 HLKGIMKPLMELSASLDTYRPN SAALFRSLSANSKSELIHNEEKDPGTDMDMSGIYSS
 KADWGAQFRAPSTLQSFNGNGGGNGLVNGANPNGLTKKQEMMVRS AIKYWVERHKHV
 VRLVAAIGDTYGAALLHMLTSTIKLTLAYQATKINGVNVYAFTVVGYLGYALAQVF
 55 HFCIFGNRLIEESSVMEAAYSCHWYDGSSEAKTFVQIVCQQCQKAMSISGAKFFTVS
 LDLFASVLGAVVTYFMVLVQLK

DORA45nt

ES 2 315 227 T3

GGCACGAGCTGGTTCCGGAAAGCCTCATATCTCGTATCTTAAAGTATCCCGGTTAAGC
 CTAAAGAGTGAAATGATTGCCTAGACGATTGCTGCATTACTGGCACTCAATTAACCC
 5 AAGTGTACCAGACAACAATTACATTTGTATTTTAAAGTTCAATAGCAAGGATGACAA
 CCTCGATGCAGCCGAGCAAGTACACGGGCCCTGGTCGCCGACCTGATGCCAACATCCG
 GGCGATGAAGTACTCCGGCCTGTTTCATGCACAACCTTCACGGGCGGCAGTGCCTTCATG
 10 AAGAAGGTGTACTCCTCCGTGCACCTGGTGTTCTCCTCATGCAGTTCACCTTCATCC
 TGGTCAACATGGCCCTGAACGCCGAGGAGGTCAACGAGCTGTCCGGCAACACGATCAC
 GACCCTCTTCTTCACCCACTGCATCACGAAGTTTATCTACCTGGCTGTAAACCAGAAG
 15 AATTTCTACAGAACATTGAATATATGGAACCAGGTGAACACGCATCCCTTGTTCCGCCG
 AGTCGGATGCTCGTTACCATTTCGATCGCACTGGCGAAGATGAGGAAGCTGTTCTTTCT
 GGTGATGCTGACCACAGTCGCCTCGGCCACCGCCTGGACCACGATCACCTTCTTTGGC
 20 GACAGCGTAAAAATGGTGGTGGACCATGAGACGAACTCCAGCATCCCGGTGGAGATAC
 CCCGGCTGCCGATTAAGTCCTTCTACCCGTGGAACGCCAGCCACGGCATGTTCTACAT
 GATCAGCTTTGCCTTTCAGATCTACTACGTGCTCTTCTCGATGATCCACTCCAATCTA
 25 TGCGACGTGATGTTCTGCTCTTGGCTGATATTGCCTGCGAGCAGCTGCAGCACTTGA
 AGGGCATCATGAAGCCGCTGATGGAGCTGTCCGCCTCGCTGGACACCTACAGGCCCAA
 CTCGGCGGCCCTCTTCAGGTCCCTGTCCGCCAACTCCAAGTCGGAGCTAATTCTAAT
 30 GAAGAAAAGGATCCCGGCACCGACATGGACATGTCCGGGCATCTACAGCTCGAAAGCGG
 ATTGGGGCGCTCAGTTTCGAGCACCTCGACACTGCAGTCCTTTGGCGGGAACGGGGG
 CGGAGGCAACGGGTTGGTGAACGGCGCTAATCCCAACGGGCTGACCAAAAAGCAGGAG
 35 ATGATGGTGCAGTGCATCAAGTACTGGGTGCGAGCGGCACAAGCACGTGGTGCGAC
 TGGTGGCTGCCATCGGCGATACTTACGGAGCCGCCCTCCTCCTCCACATGCTGACCTC
 GACCATCAAGCTGACCCTGCTGGCATAACAGGCCACCAAAATCAACGGAGTGAATGTC
 40 TACGCCTTCACAGTCGTGCGATACCTAGGATACGCGCTGGCCAGGTGTTCCACTTTT
 GCATCTTTGGCAATCGTCTGATTGAAGAGAGTTTCATCCGTCATGGAGGCCGCCCTACTC
 GTGCCACTGGTACGATGGCTCCGAGGAGGCCAAGACCTTCGTCCAGATCGTGTGCCAG
 CAGTGCCAGAAGGCGATGAGCATATCGGGAGCGAAATCTTCACCGTCTCCCTGGATT
 45 TGTGTTGCTTCGGTTCTGGGTGCCGTCCGTACCTACTTTATGGTGTGGTGCAGCTCAA
 GTAAGTTGCTGCGAAGCTGATGGATTTTGTACCAGAAAAGCGAATGCCAAGAAGCCA
 CCTACCGCCCCCTTGCCCCCTCCGCACTGTGCAACCAGCAATATCACAGAGCAATTATA
 50 ACGCAAATTATATATTTTATACCTGCGACGAGCGAGCCTCGTGGGGCATAATGGAGAC
 ATTCTGGGGCACATAGAAGCCTGCAAATACTTATCGATTTTGTACACGCGTAGAGCTT
 TTAATGTAAACTCAAGATGCAAATAAAATAATGTGTAGTGAAAAAAAAAAAAAAAAAA
 55 AAA

60 *Números de acceso del GENBANK*

Los números de acceso para las secuencias descritas en este artículo son AF127921-AF127926.

65

Bibliografia

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403-410.
- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389-3402.
- Amrein, H., Gorman, M., and Nothiger, R. (1986). The sex-determining gene *tra-2* of *Drosophila* encodes a putative RNA binding protein [published erratum appears in *Cell* 1989 Jul 28; 58(2): following 419]. *Cell* 55, 1025-1035.
- Ayer, R.K., Jr., and Carlson, J. (1992) Olfactory physiology in the *Drosophila* antenna and maxillary palp: *acj6* distinguishes two classes of odorant pathways. *J. Neurobiol.* 23, 965-962.
- Bargmann, C.I., and Kaplan, J. M. (1998). Signal transduction in the *Caenorhabditis elegans* nervous system. *Annu Rev Neurosci* 21, 279-308.
- Bargmann, C. I., Hartweg, E., and Horvitz, H. R. (1993). Odorant-selective genes and neurons mediate olfaction in *C. elegans*. *Cell* 74, 515-527.
- Bargmann, C. I., and Horvitz, H. R. (1991). Chemosensory neurons with overlapping functions direct chemotaxis to multiple chemicals in *C. elegans*. *Neuron* 7, 729-742.
- Ben-Arie, N., Lancet, D., Taylor, C., Khen, M., Walker, N., Ledbetter, D. H., Carrozzo, R., Patel, K., Sheer, D., Lehrach, H., and *et al.* (1994). Olfactory receptor gene cluster on human chromosome 17: possible duplication of an ancestral receptor repertoire. *Hum. Mol. Genet.* 3, 229-235.
- Buck, L. B. (1996). Information coding in the vertebrate olfactory system. *Annu Rev Neurosci* 19, 517-44.
- Buck, L., and Axel, R. (1991). A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell* 65, 175-187.
- Burge, C., and Karlin, S. (1997). Prediction of complete gene structures in human genomic DNA. *J. Mol. Biol.* 268, 78-94.
- Carlson, J. R. (1996). Olfaction in *Drosophila*: from odor to behavior. *Trends Genet.* 12, 175-160.
- Chess, A., Simon, I., Cedar, H., and Axel, R. (1994). Allelic inactivation regulates olfactory receptor gene expression. *Cell* 78, 823-834.
- Colbert, H.A., and Bargmann, C.I. (1995) Odorant-specific adaptation pathways generate olfactory plasticity in *C. elegans*. *Neuron* 14, 803-812.
- Cserzo, M., Wallin, E., Simon, I., von Heijne, G., and Elofsson, A. (1997). Prediction of transmembrane-helices in prokaryotic membrane proteins: the dense alignment surface method. *Protein Eng.* 10, 673-676.
- Doe, C. Q., and Skeath, J. B. (1996). Neurogenesis in the insect central nervous system. *Curr. Opin. Neurobiol.* 6, 18-24.
- Dulac, C., and Axel, R. (1995). A novel family of genes encoding putative pheromone receptors in mammals. *Cell* 83, 195-206.
- Faber, T., Joerges, J., and Menzel, R. (1996). Associative learning modifies neural representations of odors in the insect brain. *Nature Neurosci.* 2, 74-78.
- Friedrich, R. W., and Korsching, S.I. (1997). Combinatorial and chemotopic odorant coding in the zebrafish olfactory bulb visualized by optical imaging. *Neuron* 18, 737-752.
- Grillenzoni, N., van Helden, J., Dambly-Chaudiere, C., and Ghysen, A. (1998). The iroquois complex controls the somatotopy of *Drosophila notum* mechanosensory projections. *Development* 125, 3563-9.
- Hartl, D. L., Nurminsky, D. I., Jones, R. W., and Lozovskaya, E. R. (1994). Genome structure and evolution in *Drosophila*: applications of the framework P1 map. *Proc. Natl. Acad. Sci. U S A* 91, 6824-6829.
- Herrada, G., and Dulac, C. (1997). A novel family of putative pheromone receptors in mammals with a topographically organized and sexually dimorphic distribution. *Cell* 90, 763-773.

- Imamura, K., Mataga, N., and Mori, K. (1992).** Coding of odor molecules by mitral/tufted cells in rabbit olfactory bulb. I. Aliphatic compounds. *J. Neurophysiol.* 68, 1986-2002.
- Joerges, J., Kuttner, A., Galizia, C. G., and Menzel, R. (1997).** Presentations of odours and odour mixtures visualized in the honeybee brain. *Nature* 387, 285-288.
- Kato, K., Koshimoto, H., Tani, A., and Mori, K. (1993).** Coding of odor molecules by mitral/tufted cells in rabbit olfactory bulb. II. Aromatic compounds. *J. Neurophysiol.* 70, 2161-2175.
- Kauer, J. S., Senseman, D. M., and Cohen, L. B. (1987).** Odor-elicited activity monitored simultaneously from 124 regions of the salamander olfactory bulb using a voltage-sensitive dye. *Brain Res.* 418, 255-261.
- Kim, M. S., Repp, A., and Smith, D. P. (1998).** LUSH odorant-binding protein mediates chemosensory responses to alcohols in *Drosophila melanogaster*. *Genetics* 150, 711-21.
- Kimmerly, W., Stultz, K., Lewis, S., Lewis, K., Lustre, V., Romero, R., Benke, J., Sun, D., Shirley, G., Martin, C., and Palazzolo, M. (1996).** A P1-based physical map of the *Drosophila* euchromatic genome. *Genome Res.* 6, 414-430.
- Kyte, J., and Doolittle, R. F. (1982).** A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157, 105-132.
- Laissue, P. P., Reiter, C., Hiesinger, P. R., Halter, S., Fischbach, K. F., and Stocker, R. F. (1999)** 3D reconstruction of the antennal lobe in *Drosophila melanogaster*. *J. Comp. Neurol.*, in press.
- Lancet, D., Greer, C. A., Kauer, J. S., and Shepherd, G. M. (1982).** Mapping of odor-related neuronal activity in the olfactory bulb by high-resolution 2-deoxyglucose autoradiography. *Proc. Natl. Acad. Sci. USA* 79, 670-674.
- Levy, N. S., Bakalyar, H. A., and Reed, R. R. (1991).** Signal transduction in olfactory neurons. *J. Steroid Biochem. Mol. Biol.* 39, 633-637.
- Lin, D. M., and Goodman, C. S. (1994).** Ectopic and increased expression of Fasciclin II alters motoneuron growth cone guidance. *Neuron* 13, 507-523.
- Matsunami, H., and Buck, L. B. (1997).** A multigene family encoding a diverse array of putative pheromone receptors in mammals. *Cell* 90, 775-84.
- McKenna, M. P., Hekmat-Safe, D. S., Gaines, P., and Carlson, J. R. (1994).** Putative *Drosophila* pheromone-binding proteins expressed in a subregion of the olfactory system. *J. Biol. Chem.* 269, 16340-16347.
- McLatchie, L. M., Fraser, N. J., Main, M. J., Wise, A., Brown, J., Thompson, N., Solari, R., Lee, M. G., and Foord, S. M. (1998).** RAMPs regulate the transport and ligand specificity of the calcitonin-receptor-like receptor. *Nature* 393, 333-9.
- Merritt, D. J., and Whittington, P. M. (1995).** Central projections of sensory neurons in the *Drosophila* embryo correlate with sensory modality, soma position, and proneural gene function. *J. Neurosci.* 15, 1755-67.
- Mitchell, R., McCulloch, D., Lutz, E., Johnson, M., MacKenzie, C., Fennell, M., Fink, G., Zhou, W., and Sealfon, S. C. (1998).** Rhodopsin-family receptors associate with small G proteins to activate phospholipase D. *Nature* 392, 411-4.
- Mombaerts, P., Wang, F., Dulac, C., Chao, S. K., Nemes, A., Mendelsohn, M., Edmondson, J., and Axel, R. (1996).** Visualizing an olfactory sensory map. *Cell* 87, 675-686.
- Mori, K., Mataga, N., and Imamura, K. (1992).** Differential specificities of single mitral cells in rabbit olfactory bulb for a homologous series of fatty acid odor molecules. *J. Neurophysiol.* 67, 786-789.
- Ngai, J., Chess, A., Dowling, M. M., Necles, N., Macagno, E. R., and Axel, R. (1993).** Coding of olfactory information: topography of odorant receptor expression in the catfish olfactory epithelium. *Cell* 72, 667-680.
- Parmentier, M., Libert, F., Schurmans, S., Schiffmann, S., Lefort, A., Eggericks, D., Ledent, C., Molleareau, C., Gerard, D., and et al. (1992).** Expression of members of the putative olfactory receptor gene family in mammalian germ cells. *Nature* 355, 453-455.
- Pelosi, P. (1994).** Odorant-binding proteins. *Crit. Rev. Biochem. Mol. Biol.* 29, 199-228.
- Persson, B., and Argos, P. (1994).** Prediction of transmembrane segments in proteins utilizing multiple sequence alignments. *J. Mol. Biol.* 237, 182-192.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

ES 2 315 227 T3

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

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

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

10

15

20

25

30

35

40

45

50

55

60

65

REIVINDICACIONES

1. Un ácido nucleico aislado que codifica un receptor de odorizantes de insecto, donde el ácido nucleico codifica un polipéptido que comprende aminoácidos consecutivos, cuya secuencia es idéntica a la expuesta para DORA45 en la SEC ID N°: 107.

2. El ácido nucleico de la reivindicación 1, donde el receptor de odorizantes de insecto comprende siete dominios transmembrana.

3. El ácido nucleico de una cualquiera de las reivindicaciones 1 o 2, donde el ácido nucleico es ADN o ARN.

4. El ácido nucleico de la reivindicación 3, donde el ADN es ADNc, ADN genómico o ADN sintético.

5. El ácido nucleico de una cualquiera de las reivindicaciones 1-4, donde el ácido nucleico codifica un receptor de odorizantes de *Drosophila*.

6. Un vector que comprende el ácido nucleico de una cualquiera de las reivindicaciones 1-5.

7. El vector de la reivindicación 6, donde el ácido nucleico aislado está unido operativamente a un elemento regulador.

8. Un plásmido que comprende el vector de la reivindicación 6.

9. Un sistema de vector-hospedador para la producción de un polipéptido que tiene la actividad biológica de un receptor de odorizantes de insecto, que comprende el vector de la reivindicación 6 y una célula hospedadora adecuada.

10. El sistema de vector-hospedador de la reivindicación 9, donde el hospedador adecuado es una célula bacteriana, una célula de levadura, una célula de insecto o una célula animal.

11. Un método de producción de un polipéptido que tiene la actividad biológica de un receptor de odorizantes de insecto, que comprende cultivar el sistema de vector-hospedador de la reivindicación 9 en condiciones que permitan la producción del polipéptido y recuperar el polipéptido producido de este modo.

12. Un receptor de odorizantes de insecto purificado codificado por el ácido nucleico de una cualquiera de las reivindicaciones 1-5.

13. Un anticuerpo que se une específicamente a un receptor de odorizantes de insecto codificado por el ácido nucleico aislado de una cualquiera de las reivindicaciones 1-5.

14. El anticuerpo de la reivindicación 13, donde el anticuerpo es un anticuerpo monoclonal.

15. Un método para identificar un inserto de ADNc que codifica un receptor de odorizantes de insecto, que comprende:

(a) generar una genoteca de ADNc que comprende clones que llevan insertos de ADNc a partir de neuronas sensoriales de la antena o del palpo maxilar de un insecto;

(b) hibridar ácidos nucleicos de los clones de las genotecas de ADNc generadas en la etapa (a) con una sonda que tiene una secuencia exclusivamente presente en un ácido nucleico de la reivindicación 1; y

(c) aislar los ácidos nucleicos hibridados resultantes para identificar de este modo el inserto de ADNc que codifica el receptor de odorizantes de insecto.

16. Un método para obtener un ácido nucleico que codifica un receptor de odorizantes a partir de un insecto, que comprende:

(a) poner en contacto una muestra que contiene ácido nucleico de origen de insecto con cebadores de reacción en cadena de la polimerasa que hibriden específicamente con un ácido nucleico de la reivindicación 1 en condiciones apropiadas que permiten la hibridación de los cebadores con el ácido nucleico para producir un producto de hibridación;

(b) amplificar el producto de hibridación resultante usando la reacción en cadena de la polimerasa; y

(c) aislar las moléculas amplificadas, obteniendo así el ácido nucleico que codifica el receptor de odorizantes del insecto.

ES 2 315 227 T3

17. Un método de transformación de una célula que comprende transfectar una célula hospedadora con el vector de la reivindicación 6.

5 18. Una célula transformada producida por el método de la reivindicación 17 donde antes de transfectarse con el vector, la célula hospedadora no expresa un receptor de odorizantes de insecto.

10 19. Un método de identificación de un compuesto que se une específicamente a un receptor de odorizantes de insecto, que comprende poner en contacto la célula transformada de la reivindicación 18, o una fracción de membrana de dicha célula, con el compuesto en condiciones que permitan la unión del compuesto al receptor de odorizantes, detectar la presencia de dicho compuesto unido específicamente al receptor y de este modo identificar el compuesto como un compuesto que se une específicamente a un receptor de odorizantes de insecto.

15 20. Un método de identificación de un compuesto que se une específicamente a un receptor de odorizantes de insecto, que comprende poner en contacto el receptor de odorizantes purificado de la reivindicación 12 con el compuesto en condiciones que permitan la unión del compuesto al receptor de odorizantes purificado, detectar la presencia de dicho compuesto unido específicamente al receptor y de este modo identificar el compuesto como un compuesto que se une específicamente a un receptor de odorizantes de insecto.

20 21. El método de la reivindicación 20, en el que el receptor de odorizantes de insecto purificado está incorporado en una bicapa lipídica.

25

30

35

40

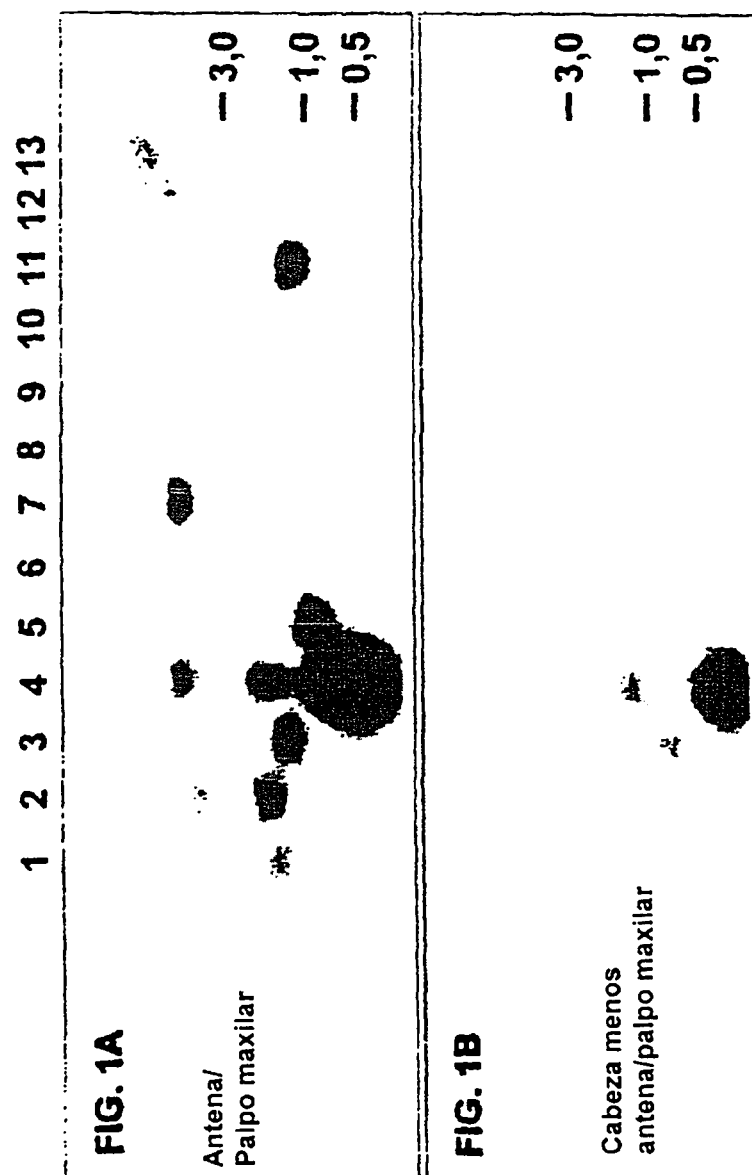
45

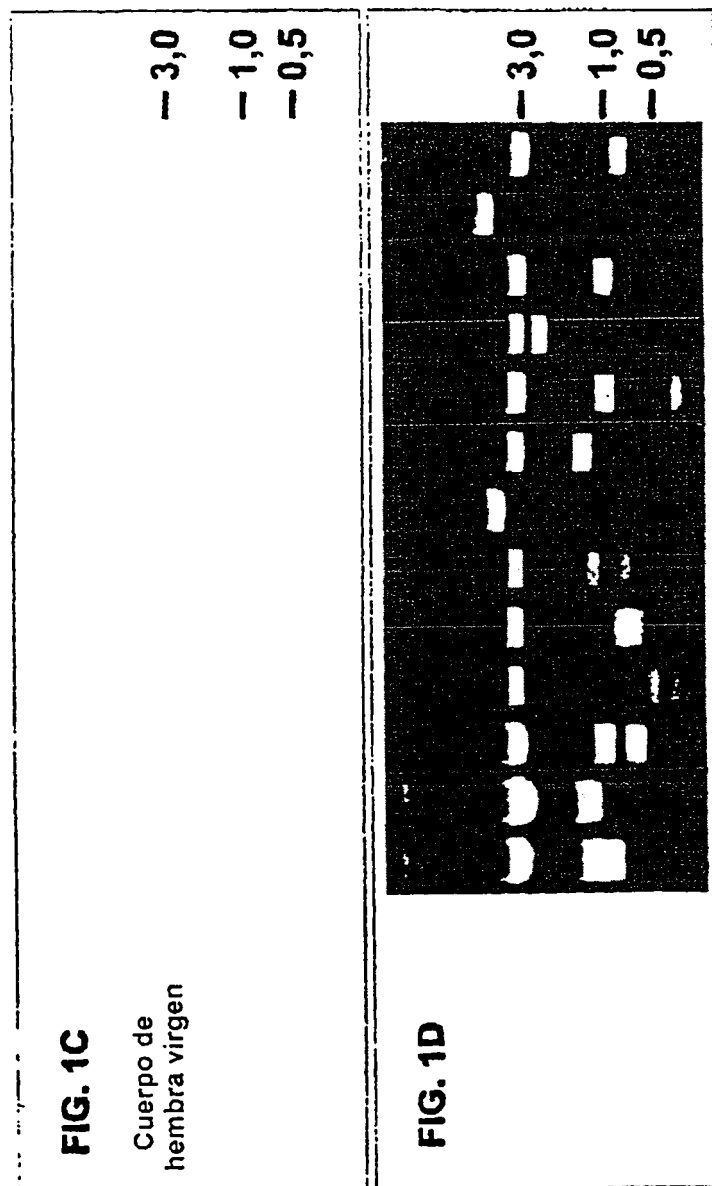
50

55

60

65





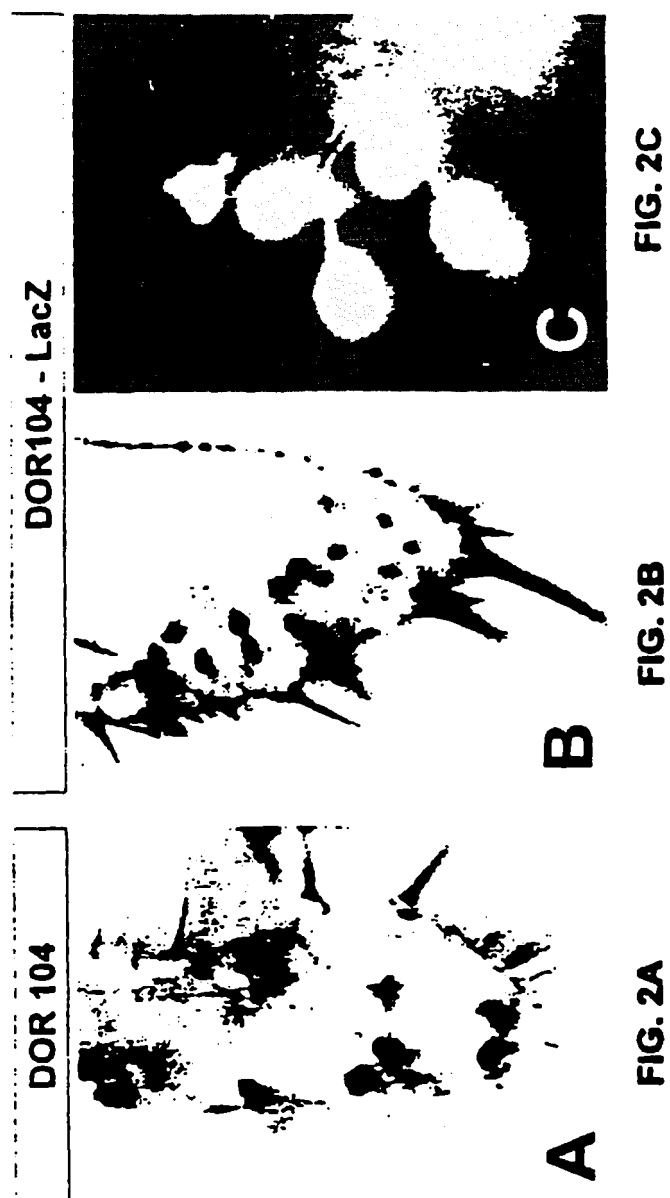
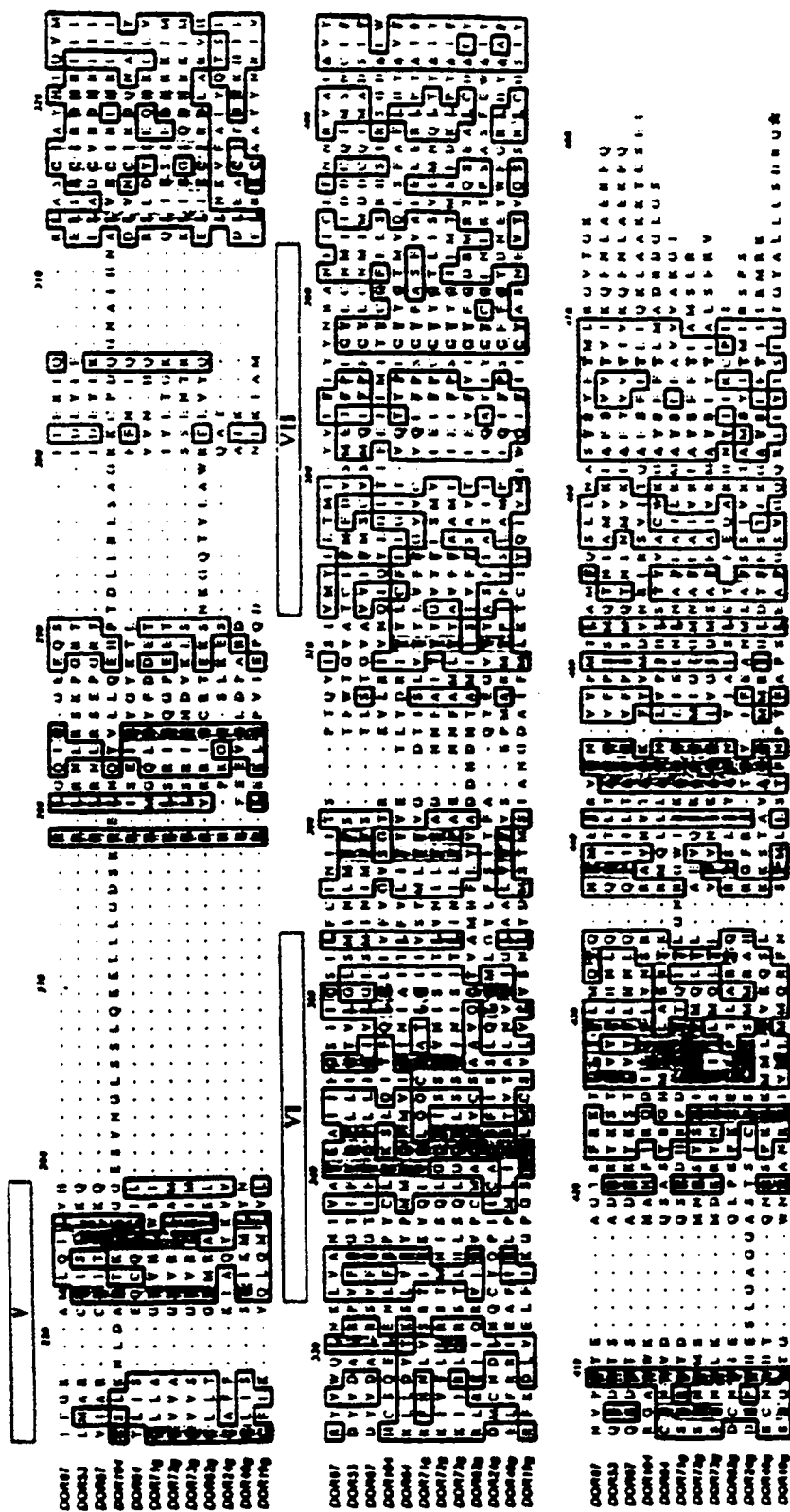
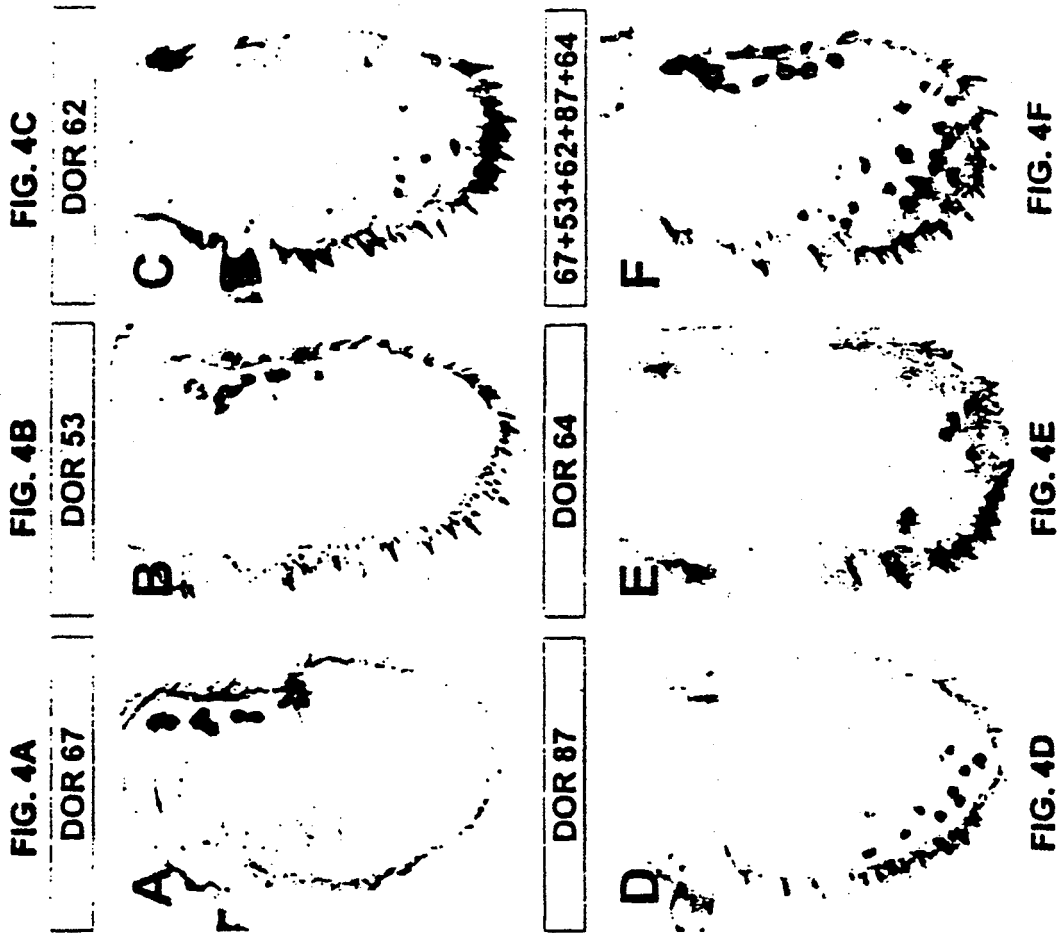


FIG. 3B





DOR 71

DOR 72

DOR 73

G



FIG. 4G

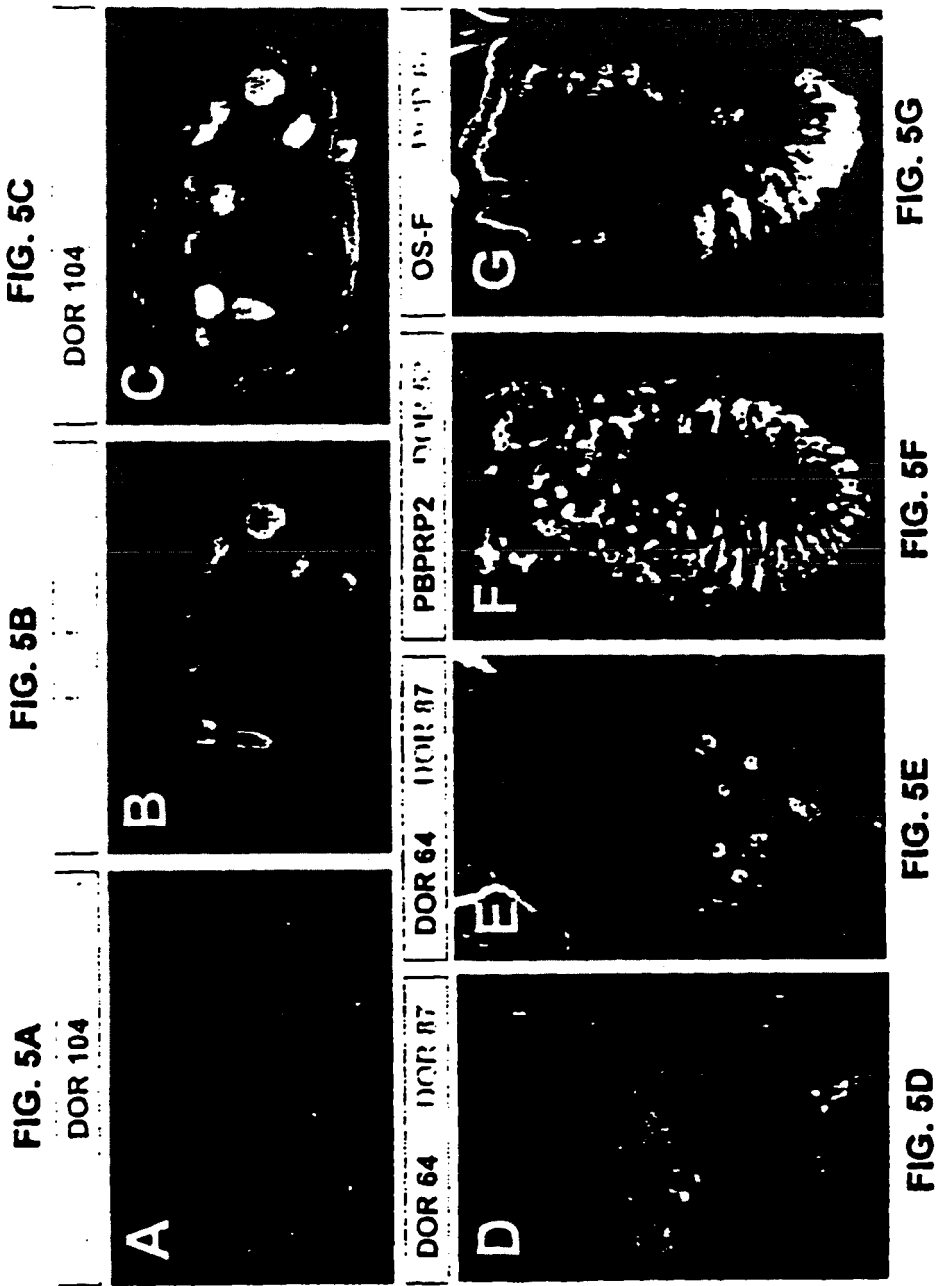
H



FIG. 4H



FIG. 4I



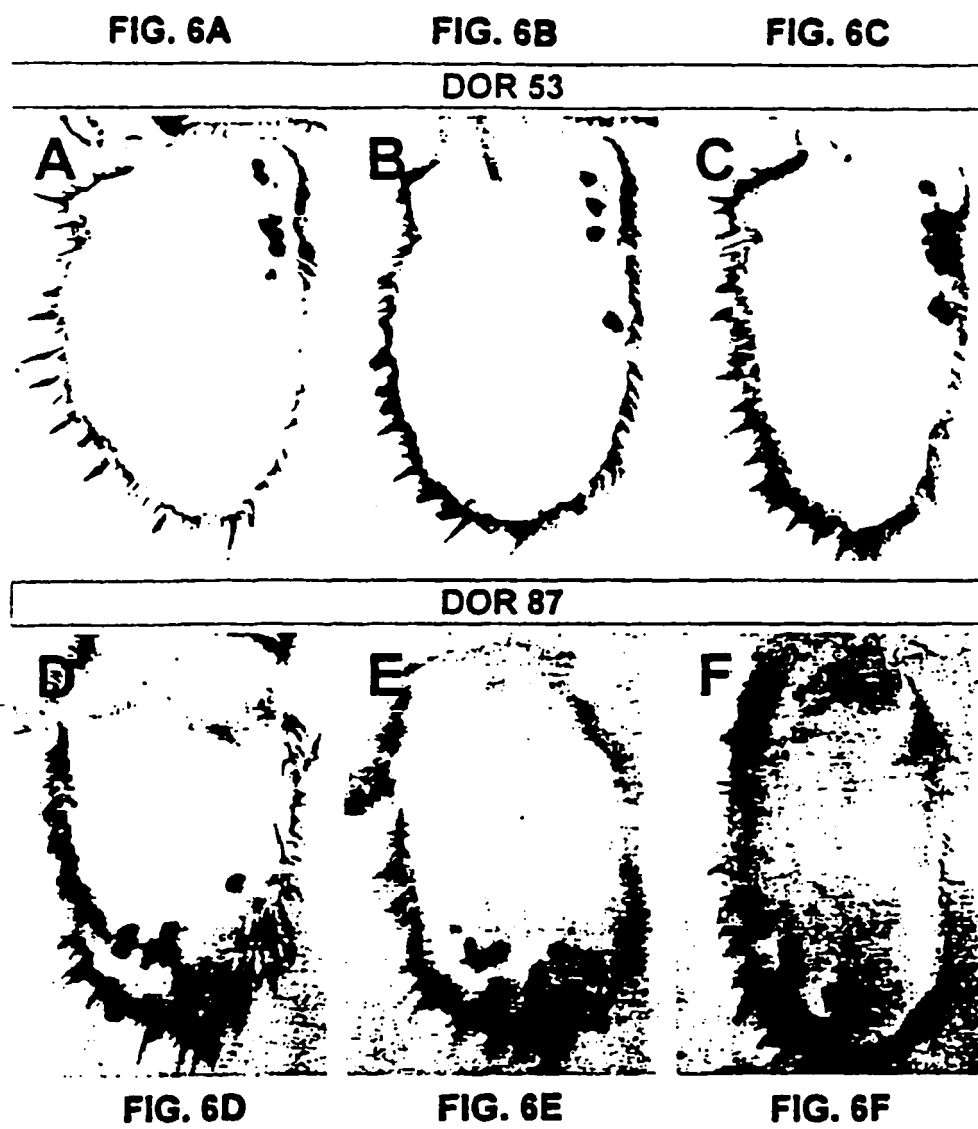


FIG. 7A

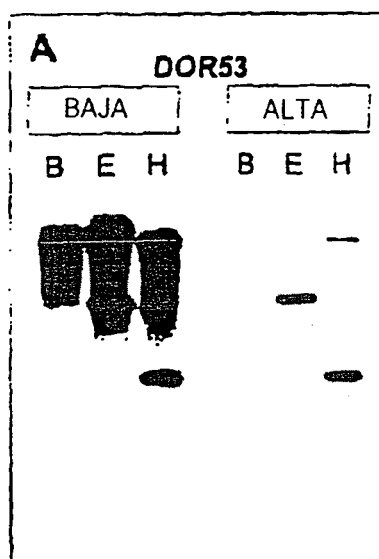


FIG. 7B

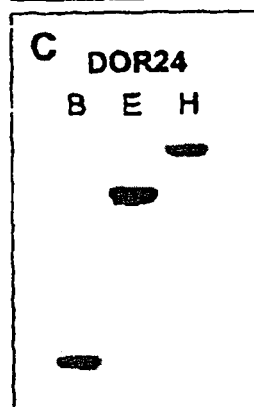
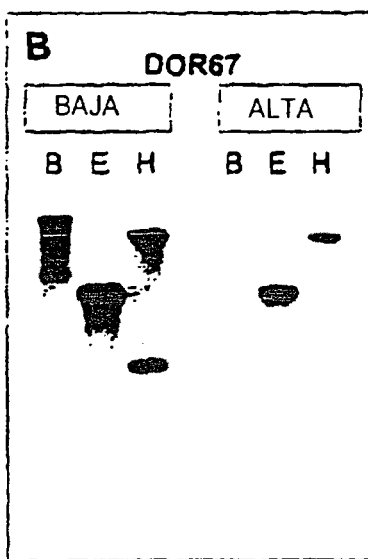


FIG. 7C

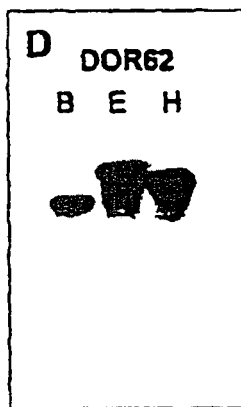


FIG. 7D

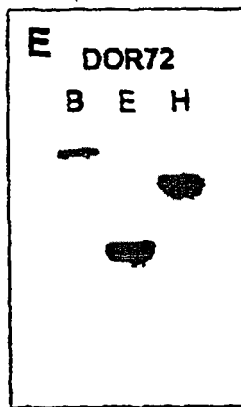


FIG. 7E

FIG. 8A

DOR62 [AF127921]
 LOCUS xxxxx 553 pb ARNm INV 2 de feb de 1999
 DEFINICIÓN ARNm de supuesto receptor de odorizante de *Drosophila melanogaster* DOR62 cds parcial
 ACCESO AF127921
 PALABRAS CLAVE
 FUENTE mosca de la fruta
 ORGANISMO *Drosophila melanogaster*
 Eukaryota: Metazoa; Arthropoda; Tracheata; Hexapoda;
 Insecta; Pterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; *Drosophila*.
 REFERENCIA 1 (bases 1 a 553)
 AUTORES Leslie B. Vosshall, Hubert Amrein, Pavel S. Morozov,
 Andrey Rzhetsky y Richard Axel.
 TÍTULO Un mapa espacial de la expresión de receptor olfatorio
 en la antena de *Drosophila*
 REVISTA Cell (1999) en prensa
 REFERENCIA 2 2 (bases 1 a 553)
 AUTORES Vosshall, L. B., Amrein, H., Morozov, P. S., Rzhetsky,
 A. y Axel Richard.
 TÍTULO Presentación directa
 REVISTA Presentado (02-FEB-1999) Departamento de Bioquímica y
 Biofísica Molecular, HHMI-Universidad de Columbia, 701
 West 168th Street, Nueva York, NY 10032, Estados Unidos
 CARACTERÍSTICAS Localización/Clasificadores
 Comentarios: /producto = 'receptor de odorizante DOR62'
 /tipo_tisular = 'antena de adulto'
 /nota 'supuesto receptor acoplado a proteínas G
 con siete dominios transmembrana; ADN genómico
 secuenciado por el proyecto genoma de *Drosophila*
 europeo y asignado el número de acceso 62D9.q'
 Fuente 1..553
 /organismo = "*Drosophila melanogaster*"
 /cepa = "Oregon R"
 /db_xref = "taxón: 7227"
 /cromosoma = "X"
 /mapa = "2F"
 Gen <1..553
 /nota = "receptor de odorizante DOR62 de
Drosophila melanogaster"
 /gen = "DOR62"
 CDS <1..459
 /gen = "DOR62"
 /nota = "receptor de odorizante"
 /codón_inicio = 1
 /producto = "DOR62"
 /traducción = "QELIECIRDLARVHRLREIIQRVLSVPCMAQFVCSAAVOCTVAM
 #FLYVADDHNDHTAMIIIVFFSAVTLEVFVICYFGDRMRTQSEALCDAFYDCNWIEQL
 PKFKRELLFTLARTQRPSTLIYAGNYIALSLETFEQVMRFTYSVFTLLLRAX"
 RECUENTO DE BASES 141 a 145 c 132 g 135 t
 ORIGEN
 1 caggaactca tcgagtgcac ccgcgatctg gcgcgggccc atccggctgag ggagattcat
 61 cagcgggtcc tttcagtgcc ctgcattggcc cagttcgtct gctccgcgc cgtccagtg
 121 accgtcgcga tgcacttctt gtacgtagcg gatgaccacg accacaccgc catgatcat
 181 tcgattgtat tttctcggc cgtcaccctg gaggtgtttg taattcgtca ttttggggac
 241 aggatgcgga cacagagcga ggcgtgtgac gatgcttctt acgattgcaa ctggatagaa
 301 cagctgcccga agttcaagcg cgaactgtct tccaccctgg ccaggacgca gcggccttct
 361 cttacttacc caggcaacta catcgacttc tcgtgggaga ctttcgagca ggtcatgagg
 421 ttcacatact ctgttttcac actcttctgt agggccaaagt aagaacttta taattcttct
 481 ttggggagaa aaattctaaa gcacaatagc agaaaaatat atcagataat ataacaaaaa
 541 aaaaaaaaaa aaa

ES 2 315 227 T3

FIG. 8B-1

DOR104 [AF127922]

LOCUS xxxxx 1493 pb ARNm INV 2 de feb de 1999
DEFINICIÓN ARNm de supuesto receptor de odorizante de *Drosophila melanogaster* DOR104, cds completa
ACCESO AF127922

PALABRAS CLAVE

FUENTE

ORGANISMO

mosca de la fruta

Drosophila melanogaster

Eukaryota: Metazoa; Arthropoda; Tracheata; Hexapoda;
Insecta; Pterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; *Drosophila*.

REFERENCIA

AUTORES

TÍTULO

REVISTA

REFERENCIA 2

AUTORES

TÍTULO

REVISTA

1 (bases 1 a 1493)

Leslie B. Vosshall, Hubert Amrein, Pavel S. Morozov,
Andrey Rzhetsky y Richard Axel.

Un mapa espacial de la expresión de receptor olfatorio
en la antena de *Drosophila*

Cell (1999) en prensa

2 (bases 1 a 1493)

Vosshall, L. B., Amrein, H., Morozov, P. S., Rzhetsky,
A. y Axel R.

Presentación directa

Presentado (02-FEB-1999) Departamento de Bioquímica y
Biofísica Molecular, HHMI-Universidad de Columbia, 701
West 168th Street, Nueva York, NY 10032, Estados Unidos

CARACTERÍSTICAS

Comentarios:

Fuente

Gen

CDS

Localización/Clasificadores

/tipo_tisular = 'palpo maxilar de adulto'
/nota 'supuesto receptor acoplado a proteínas G
con siete dominios transmembrana'

1..1493

/organismo = "*Drosophila melanogaster*"

/cepa = "Oregon R"

/db_xref = "taxón: 7227"

/cromosoma = "III"

/mapa = "85B"

1..1493

/nota = "receptor de odorizante"

/gen = "DOR104"

21..1421

/gen = "DOR104"

/nota = "receptor de odorizante"

/codón_inicio = 1

/producto = "DOR104"

/traducción = " MASLQFHGNVDADIRYDISLDPARESNLFRLLMGLQLANGTKPS

PRLPKWWPKRLEMIGKVLKAYCSMVIFTSLHLGVLFKTTLDVLPTELQAITDALT
MTIIYFTGYGTIYWCILSRRLAYMEHMRREYRHHSLAGVTFVSSHAAFRMSRNFV
VWIMSCLLGVISWGVSPMLGIRMLPLQCWYFFDALGPGTYTAVYATQLFGQIMVGMT
FGFGGSLFVTLSSLLLGQFDVLYCSLKNLDAHTKLLGGESVNGLSLQELLLGDSKR
ELNQYVLLQEHPTDLLRLSAGRKCPDQGNAFHNAIVECIRLHRFILHCSQELNLFSP
YCLVKSLQITFQLCLLVFVGVSQGTREVLRIVNLQYLGLTIFELLMFTYCGELLSRHS
CRSGDAFWRGAWWKHAHFIRQDILIFLVNSRRVHVHTAGKFYVMDVNRLSRVITQAFS
FLTLQKLAAKKTESEL"

RECuento DE BASES 332 a 373 c 407 g 381 t

FIG. 8B-2

ORIGEN

```

1  ggacacgagca gtcgatggcc agtctccagt tccacggcaa cgtcgatgcg gacatcaggc
61  atgatatattag cctggatccg gctaggggaat cgaatctctt ccgtctgcta atgggactcc
121 agttggcgaa tggcacgaag ccaccggccg ggttacccaa atgggtggcca aagcggctcg
181 aaatgatttg taaagtgcctg cccaaagcct attgttccat ggtgattttc acctccctgc
241 atttgggtgt cctgttcacg aaaaccacac tggatgtccc gcgacgggg gagctgcagg
301 ccataacgga tgccttcacc atgaccataa tatacttttt cacgggctac ggcaccatcc
361 actggtgcct gcgctcccg gcctctcttg cctacatgga gcacatgaac cgggagttac
421 gccatcattc gctggccgg gtgacctttg tgagttagcca tgcggccttc aggatgtcca
481 gaaactttac ggtggtgtcg ataattgctct gcctgctggg cgtgattttc tggggcgttc
541 cgccactgat gctgggcac cggatgctgc cgtcccaatg ttggtatccc ttcgacgccc
601 tgggtcccg caccatatac gcggtctatg ctacacaacc ttccggtcag atcatggtgg
661 gcatgacctt tggatttcgg ggatcactgt ttgtcacccc gagcctgtca ctcctgggac
721 aattcgatgt gctctactgc agcctgaaga acctggatgc ccataccaag ttgctgggcg
781 gggagttctg aaatggcctg agtccgctgc aagaggagtc gctgctgggg gactcgaaga
841 gggaattaaa tcagtacgtt ttgtccagg agcatccgac ggtctgtctg agattgtcgg
901 caggacgaaa atgtcctgac caaggaaaag cgttccacaa cgccttggtg gaatgcattc
961 gcttgcatcg ctccattctg cactgctcac aggagttgga gaatctattc agtccatatt
1021 gtctgggtcaa gtcaactgcag atcaccttct agctttgccc gctggctctt gtgggctctt
1081 cgggtactcg agaggtcctg cggattgtca accagctaca gtacttggga ctgacctctt
1141 tcgagctcct aatgttcacc tattgtggcg aactcctcag tcggcatagt attcgatctg
1201 gcgacgccc ttggaggggt gcgtgggtgga agcacgcca ttcaatccgc caggacatcc
1261 tcatctttct ggtcaatagt agacgtgcag ttcacgtgac tgcgggcaa ttttatgtga
1321 tggatgtgaa tcgtctaaag tcggttataa cgcaggcgtt cagcttcttg accttgcctg
1381 aaaagtctgg tcccaagaag acggaatcgg agctctaaac cggtaaccag catcgatatt
1441 tatttagcgc attaaaaaaa agtcgagtaa aagcaaaaaa aaaaaaaa aa

```

ES 2 315 227 T3

FIG. 8C-1

DOR87 [AF127926]

LOCUS	xxxxx 1556 pb ARNm INV 2 de feb de 1999
DEFINICIÓN	ARNm de supuesto receptor de odorizante de <i>Drosophila melanogaster</i> DOR87, cds completa AF127926
ACCESO	
PALABRAS CLAVE	
FUENTE	mosca de la fruta
ORGANISMO	<i>Drosophila melanogaster</i> Eukaryota: Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; <i>Drosophila</i> .
REFERENCIA	1 (bases 1 a 1556)
AUTORES	Leslie B. Vosshall, Hubert Amrein, Pavel S. Morozov, Andrey Rzhetsky y Richard Axel.
TÍTULO	Un mapa espacial de expresión de receptor olfatorio en la antena de <i>Drosophila</i>
REVISTA	Cell (1999) en prensa
REFERENCIA 2	2 (bases 1 a 1556)
AUTORES	Vosshall, L. B., Amrein, H., Morozov, P. S., Rzhetsky, A. y Axel R.
TÍTULO	Presentación directa
REVISTA	Presentado (02-FEB-1999) Departamento de Bioquímica y Biofísica Molecular, HHMI-Universidad de Columbia, 701 West 168th Street, Nueva York, NY 10032, Estados Unidos
CARACTERÍSTICAS	Localización/Clasificadores
Comentarios:	/tipo_tisular = 'antena de adulto' /nota 'supuesto receptor acoplado a proteínas G con siete dominios transmembrana'
Fuente	1..1556 /organismo = " <i>Drosophila melanogaster</i> " /cepa = "Oregon R" /db_xref = "taxón: 7227" /cromosoma = "II" /mapa = "43B1-2"
Gen	1..1556 /nota = "receptor de odorizante" /gen = "DOR87"
CDS	30..1157 /gen = "DOR87" /nota = "receptor de odorizante" /codón_inicio = 1 /producto = "DOR87"
/traducción = "MTIEDIGLVGINVRMWRHLAVLYPTPGSSWRKFAFVLPVTAMNL MQFVYLLRMWGDLPFILNMFFFSIAIFNALMRTWLVIKRRQFEFLGQLATLFHSIL DSTDEWGRGILRRAEREARNLAILNLSASFLLDIVGALVSPFREERAHFFGVALPGVS MTSSPVYEVITYLAQLPTPLLSMMYMPFVSLFAGLAIFGKAMLQILVHRLGQIGGEEQ SEERFORLASCIAHYHTQVMRYVWQLNKLVANIVAVEAIIFGSIIICSLFCLNIITSP TQVISIVMYILTMLYVLFYTYNRANEICLENRVAAEVNVPWYEAGTRFRKTLIFL MQTQHPMEIRVGNVYPMTLAMFQSLNASYSYFTMLRGVTGK"	
RECuento de bases	329 a 413 c 408 g 406 t

FIG. 8C-2

ORIGEN

```

1  ggcacgagggc ttatagaaag tgccgagcaa tgacaatcga ggatatcggc ctcggcgggca
61  tcaacgtggc gatgtggcga cacttgggcg cgctgtaccc cactccgggc tccagctggc
121 gcaagttcgc cctcgtgctg ccggcgactg cgatgaacct gatgcagtc gcttacctgc
181 tcgggatgtg gggcgacctg ccgccttca ttctgaacat gtcttcttc tcggccattc
241 tcaacgccct gatgcgcacg tggctgggtca taatcaagcg gcgccagtc gaggagtttc
301 tcggccaact ggccactctg ttccattcga ttctcgactc caccgacgag tgggggcgtg
361 gcatcctgcg gagggcgga cgggaggctc ggaacctggc catccttaat tcgagtgcct
421 ccttcttggg cattgtcggc gctctgggtat cgccgctctt cagggaggag agagctcatc
481 ccttcggcgt agctctacca ggagtgaaga tgaccagtc acccgctctac gaggttatct
541 acttggtcca actgcctacg cccctgctgc tgtccatgat gtacatgcct ttcgtcagcc
601 tttttgcccg cctggccatc tttgggaagg ccatgctgca gatccttgta cacaggctgg
661 gccagattgg cggagaagag cagtcggagg aggagcgctt ccaaaggctg gcctcctgca
721 ttgctacca cagcgaggtg atgcgctatg tgtggcagct caacaaactg gtggccaaca
781 ttgtggcggc ggaagcaatt attttggct cgataatctg ctactgctc tctgtctga
841 atattataac cccaccacc caggtgatct cgatagtgat gtacattctg accatgctgt
901 acgttctctt cactactac aatcgggcca atgaaatatg cctcgagaac aactgggtgg
961 cggaggctgt ttacaatgtg ccttggtagc aggcaggaaac tcggtttcgc aaaacccctc
1021 tgatctctct gatgcaaca caacaccga tggagataag agtcggcaac gtttacccca
1081 tgacattggc catgttccag agtctgttga atgcgtccca cctctacttt accatgctgc
1141 gtggcgctac cggcaaatga gctgaaagac cgaaaaaacc ggagtatccc cttccatatt
1201 cccctgctc ctttattctt ctttctcttt ccttttccgt tttcccatc gcttttccag
1261 caatccgggt aatgcaaaaa gttgttgcgt gctgtggctc tggctgctt tttggcattt
1321 gcatacgctt gtcgtttgaa aggatttaac cggactgctg gcacggagtc ggcatctgg
1381 ctcttggtc ctggcatgca aatagttggc tctttagatt gttacacaaa atagattgca
1441 gatcgagct gaatgttctg cttggaataa agtcaaaagg atgtggagtc ggcccaaggc
1501 tctgcccatt ctgcttgcctc gggatgcccc aaagtatgaa aaaaaaaaaa aaaaaa

```

FIG. 8D-1

DOR53 [AF127923]

LOCUS xxxxx 1305 pb ARNm INV 6 de feb de 1999
 DEFICINICIÓN ARNm de supuesto receptor de odorizante de *Drosophila melanogaster* DOR53, cds completa
 AF127923

ACCESO
 PALABRAS CLAVE
 FUENTE mosca de la fruta
 ORGANISMO *Drosophila melanogaster*
 Eukaryota: Metazoa; Arthropoda; Tracheata; Hexapoda;
 Insecta; Pterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; *Drosophila*.

REFERENCIA 1 (bases 1 a 1305)
 AUTORES Leslie B. Vosshall, Hubert Amrein, Pavel S. Morozov, Andrey Rzhetsky y Richard Axel.
 TÍTULO Un mapa espacial de expresión de receptor olfatorio en la antena de *Drosophila*
 REVISTA Cell (1999) en prensa

REFERENCIA 2 (bases 1 a 1305)
 AUTORES Vosshall, L. B., Amrein, H., Morozov, P. S., Rzhetsky, A. y Axel R.
 TÍTULO Presentación directa
 REVISTA Presentado (06-FEB-1999) Departamento de Bioquímica y Biofísica Molecular, HHMI-Universidad de Columbia, 701 West 168th Street, Nueva York, NY 10032, Estados Unidos

CARACTERÍSTICAS Localización/Clasificadores

Fuente 1..1305
 /organismo = "*Drosophila melanogaster*"
 /cepa = "Oregon R"
 /db_xref = "taxón: 7227"
 /nota="tipo tisular = 'antena de adulto'
 _nota='supuesto receptor acoplado a proteínas G con siete dominios transmembrana'
 /cromosoma = "II"
 /mapa = "22A2-3"

Gen 1..1305
 /nota = "receptor de odorizante"
 /gen = "DOR53"

CDS 79..1179
 /gen = "DOR53"
 /nota = "receptor de odorizante"
 /codón_inicio = 1
 /producto = "DOR53"

/traducción = "MWSFGWTEPENKRWILPYKLWLA FVNIVMLILLPISISIEYLHR
 FKTF SAGEFLSSLEIGVMY GSSFKCAFTLIGFKKRQEAKVLLDQLDKRCLSDKERST
 VHR YVAMGNFFDILYHIFYSTFVVMNFPYLLERRHAWRMYPYIDSDEQFYISSIAE
 CFLMTEAIYMDLCTDVCPLISMLMARCHISLLKQRLRNLRSKPGRTEDEYLEELTECI
 RDHRLLLDYVDALRPVFSGTIFVQFLLIGTVLGLSMINLMFFSTFWTGVATCLFMFDV
 SMETFFCYLCNMIIDDCQEMSNC LFQSDWTSADRRYKSTLVYFLHNLQQPITLTAGG
 VFPISMQTNLAMVKLAFSVVTVIKQFNLAERFQ"

RECuento DE BASES 331 a 186 c 279 g 409 t

FIG. 8D-2

ORIGIN

```

1  ttttttcccc acataaaaaga aaagccattg agcgagcggg ctaagtcccc agatgccttc
61  atttacctgg atcgggtgat gctgtccctt ggctggacag agcctgaaaa caaaaggctg
121 atcctccctt ataaactgtg gttagcgttc gtgaacatag taatgctcat ccttctgccc
181 atctcgataa gcattcgagta cctccaccga tttaaaaacct tctcggcggg ggagttctct
241 agtccctctg agattggagt caacatgtac ggaagctctt ttaagtgcgc cctcaccttg
301 attggattca agaaaagaca ggaagctaag gctttacttg atcagctgga caagagatgc
361 cttagcgata aggagaggtc cactgttcac cgctatgtcg ccatgggaaa ccttttcgat
421 attctgtatc acatttttta ctccaccttc gtggtaatga acttcccgta tctctctgct
481 gagagacgcc atgctggcg catgtacttc ccataatcgc attccgacga acagttctac
541 atctccagca tcgctgagtg tttctctgat acggaggcca tctacatgga tctctgtacg
601 gacgtgtgtc ccttgatctc catgcttatg gctcgtatgc acatcagcct cctgaaacag
661 cgactgagaa atctccgacg gaagccagga aggaccgaag atgagtactt ggaggagctc
721 accgagtgca ttcgggatca tcgattgcta tggactatg ttgacgcatt gcgacctgtc
781 ttttcgggaa ccattcttgt gcagttcctc ctgacgggta ctgtactggg tctctcaatg
841 ataaatctaa cgtctctctc gacattttcg actggtgtcg ccacttgcct tctctatgtc
901 gacgtgtcca tggagacgtt ccccttttgc tatttctgca acatgattat cgatgactgc
961 caggaaatgt ccaattgect ctttcaatcg gactggacct ctgccgatcg tcgctacaaa
1021 tccacttttg tatactttct ccacaatctt cagcaacca ttactctcac ggctggctga
1081 gtgtttctta tttccatgca aacaaatttg gctatggtga agctggcatt tctctgtggt
1141 acggttaatt agcaatttaa cttggccgaa aggtttcaat aagttgagag ggacgagctc
1201 tgctactatt atattatata ctataattata ttacatatat attactttat attatattat
1261 gctgtacctt aataaattat tagtaataaa aaaaaaaaaa aaaaaa

```

ES 2 315 227 T3

FIG. 8E-1

DOR67 [AF127924]

LOCUS xxxxx 1321 pb ARNm INV 6 de feb de 1999
DEFINICIÓN ARNm de supuesto receptor de odorizante de *Drosophila melanogaster* DOR53, cds completa
ACCESO AF127924

PALABRAS CLAVE

FUENTE

ORGANISMO

mosca de la fruta

Drosophila melanogaster

Eukaryota: Metazoa; Arthropoda; Tracheata; Hexapoda;
Insecta; Pterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; *Drosophila*.

REFERENCIA

AUTORES

TÍTULO

REVISTA

REFERENCIA 2

AUTORES

TÍTULO

REVISTA

1 (bases 1 a 1321)

Leslie B. Vosshall, Hubert Amrein, Pavel S. Morozov,
Andrey Rzhetsky y Richard Axel.

Un mapa espacial de expresión de receptor olfatorio en
la antena de *Drosophila*

Cell (1999) en prensa

2 (bases 1 a 1321)

Vosshall, L. B., Amrein, H., Morozov, P. S., Rzhetsky,
A. y Axel R.

Presentación directa

Presentado (06-FEB-1999) Departamento de Bioquímica y
Biofísica Molecular, HHMI-Universidad de Columbia, 701
West 168th Street, Nueva York, NY 10032, Estados Unidos

CARACTERÍSTICAS

Comentarios:

Fuente

Gen

CDS

Localización/Clasificadores

/tipo_tisular = 'antena de adulto'

/nota 'supuesto receptor acoplado a proteínas G
con siete dominios transmembrana'

1..1321

/organismo = "*Drosophila melanogaster*"

/cepa = "Oregon R"

/db_xref = "taxón: 7227"

/cromosoma = "II"

/mapa = "22A3"

1..1321

/nota = "receptor de odorizante"

/gen = "DOR67"

102..1202

/gen = "DOR67"

/nota = "receptor de odorizante"

/codón_inicio = 1

/producto = "DOR67"

/traducción = "MWSFGWTVPENKRWDLHYKLWSTFVTLVIFILLPISVSVEYIQR

FKTFSAGEFLSSIQIGVNMYGSSFKSYLTMMGYKKRQEAQMSLDKRCVCDEERTI

VHRHVALGNFCYIFYHIAYSFLISNFLSFIMKRIHAWRMYPYVDPEKQFYISSIAE

VILRGWAVFMDLCTDVCPLISMVIARCHITLLKQRLRLNRSEPGRTEDYELKELADCV

RDHRLILDYVDALRSVFSGTIFVQFLLIGIVLGLSMINIMFFSTLSTGVAVVLFMSCV

SMQTFPFCYLCNMIMDDCQEMADSLFQSDWTSADRRYKSTLVYFLHNLQQPIILTAGG

VFPISMQTNLMVMVLAFTVVTIVKQFNLAKEFKQ"

RECuento DE BASES 348 a 281 c 294 g 398 t

FIG. 8E-2

ORIGEN

```

1  ggacgagga aatgttaagc cagttcttcc cccacattaa agaaaagcca ttgagcagac
61  gggcttaagtc ccgagatgcc ttcgtcttacc tagatcgggt gacgtgggcc ttcggcttga
121  cagtgccctga aaacaaaagg tgggacctac attacaaact gtcgtcaact ttcgtgacac
181  tggtgatatt cttctctctg ccgataccgg taagcgttga gtatattcag cggcttcaaga
241  cctctctcggc gggtgagttc cttagctcaa tccagattcg cgttaacatg tacggaagca
301  gcttcaaaaag ttatttgacc atgatgggat ataagaagag acaggaggct aagatgtcac
361  tggatgagct ggacaagaga tgcgtctgtg atgaggagag gaccattgta catcgacatg
421  tccgccctggg aaactcttgc tatattttct attacattgc gtacactagc tttctgacct
481  caaactcttct gtcatttata atgaagagaa tccatgcctg gcgcatgtac tttccctacg
541  tccgaccccgaa aaagcaattt tacatctcta gcacgcgcca agtcattctt agggggctggg
601  ccgtcttccat ggcctctctg acggatgtgt gtcctctgat ctccatggta atagcacgat
661  gccacatcac ccttctgaaa cagcgccctg gaaatctacg atcggaaacca ggaaggacgg
721  aagatgagta cttgaaggag ctccgcgact gcgttcgaga tcaccgcttg atattggact
781  atgctgacgc attgcgatcc gtctttctcg ggacaatttt tgcgcagttc ctcttgatcg
841  gtattgtact gggctctgtca atgataaata taatgttttt ctcaacactt tcgactggcg
901  tcgcccgtgt ccttttctatg tccgtcggtat ctatgcagac gtctccctct tgcatttgt
961  gtaacatgat catggatgac tgccaagaga tggccgactc cctttttcaa tcggactgga
1021  catctgccga tgcctgctac aaatccactt tggatatact tcttcacaa cttcagcagc
1081  ccattattct tacggctggg ggagtcttct ctatttccat gcaaacaaat tcaaatatgg
1141  tgaagctggc ctttactgtg gttacaatag tgaacaatt caacttggca gaaaagtctc
1201  aataagctaa gatatgcaag ctctgctatt ataaacctac actcgagaaa atattctctc
1261  acattaataa accttcagta cttactgctt gtggcgcccc cggaaaaaaa aaaaaaaaaa
1321  a

```

ES 2 315 227 T3

FIG. 8F-1

DOR64 [AF127925]

LOCUS

DEFINICIÓN

ACCESO

PALABRAS CLAVE

FUENTE

ORGANISMO

REFERENCIA

AUTORES

TÍTULO

REVISTA

REFERENCIA 2

AUTORES

TÍTULO

REVISTA

CARACTERÍSTICAS

Comentarios:

Fuente

Gen

CDS

/traducción = "MKLSETLKIDYFRVQLNAWRICGALDLSEGRYWSWSMLLCILVY

LPTPMLLRGVYSFEDFVENNFSLSTVTSLSNLMKFCMYVAQLTKMVEVQSLIGQLDA
RVSGESQSERHRNMTEHLLRMSKLFQITYAVVFIIAAVPFVFETELSLPMPMWFPPDW
KNSMVAYIGALVFQEIQYVQIMQCFAADSFPPLVLYLISEQCQLLIRISEIGYGYK
TLEENEQDLVNCIRDQNALYRLLDVTKSLVSYPMVQFMVIGINIAITLFLVLIYFVET
LYDRIYYLCFLGIVTQTYPLCYGTMVQESFAELHYAVFCSNWVDQSASYRGHMLIL
AERTKRMQLLLAGNLVPIHLSTYVACWKGAYSFFTLMDRDLGLS"

RECuento DE BASES 331 a 314 c 305 g 358 t

xxxxx 1308 pb ARNm INV 6 de feb de 1999
ARNm de supuesto receptor de odorizante de *Drosophila melanogaster* DOR53, cds completa
AF127925

mosca de la fruta

Drosophila melanogaster

Eukaryota: Metazoa; Arthropoda; Tracheata; Hexapoda;
Insecta; Pterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; *Drosophila*.

1 (bases 1 a 1308)

Leslie B. Vosshall, Hubert Amrein, Pavel S. Morozov,
Andrey Rzhetsky y Richard Axel.

Un mapa espacial de expresión de receptor olfatorio en
la antena de *Drosophila*

Cell (1999) en prensa

2 (bases 1 a 1308)

Vosshall, L. B., Amrein, H., Morozov, P. S., Rzhetsky,
A. y Axel R.

Presentación directa

Presentado (06-FEB-1999) Departamento de Bioquímica y
Biofísica Molecular, HHMI-Universidad de Columbia, 701
West 168th Street, Nueva York, NY 10032, Estados Unidos

Localización/Clasificadores

/tipo_tisular = 'antena de adulto'

/nota 'supuesto receptor acoplado a proteínas G
con siete dominios transmembrana'

1..1308

/organismo = "*Drosophila melanogaster*"

/cepa = "Oregon R"

/db_xref = "taxón: 7227"

/cromosoma = "II"

/mapa = "23A1-2"

1..1308

/nota = "receptor de odorizante"

/gen = "DOR64"

22..1158

/gen = "DOR64"

/nota = "receptor de odorizante"

/codón_inicio = 1

/producto = "DOR64"

FIG. 8F-2

ORIGIN

```

1  ggcacgagcc aagaattcaa aatgaaactc agcgaaaccc taaaaatcga ctattttcga
61  gtccagttga atgcctggcg aatttctggg gccctggatc ccagcgaggg taggtactgg
121 agttgggtcga tgcctattgtg catcttgggtg tacctgccga caccatgct actgagagga
181 gtatacagtt tcgaggatccc ggtggaaaac aatttcagct cgagcctgac ggtcacatcg
241 ctgtccaatc tcatgaagtt ctgcatgtac gtggcccaac taacaaagat ggtcgaagtc
301 cagagtcctta ttggtcagct ggaatgcccgg gtctctggcg agagccagtc tgagcgtcac
361 agaaatatga ccgagcacct gctaaggatg tccaagctgt tccagatcac ctacgctgta
421 gtcttcatca ttgctgcagt tcccttcggt ttcgaaactg agctaagctt acccatgccc
481 atgtgggttcc ccttcgactg gaagaactcg atggtggcct acatcggagc tctggttttc
541 caggagattg gctatgtctt tcaaattatg caatgctttg cagctgactc gtctcccccg
601 ctcgtactgt acctgatctc cgagcaatgt caattgctga tcccgagaat cctcgaaatc
661 ggatatgggtt acaagactct ggaggagaac gaacaggatc tggtcactg catcagggat
721 caaacgcgcg tgtatagatt actcgtatgt accaagagtc tcgtttcgta tcccatgatg
781 gtgcagtcct tggctattgg catcaacatc gccatcacc ctttctgctc gatattttac
841 gtggagacct tgtacgatcg catctattat ctttgccttc tcttgggcat caccgtgcag
901 acatatccat tgtgctacta tggaaacctg gtgcaggaga gtcttgctga gcttactatc
961 gcgggtattct gcagcaactg ggtggatcaa agtgccagct atcgtgggca catgctcatc
1021 ctggcgggagc gcactaagcg gatgcagctt ctcctcgccg gcaacctggg gcccatccac
1081 ctgagcacct acgtggcctg ttggaaggga gcctactcct tcttcacctt gatggcctac
1141 cgagatggcc tgggttctta gtgcccagc catttcactc acattctaca tcaagtagta
1201 ctaccactga acacgaacac gaataatttc aaagttaaca cataatattc acaatagttg
1261 atccacttcaa taaaattctt ggttaccatg aaaaaaaaaa aaaaaaaa

```

ES 2 315 227 T3

FIG. 8G-1
DOR71g
LOCUS xxxxx 1252 pb ADN INV 7 de feb de 1999
DEFICINICIÓN Receptor de odorizante de *Drosophila melanogaster*
DOR71g, cds completa predicha con GENSCAN.

ACCESO
PALABRAS CLAVE
FUENTE mosca de la fruta
ORGANISMO *Drosophila melanogaster*
Eukaryota: Metazoa; Arthropoda; Tracheata; Hexapoda;
Insecta; Pterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; *Drosophila*.

REFERENCIA 1 (bases 1 a 1252)
AUTORES Leslie B. Vosshall, Hubert Amrein, Pavel S. Morozov,
Andrey Rzhetsky y Richard Axel.

TÍTULO Un mapa espacial de expresión de receptor olfatorio en
la antena de *Drosophila*

REVISTA Cell (1999) en prensa

REFERENCIA 2 (bases 1 a 1252)
AUTORES Celniker, S. E., Agbayani, A., Arcaina, T. T., Baxter,
E., Brazej, R. G., Butenhoff, C., Champe, M., Chavez,
C., Chew, M., Ciesiolka, L., Doyle, C. M., Farfan, D.
E., Galle, R., George, R. A., Harris, N. L., Hoskins, R.
A., Houston, K. A., Hammasti, S. R., Karra, K., Kearney,
L., Kim, E., Lee, B., Lewis, S., Li, P., Lomotan, M. A.,
Mazda, P., Moshrefi, A. R., Moshrefi, M., Nixon, K.,
Pacleb, J. M., Park, S., Pfeiffer, B., Poon, L., Punch,
E., Sequeira, A., Sethi, H., Snir, E., Svirska, R. R.,
Twomey, B., Wan, K. H., Weinburg, T., Zhang, R., Zieran,
L. L. y Rubin, G. M.

TÍTULO Secuenciación de *Drosophila melanogaster*

REVISTA Sin publicar

REFERENCIA 3 (bases 1 a 1252)
AUTORES Vosshall, L. B., Amrein, H., Morozov, P. S., Rzhetsky,
A. y Axel R.

TÍTULO Presentación directa

REVISTA Presentado (07-FEB-1999) Departamento de Bioquímica y
Biofísica Molecular, HHMI-Universidad de Columbia, 701
West 168th Street, Nueva York, NY 10032, Estados Unidos

CARACTERÍSTICAS Localización/Clasificadores

Comentarios: /tipo_tisular = 'palpo maxilar de adulto'
/nota 'supuesto receptor acoplado a proteínas G
con siete dominios transmembrana'
/nota 'anotacion de secuencia para BDGP P1 clon
DS07071; predicha usando GENSCAN; unir (26150-
24899)'

Fuente 1..1252
/organismo = "*Drosophila melanogaster*"
/db_xref = "taxón: 7227"
/cromosoma = "II"
/mapa = "33B1-2"

exón 1..731
/gen = "DOR71g"
/nota = "exón"

ARNm unir(1..731, 811..1234)
/gen = "DOR71g"

CDS unir(1..731, 811..1234)
/gen = "DOR17g"
/nota = "receptor de odorizante"
/codón_inicio = 1
/producto = "DOR71g"

FIG. 8G-2

/traducción="MVIIDSLSFYRPFWICMRLLVPTFFKDSSRPVQLVVLHILVT
LWFLPLHLLHLLLPSTAEFFKNLTMSLTCVACSLKHVAHLYHLPQIVEIESLIEQLD
TFIASEQEHRYRDHVHCHARRFTRCLYISFGMIYALFLGCVFVQVISGNWELLYPAY
FPFDLESNRLGVALGYQVFSMLVEGFQGLGNDTYTPLTLCLLAGHVHLWSIRMGQL
GYFDDETVVNHQRLLDYIEQHKLLVRFHNLVSRITISEVQLVQLGGCGATLCTIVSYML
FFVGDITISLVYYLVFFGVVVCVQLFSPCYFASEVAEELERLPYAIFSSRWYDQSRDHRF
DLLIFTQLTLGNRGWIKAGGLIELNLNAFFATLKMAYSLFAVVVRAGI"

```

gen          1..1234
              /nota = "receptor de odorizante"
              /gen = "DOR7lg"
intrón       732..810
              /gen = "DOR7lg"
              /nota = "intrón"
exón         811..1231
              /gen = "DOR7lg"
              /nota = "exón"
señal_PoliA 1246..1252
              /nota = "señal_PoliA"
RECuento DE BASES 269 a 299 c 305 g 379 t
ORIGEN
1 atgggcatta cgcacagctc tagcttttat cgtccattct ggatctgcat gcgattgctg
61 gtaccgactt ccttcaagga tccctcacgt cctgtccagc tgtacgtggc gttgctgcac
121 atccctggta ccttctgggt tccactgcat ctgctgctgc atcttctgct acttccatct
181 accgtgagtc cctttaagaa cctgaccatg tctctgactt gctgtggcctg cagtctgaag
241 catgtggccc acctgtatca cctgcccagc attgtggaaa tcgaatcact gatcggagcaa
301 ttagacacat ttattgccag cgaacaggag catcgttact accgggatca cgtacattgc
361 catgctaggc gctttacaag atgtctctat attagctttg gcatgatcta tgcgctcttc
421 ctgttcggcg tcttcgttca ggttattagc ggaaattggg aacttctcta tccagcctat
481 tttccattcg acttggagag caatcgcttc ctgggcgcag tagccttggg ctatcaggta
541 ttcagcatgt tagttgaagg cttccagggg ctgggcaacg atacctatac cccactgacc
601 ctatgccttc tggccggaca tgtccatttg tggcccatc gaattgggtca actgggatac
661 ttcgatgacg agacggtggt gaatcatcag cgtttgctgg attacattga gcagcataaa
721 cttcttggtg ggtaagcttc gatcaactaa cttttgacaa gaagtttatt cactttaact
781 ggttccaaaa acgatgcact caatgtgcag attccacaac ctggtgagcc ggaccatcag
841 cgaagtgcga ctggtgcagc tgggcggatg tggagccact ctgtgcatca ctgtctctta
901 catgctcttc tttgtgggag acacaatctc gctgggtctac tacttgggtg tctttggagt
961 ggtctgctg cagctctttc ccagctgcga ttttgcagc gaagttagccg aggagctgga
1021 acggctgcga tatgcgatct tctccagcag atggtagcat caatcgccgg atcatcgatt
1081 cgattctgctt tctcttacac aactaacact gggaaaccgg gggtaggata ccaaggcagg
1141 aggtcttctc gagctgaatt tgaatgcctt tttcgcacc ctgaagatgg cctattcccc
1201 tcttcgagtt gtggtgcggg caaagggtat atagagagtc tgttcaatta aa

```

ES 2 315 227 T3

FIG. 8H-1

DOR72g

LOCUS

DEFINICIÓN

xxxxxx 1321 pb ADN INV 8 de feb de 1999
Receptor de odorizante de *Drosophila melanogaster*
DOR72g, cds completa predicha con GENSCAN.

ACCESO

PALABRAS CLAVE

FUENTE

ORGANISMO

mosca de la fruta

Drosophila melanogaster

Eukaryota: Metazoa; Arthropoda; Tracheata; Hexapoda;

Insecta; Pterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; *Drosophila*.

REFERENCIA

AUTORES

1 (bases 1 a 1321)

Leslie B. Vosshall, Hubert Amrein, Pavel S. Morozov,
Andrey Rzhetsky y Richard Axel.

TÍTULO

Un mapa espacial de expresión de receptor olfatorio en
la antena de *Drosophila*

REVISTA

Cell (1999) en prensa

REFERENCIA 2

AUTORES

2 (bases 1 a 1321)

Celniker, S. E., Agbayani, A., Arcaina, T. T., Baxter,
E., Brazej, R. G., Butenhoff, C., Champe, M., Chavez,
C., Chew, M., Ciesiolka, L., Doyle, C. M., Farfan, D.
E., Galle, R., George, R. A., Harris, N. L., Hoskins, R.
A., Houston, K. A., Hammasti, S. R., Karra, K., Kearney,
L., Kim, E., Lee, B., Lewis, S., Li, P., Lomotan, M. A.,
Mazda, P., Moshrefi, A. R., Moshrefi, M., Nixon, K.,
Pacleb, J. M., Park, S., Pfeiffer, B., Poon, L., Punch,
E., Sequeira, A., Sethi, H., Snir, E., Svirskas, R. R.,
Twomey, B., Wan, K. H., Weinburg, T., Zhang, R., Zieran,
L. L. y Rubin, G. M.

TÍTULO

Secuenciación de *Drosophila melanogaster*

REVISTA

Sin publicar

REFERENCIA

AUTORES

3 (bases 1 a 1321)

Vosshall, L. B., Amrein, H., Morozov, P. S., Rzhetsky,
A. y Axel R.

TÍTULO

Presentación directa

REVISTA

Presentado (08-FEB-1999) Departamento de Bioquímica y
Biofísica Molecular, HHMI-Universidad de Columbia, 701
West 168th Street, Nueva York, NY 10032, Estados Unidos

CARACTERÍSTICAS

Comentarios:

/tipo_tisular = 'antena de adulto'

/nota 'supuesto receptor acoplado a proteínas G
con siete dominios transmembrana'

/nota 'anotación de secuencia para BDGP P1 clon
DS07071; predicha usando GENSCAN; unir (27938-
26618)'

Fuente

1..1321

/organismo = "*Drosophila melanogaster*"

/db_xref = "taxón: 7227"

/cromosoma = "II"

/mapa = "33B1-2"

ARNm

unir(1..728, 910..1321)

/gen = "DOR72g"

gen

1..1321

/nota = "receptor de odorizante"

/gen = "DOR72g"

exón

<1..728

/gen = "DOR72g"

/nota = "exón"

CDS

unir(1..728, 910..1321)

/gen = "DOR72g"

/nota = "receptor de odorizante"

/codón_inicio = 1

/producto = "DOR72g"

FIG. 8H-2

/traducción = "MDLKPRVIRSEDIVRTYWLWHLGLLESNFFLNRLDLVITIFV
 TIWYPIHLILGLFMERSLGDVCKGLPITAACFFASFKFICFRFKLSEIKEIEILFKEL
 DQRALSREECEFFNQNTREANFIWKSFIVAYGLSNISAIASVLFGGGHKLLYPWFPP
 YDVQATELIFWLSVTYQIAGVSLAILQNLANDSYPPMTFCVVAGHVRLAMRLSRIGQ
 GPEETIYLTGKQLIESTEDHRKLMKIVELLRSTNMNISQLGQFISSGVNISITLVNILF
 FADNNFAITYYGVYFLSMVLELPCCYVGTLSIVEMNQLTYAIYSSNMMSMNRYSRI
 LLIFMQLTLAEVQIKAGGMIGIGMNAFFATVRLAYSFFTLAMSLR"

intrón 729..909
 /gen = "DOR72g"
 /nota = "intrón"
 exón 910..1321
 /gen = "DOR72g"
 /nota = "exón"

RECuento DE BASES 361 a 255 c 283 g 422 t

ORIGEN

```

1 atggacttaa aaccgcgagt cattcgaagt gaagatatct acagaacctt ctggttatat
61 tggcacccttc tgggctcggg aagcaatttc cttctgaatc gcttgctgga ctcgggtgatt
121 acaatttttcg taaccatttcg gtatccaatt cacctgattc tgggactgtt catggaaaga
181 cctctggggg atgtctgcaa gggctctacca attacggcag catgcttttt cgcagctttt
241 aaatttattt gtcttcgctt caagctatct gaaattaaag aaatcgaatc attatttttt
301 gagctggatc agcgagcttt aagtcgagag gaatgcgagt ttttcaatca aaatcagaga
361 cgtgaggcga atttcatttg gaaaagtctt attgtggcct atggactgtc gaatatctcg
421 gctattgcat cagttctttt cggcgggtgga cataagctat tatatcccg cctgggttcca.
481 tacgatgtgc aggccacgga actaatattt tggctaagtg caacatacca aatcgccgga

541 gtaagtcttg ccatacttca gaatttggcc aatgattcct atccaccgat gacatttttc
601 gtggttgccg gtcattgaag accttctggc atgcgcttga gtagaattgg ccaagggtcca
661 gaggaaccaa taccctaac cggaaagcaa ttaatcgaaa gcatcgagga tcaccgaaaa
721 ctaatgaagt aatgtacata tatagaatgg tttttagtta tcatcattaa atgaacgtgt
781 tgtaggaaaa ccattctgtt tgtcgggtgt caccgaaatc gatcttcttt aattttacata
841 cgatattaaa tacttctctg caaacaatta tcatattagt aatttagaat cttttattat
901 ctttccaga atagtggat tactgcgcag caccatgaat atttcgcagc tcggccagtt
961 ctttcaagt ggtgttaata tttccataac actagtcaac atttctttct ttcgggataa
1021 caatttcgct ataacctact acggagtgtt cttcttatcg atggtgttgg aattattccc
1081 gtgctgctat tacggcacc tgatatccgt ggagatgaac cagctgacct atgcgattta
1141 ctcaagtaac tggatgagta tgaatcggag ctacagccgc atcctactga ctttcatgca
1201 attcacctgc gcggaagtgc agatcaaggc cggtaggatg atcggcatcg gaatgaacgc
1261 cttctttgcc accgtgcgat tggcctact cttcttact ttcggcatgt cgtgcgtta
1321 a

```

ES 2 315 227 T3

FIG. 8I-1

DOR73g

LOCUS

DEFINICIÓN

ACCESO

PALABRAS CLAVE

FUENTE

ORGANISMO

REFERENCIA

AUTORES

TÍTULO

REVISTA

REFERENCIA 2

AUTORES

TÍTULO

REVISTA

REFERENCIA

AUTORES

TÍTULO

REVISTA

CARACTERÍSTICAS

Comentarios:

Fuente

ARNm

gen

exón

CDS

xxxxx 1212 pb ADN INV 8 de feb de 1999
Receptor de odorizante de *Drosophila melanogaster*
DOR73g, cds completa predicha con GENSCAN.

mosca de la fruta
Drosophila melanogaster
Eukaryota: Metazoa; Arthropoda; Tracheata; Hexapoda;
Insecta; Pterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; *Drosophila*.

1 (bases 1 a 1212)

Leslie B. Vosshall, Hubert Amrein, Pavel S. Morozov,
Andrey Rzhetsky y Richard Axel.

Un mapa espacial de expresión de receptor olfatorio en
la antena de *Drosophila*

Cell (1999) en prensa

2 (bases 1 a 1212)

Celniker, S. E., Agbayani, A., Arcaina, T. T., Baxter,
E., Brazej, R. G., Butenhoff, C., Champe, M., Chavez,
C., Chew, M., Ciesiolka, L., Doyle, C. M., Farfan, D.
E., Galle, R., George, R. A., Harris, N. L., Hoskins, R.
A., Houston, K. A., Hammasti, S. R., Karra, K., Kearney,
L., Kim, E., Lee, B., Lewis, S., Li, P., Lomotan, M. A.,
Mazda, P., Moshrefi, A. R., Moshrefi, M., Nixon, K.,
Pacleb, J. M., Park, S., Pfeiffer, B., Poon, L., Punch,
E., Sequeira, A., Sethi, H., Snir, E., Svirska, R. R.,
Twomey, B., Wan, K. H., Weinburg, T., Zhang, R., Zieran,
L. L. y Rubin, G. M.

Secuenciación de *Drosophila melanogaster*

Sin publicar

3 (bases 1 a 1212)

Vosshall, L. B., Amrein, H., Morozov, P. S., Rzhetsky,
A. y Axel R.

Presentación directa

Presentado (08-FEB-1999) Departamento de Bioquímica y
Biofísica Molecular, HHMI-Universidad de Columbia, 701
West 168th Street, Nueva York, NY 10032, Estados Unidos

Localización/Clasificadores

/tipo_tisular = 'antena de adulto'
/nota 'supuesto receptor acoplado a proteínas G
con siete dominios transmembrana'
/nota 'anotación de secuencia para BDGP P1 clon
DS07071; predicha usando GENSCAN; unir (29614-
28403) '

1..1212
/organismo = "*Drosophila melanogaster*"
/db_xref = "taxón: 7227"
/cromosoma = "II"
/mapa = "33B1-2"

unir(1..722, 798..1212)
/gen = "DOR73g"

1..1212
/nota = "receptor de odorizante"
/gen = "DOR73g"

<1..722
/gen = "DOR73g"
/nota = "exón"

unir(1..722, 798..1212)
/gen = "DOR73g"
/nota = "receptor de odorizante"
/codón_inicio = 1
/producto = "DOR73g"

FIG. 8I-2

```

/traducción = "MSRRKVRSENLYKTYWLYWRLLGVEGDYPFRRLVDFITTSFIT
ILFPVHLILGMYKKPQIQVFRSLHFTSECLFCSYKFFCFRWKLKEIKTIEGLLQDLDS
RVESEEEERNYFNQNPSPRVARMLSKSYLVAAISAIITATVAGLFSTGRNLMYLGWFPYD
FOATAAIYWISFSYQAIGSSLLILENLANDSYPPITFCVVSGHVRLIMRLSRIGHDV
KLSSSENTKRKLEGIQDHRKLMKIIRLLRSTLHLSQLGQFLSSGINISITLINILFFA
ENNFAMLYYAVFFAAMLIELFPSCYYGILMTMEFDKLPYAFSSNWLKMDKRYNRSII
ILMQLTLVPVNIKAGGIVGIDMSAFFATVRMAYSFYTLALSFRV"

```

```

intrón      723..797
            /gen = "DOR73g"
            /nota = "intrón"
exón        798..1212
            /gen = "DOR73g"
            /nota = "exón"

```

```

RECuento DE BASES  352 a    231 c    242 g    387 t
ORIGEN

```

```

1  atggattcaa gaaggaaagt ccgaagtga aatctttaca aaacctatcg gctctactcg
61 cgacttctcg gagtcgaggg cgattatccc ttctcgacggc tagtggattc tacaatcacg
121 tctttcatta cgattttatt ccccgctgcat cttatactgg gaatgtataa aaagcccccag
181 attcaagtct tcaggagtcct gcatttcaca tcggaatgcc tttctcgag ctataagttt
241 tctctgtttc gttggaaact caaagaaata aagaccatcg aaggattgct ccaggatctc
301 gatagtcgag ttgaaagtga agaagaacgc aactacctta atcaaaaatcc aagtcgctcg
361 gctcgaatgc tctcgaaaag ttacttggtg gctgctatat cggccataat cactgcaact
421 gcagctggtt tatttagtac tggctgaaat ttaactgtac tgggttggct tccctacgac
481 tttcaagcaa ccgccgcaat ctattggatt agtttttccc atcaggctgat tggccctagt
541 ctgttgatct tggaaaatct ggccaacgat tcatatccgc cgattacatt tctgtgtggtc
601 tctggacatg tgagactatt gataatgcgt ttaagtcgaa ttggctcaca tctaaaatca
661 tcaagttcgg aaaataccag aaaactcacc gaaggtatcc aggatcacag gaaactaatg
721 aagtaagaat aaagattcaa gaaccgcatg ttctgtagct cagagaactg ataattcaac
781 aaatgtaact tctccaggat aatcacgcta cttcgagca ctttacatct tagccaactg
841 ggccagttcc tttctagtgg aatcaacatt tccataaac tcatcaacat cctgttctct
901 gctgaaaaca actttgcaat gctttattat gcggtgtctt tctgtgcaat gtttaacagaa
961 ctatttccaa gttgttacta tggaaattcg atgacaatgg agttcgataa gctaccatac
1021 gccatcttct ccagcaactg gcttaaaatg gataaaagat acaattcgac cttgataact
1081 ctgactgcaac taacactggt tccagtgaat ataaaagcag gtggtattgt tggcatcgat
1141 atgagtgcat tttctgccac agttcggatg gcataatccc ttacactttt agccctcgca
1201 tttcgagtat ag

```

ES 2 315 227 T3

FIG. 8J-1

DOR46

LOCUS

DEFINICIÓN

ACCESO

PALABRAS CLAVE

FUENTE

ORGANISMO

REFERENCIA

AUTORES

TÍTULO

REVISTA

REFERENCIA 2

AUTORES

TÍTULO

REVISTA

REFERENCIA

AUTORES

TÍTULO

REVISTA

CARACTERÍSTICAS

Comentarios:

Fuente

ARNm

gen

exón

CDS

xxxxx 1198 pb ADN INV 8 de feb de 1999
Receptor de odorizante de *Drosophila melanogaster*
DOR46g, cds completa predicha con GENSCAN.

mosca de la fruta

Drosophila melanogaster

Eukaryota: Metazoa; Arthropoda; Tracheata; Hexapoda;
Insecta; Pterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; *Drosophila*.

1 (bases 1 a 1198)

Leslie B. Voss hall, Hubert Amrein, Pavel S. Morozov,
Andrey Rzhetsky y Richard Axel.

Un mapa espacial de expresión de receptor olfatorio en
la antena de *Drosophila*

Cell (1999) en prensa

2 (bases 1 a 1198)

Celniker, S. E., Agbayani, A., Arcaina, T. T., Baxter,
E., Brazej, R. G., Butenhoff, C., Champe, M., Chavez,
C., Chew, M., Ciesiolka, L., Doyle, C. M., Farfan, D.
E., Galle, R., George, R. A., Harris, N. L., Hoskins, R.
A., Houston, K. A., Hammasti, S. R., Karra, K., Kearney,
L., Kim, E., Lee, B., Lewis, S., Li, P., Lomotan, M. A.,
Mazda, P., Moshrefi, A. R., Moshrefi, M., Nixon, K.,
Pacleb, J. M., Park, S., Pfeiffer, B., Poon, L., Punch,
E., Sequeira, A., Sethi, H., Snir, E., Svirskas, R. R.,
Twomey, B., Wan, K. H., Weinburg, T., Zhang, R., Zieran,
L. L. y Rubin, G. M.

Secuenciación de *Drosophila melanogaster*

Sin publicar

3 (bases 1 a 1198)

Voss hall, L. B., Amrein, H., Morozov, P. S., Rzhetsky,
A. y Axel R.

Presentación directa

Presentado (08-FEB-1999) Departamento de Bioquímica y
Biofísica Molecular, HHMI-Universidad de Columbia, 701
West 168th Street, Nueva York, NY 10032, Estados Unidos

Localización/Clasificadores

/tipo_tisular = 'antena de adulto'

/nota 'supuesto receptor acoplado a proteínas G
con siete dominios transmembrana'

/nota 'anotación de secuencia para BDGP P1 clon
DS07462; predicha usando GENSCAN; unir (29614-
28403)'

1..1198

/organismo = "*Drosophila melanogaster*"

/db_xref = "taxón: 7227"

/cromosoma = "II"

/mapa = "59D5-7"

unir(1..725, 784..1198)

/gen = "DOR46g"

1..1198

/nota = "receptor de odorizante"

/gen = "DOR46g"

1..725

/gen = "DOR46g"

/nota = "exón"

unir(1..725, 784..1198)

/gen = "DOR46g"

/nota = "receptor de odorizante"

/codón_inicio = 1

/producto = "DOR46g"

FIG. 8J-2

/traducción = "MAEVRVDSLEFFKSHWTAWRYLGVAHFRVENWKNLYVFYSIVSN
 LLVTLCYPVHLGISLFRNRTITEDIILNLTTFATCTACSVKCLLYAYNIKDVLEMERLL
 RLLDERVVGPEQRSIYGQVRVQLRNVLYVFIGIYMPALFAELSFLFKEERGLMYPAW
 FFFDWLHSTRNYYIANAYQIVGISFQLLQNYVSDCFPAVVLCLISSHIKMLYNRFEEV
 GLDPARDAEKDLEACITDCHKHILELFRRIEAFISLPMLIQFTVTALNVCIGLAALVFF
 VSEPMARMYFIFYSLAMPLQIFPSCFFGTDNEYWFGRLHYAAFSCNWHQTQNRSFKRKM
 MLFVEQSLKKSTAVAGGMMRIHLDTFFSTLKGAYSLFTIIRMRK"

intrón 726..783
 /gen = "DOR46g"
 exón 784..1198
 /gen = "DOR46g"

RECuento DE BASES 251 a 325 c 310 g 312 t

ORIGEN

```

1 atggcagagg tcagagtggg cagctctggag tttttcaaga gccattggac cgcctggcgg
61 tacttgggag tggctcattt ccgggtcgag aactggaaga acctttacgt gctttacagc
121 attgtgtcga atcttctcgt gacctgtgtc taccctgttc acctgggaat atccctctct
181 cgtaacccga ccatcaccca ggacatcttc aacctgacca cctttgcgac ctgcacagcc
241 tgttcgggtga agtgcctgct ctacgcctac aacatcaagg atgtgctgga gatggagcgg
301 ctgttgaggc tcttggatga acgcgtcggt ggtccggagc aacgcagcat ctacggacaa
361 gtgaggggtcc agctgcgaaa tgtgctatag gtgttcacat gcattctacat gccgtgtgccc
421 ctgttcgccc agctatcctt tctgttcaag gaggagcgcg gtctgatgta tcccgcctgg
481 tttccctctc actggctgca ctccaccagg aactattaca tagcgaacgc ctatcacata
541 gtgggcatct cgtttcagct gctgcaaaac tatgttagcg actgctttcc ggcggtggtg
601 ctgtgcctga tctcatccca catcaaaatg ttgtacaaca gattccgagga ggtgggcttg
661 gatccagcca gagatgcgga gaaggacctg gaggcctgca tcaccgatca caagcatatt
721 ctagagtggtg caggcggctc attgtaacct tcgtgttcta ttcactttcc aacttttttc
781 cagactattc cgacgcctcg aggccttcac ttccctgccc atgctaattc agttcacagt
841 gaccgccttg aatgtgtgca tcgggtttagc agccctggtg tttttcgta gcgagcccat
901 ggcacggatg tactttcatt tctactccct ggccatgccg ctgcagatct cttcgctctg
961 ctttttcggc accgacaacg agtactgggt cggacgcctc cactacgcgg ccttcagctg
1021 caattggcac acacagaaca ggagctttaa gcggaataat atgctgtctg ttgagcaatc
1081 gttgaagaag agcaccgctg cggctggcgg aatgatgcgt atccacctgg acacgtctct
1141 tttccacctc aaggggggctt actccctctt taccatcatt attcggatga gaaagtag

```

ES 2 315 227 T3

FIG. 8K-1

DOR19g

LOCUS

DEFINICIÓN

xxxxx 1293 pb ADN INV 9 de feb de 1999
Receptor de odorizante de *Drosophila melanogaster*
DOR19g, cds completa predicha con GENSCAN.

ACCESO

PALABRAS CLAVE

FUENTE

ORGANISMO

mosca de la fruta
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda;
Insecta; Pterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCIA

AUTORES

TÍTULO

REVISTA

REFERENCIA 2

AUTORES

1 (bases 1 a 1293)
Leslie B. Voss hall, Hubert Amrein, Pavel S. Morozov,
Andrey Rzhetsky y Richard Axel.
Un mapa espacial de expresión de receptor olfatorio en
la antena de *Drosophila*
Cell (1999) en prensa
2 (bases 1 a 1293)
Celniker, S. E., Agbayani, A., Arcaina, T. T., Baxter,
E., Brazej, R. G., Butenhoff, C., Champe, M., Chavez,
C., Chew, M., Ciesiolka, L., Doyle, C. M., Farfan, D.
E., Galle, R., George, R. A., Harris, N. L., Hoskins, R.
A., Houston, K. A., Hammasti, S. R., Karra, K., Kearney,
L., Kim, E., Lee, B., Lewis, S., Li, P., Lomotan, M. A.,
Mazda, P., Moshrefi, A. R., Moshrefi, M., Nixon, K.,
Pacleb, J. M., Park, S., Pfeiffer, B., Poon, L., Punch,
E., Sequeira, A., Sethi, H., Snir, E., Svirskas, R. R.,
Twomey, B., Wan, K. H., Weinburg, T., Zhang, R., Zieran,
L. L. y Rubin, G. M.

TÍTULO

REVISTA

REFERENCIA

AUTORES

TÍTULO

REVISTA

Secuenciación de *Drosophila melanogaster*
Sin publicar
3 (bases 1 a 1293)
Voss hall, L. B., Amrein, H., Morozov, P. S., Rzhetsky,
A. y Axel R.
Presentación directa
Presentado (09-FEB-1999) Departamento de Bioquímica y
Biofísica Molecular, HHMI-Universidad de Columbia, 701
West 168th Street, Nueva York, NY 10032, Estados Unidos

CARACTERÍSTICAS

Comentarios:

/nota 'supuesto receptor acoplado a proteínas G
con siete dominios transmembrana'
/nota 'anotación de secuencia para BDGP P1 clon
DS01913; predicha usando GENSCAN; unir (72218-
70926)'

Fuente

1..1293
/organismo = "*Drosophila melanogaster*"
/db_xref = "taxón: 7227"
/crómosoma = "II"
/mapa = "46F5-6"

ARNm

1..1293
/gen = "DOR19g"

gen

1..1293
/nota = "receptor de odorizante"
/gen = "DOR19g"

exón

1..1293
/gen = "DOR19g"

CDS

1..1293
/gen = "DOR19g"
/nota = "receptor de odorizante"
/codón_inicio = 1
/producto = "DOR19g"

FIG. 8K-2

/traducción = "MVTEDFYKYQVWYFQILGVWQLPTWAADHQRRFQSMRFGFILVI
 LFIMLLLSFSEMLNNISQVREILKVFFMFATEISCHAKLLHLKLSRKLGLVDAMLS
 BEFGVKSEQEMQMLELDRVAVVRMRNSYGIMSLGAASLILVPCFDNFGELPLAMLEV
 CSIEGWICYWSQYLFHSICLLPTCVLNITYDSVAYSLLCFLKVQLQMLVLRLEKLGPV
 IEPQDNEKIAMELRECAAYNRIVRFKDLVELFIKGGSVQLMCSVLVLSNLYDMST
 MSIANGDAlFMLKTCIYQLVMLWQIFIICYASNEVTQSSRLCHSIYSSQWTGWNRA
 NRIVLLMMQRFNSPMLLSTFNPTFAFSLEAFGSVGQKFLYISFITGYALLLSDRQLL
 LQLLRITAEARQQLNFETPQHLKIFKPIFKSTQNVMHVH"

RECuento DE BASES 290 a 322 c 307 g 374 t

ORIGEN

```

1  atggttacgg aggaccccca caagtaccag gtgtgtgtact tccaaatcct cggctgtctgg
61  cagctcccca cctgggcccgc agaccaccag cgtcgtcttc agtccatgag gttcggcttc
121 atcctgggtca cctcgttcac catgctcgtc cttctctcct tcgaaatggt gaacaacatt
181 tcccaagtta gggagatccc aaaggtatc ttcattgttc ccacggaaat atccctgcatg
241 gccaaattat tgcatttgaa gttgaagagc cgcacactcg ctggcttcgt tgcgcgatg
301 ctgtccccag agttcggcgt taaaagttaa caggaaatgc agatgctgga attggataga
361 gtggcgggtc tccgcattgag gaactcctac ggcattcatg cctcgggcgc ggcctccctc
421 atccttatag tccctcgttc cgacaacttc ggcgagctac cactggccat gttggaggta
481 tgcagcatcg agggatggat ctgctattgg tcgcagtacc tttccactc gatttgcctc
541 ctgcccactt gtgtgctgaa tataaacctac gactcgttgg cctactcgtt gctcgttcc
601 ttgaagggtc agctacaaat gctggctcct cgattagaaa agttgggtcc tctgaccgaa
661 ccccaggata atgagaaaaa cgcaatggaa ctgcgtgagt gtgccgccta ctacaacagg
721 attgttcgtt tcaaggacct ggtggagctg ttcataaagg ggccaggatc tgtgcagctc
781 atgtgtcttg tctcgggtgc ggtgtccaac ctgtacgaca tgtccaccat gctccattgca
841 aacggcgatg ccattcttat gctcaagacc tgtatctatc agctgggtgat gctctggcag
901 atcttcatca tttgtctacg cttcaacgag gtaactgtcc agagctctag gttgtgtcac
961 agcatctaca gctcccaatg gacgggatgg aacaggggcaa accgcccgat tgtcctctc
1021 atgatgcage gctttaattc cctgatgctc ctgagcacct ttaacccac ctttgccttc
1081 agcttggagg cctctgggtc tgtagggcag cagaaattcc ttatatatc atttactact
1141 ggttatgctc tctcctcttc agatcgtcaa ctgctcctac agctacttcg cactgctgaa
1201 gcgcgtcaac agttcaattc cgaacacccg cagcacctaa agattttcaa gccgattctt
1261 aaaagcactc aaaacgttat gcacgtacat taa

```

ES 2 315 227 T3

FIG. 8L-1

DOR24g

LOCUS

DEFINICIÓN

ACCESO

PALABRAS CLAVE

FUENTE

ORGANISMO

REFERENCIA

AUTORES

TÍTULO

REVISTA

REFERENCIA 2

AUTORES

TÍTULO

REVISTA

REFERENCIA

AUTORES

TÍTULO

REVISTA

CARACTERÍSTICAS

Comentarios:

Fuente

ARNm

gen

CDS

xxxxx 2075 pb ADN INV 9 de feb de 1999
Receptor de odorizante de *Drosophila melanogaster*
DOR24g, cds completa predicha con GENSCAN.

mosca de la fruta

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda;
Insecta; Pterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; *Drosophila*.

1 (bases 1 a 2075)

Leslie B. Vosshall, Hubert Amrein, Pavel S. Morozov,
Andrey Rzhetsky y Richard Axel.

Un mapa espacial de expresión de receptor olfatorio en
la antena de *Drosophila*

Cell (1999) en prensa

2 (bases 1 a 2075)

Celniker, S. E., Agbayani, A., Arcaina, T. T., Baxter,
E., Brazej, R. G., Butenhoff, C., Champe, M., Chavez,
C., Chew, M., Ciesiolka, L., Doyle, C. M., Farfan, D.
E., Galle, R., George, R. A., Harris, N. L., Hoskins, R.
A., Houston, K. A., Hammasti, S. R., Karra, K., Kearney,
L., Kim, E., Lee, B., Lewis, S., Li, P., Lomotan, M. A.,
Mazda, P., Moshrefi, A. R., Moshrefi, M., Nixon, K.,
Pacleb, J. M., Park, S., Pfeiffer, B., Poon, L., Punch,
E., Sequeira, A., Sethi, H., Snir, E., Svirskas, R. R.,
Twomey, B., Wan, K. H., Weinburg, T., Zhang, R., Zieran,
L. L. y Rubin, G. M.

Secuenciación de *Drosophila melanogaster*

Sin publicar

3 (bases 1 a 2075)

Vosshall, L. B., Amrein, H., Morozov, P. S., Rzhetsky,
A. y Axel R.

Presentación directa

Presentado (09-FEB-1999) Departamento de Bioquímica y
Biofísica Molecular, HHMI-Universidad de Columbia, 701
West 168th Street, Nueva York, NY 10032, Estados Unidos

Localización/Clasificadores

/tipo_tisular = 'antena de adulto'

/nota 'supuesto receptor acoplado a proteínas G
con siete dominios transmembrana'

/nota 'anotación de secuencia para BDGP P1 clon
DS00724; predicha usando GENSCAN; unir (62623-
64697)'

1..2075

/organismo = "*Drosophila melanogaster*"

/db_xref = "taxón: 7227"

/cromosoma = "II"

/mapa = "47D6-E2"

unir(1..185, 739..1038, 1338..1953, 2025..2075)

/gen = "DOR24g"

1..2075

/nota = "receptor de odorizante"

/gen = "DOR24g"

unir(1..185, 739..1038, 1338..1953, 2025..2075)

/gen = "DOR24g"

/nota = "receptor de odorizante"

/codón_inicio = 1

/producto = "DOR24g"

FIG. 8L-2

/traducción = "MRPTLQVLISVLCVLSAYAWDHTDCNDHYIEFMDYPDERATAYS
 NESSEWDFEFWRQVFLQVQKSTIALLGDFSENREMWKRPYRAMNVFSIAAIF
 PFILAAVLHNWKNVLLADAMVALLITILGLFKFSMILYLRDFKRLIDKFRLLMSNG
 EFFPWNHIIIRNYVLSFIWSAFASGTGVVLPVSLDTIFCSFTSNLCAFFKIAQYKVVR
 FKGGSLKESQATLNKVVALYQTSMDMCDLNQCYQPIICAQFFISSLQCLMLGYLFSI
 TFAQTEGVVYASFIATIIQAYIYCYCCGENLKTESASFWEAIYDSPWHESLGAGGAST
 SICRSLISMRAHRGFRITGYFFEANMEAFSSIVRTAMSYYTMLRSFS"

exón	1..185
	/gen = "DOR24g"
intrón	186..738
	/gen = "DOR24g"
exón	739..1038
	/gen = "DOR24g"
intrón	1039..1337
	/gen = "DOR24g"
exón	1338..1953
	/gen = "DOR24g"
intrón	1954..2024
	/gen = "DOR24g"
exón	2025..2075
	/gen = "DOR24g"

RECuento DE BASES 548 a 460 c 447 g 620 t

ORIGEN

```

1 atgcgcccaa cactccaggt cctcatcagt gtcctttgtc tggctctctgc ctacgcctgg
61 gatcataccg actgtaacga tcactatata gaattcatcg actatcccga tgagagagct
121 accgcctata gcaatgaate ctcagaatgg gatttctttg aattttggag gcaagtgttc
181 ggcctgtagc aattctgaca tcgctgacca cgctctctgt caaagcaatc acttggcgag
241 acatgtgacc ctgggatgtc agggcgtgaa taaagaacat taaaatttaa ccggagttcc
301 cacttactat gcctaaagag gtgtcaaacg gaggtattat ttgggcgtga tatattacac
361 cctctaaagg agtatcccca ccagggtgaaa cctataaaaa cctctgacgt cgtcaatgga
421 aagtactgaa ccggaaatga aagtgggcac tttctaatgg tagaaaattg tcgggatgaa
481 atataatagt agatataata aatattatat cttgattatc cagcatcaat cacaagataa
541 aaaaaaata taattcataa ttcataaatt acatgtaata ggcatttcta aatgctgtaa
601 acaaaaggag ggtattaaag agctgtactt tcgcatctgt atttcattta ttcctactca
661 atcgaagggt ttaaaacccc tgaactgaac acacttgact tagtgtgagg ccgaattaac
721 ccttctctgc atggacagtt tcttgcaagt acagaagagc accatttccc tctctgggtc
781 tgatctcttt agtgaaaaat gagaaatgtg gaaacgcccc tatagagcaa ggaatgtgtc
841 tagcatagct gccatttttc cctttatcct ggcagctgtg cttcataatt ggaagaatgt
901 atctgtctgt gccgatgcca tgggtggccc actaataacc attctggggt tattcaagtc
961 tagcatgata ctttacttac gtctgcgatt caagcgactg attgacaaat tccgttctgt
1021 catgtcgaat ggtgagttgt aatccatttc ggccagaatg tgcattctt catttattat
1081 tttatagagg cggaaacagg cgaggaatac gccgagattc tcaacgcagc aaacaagcag
1141 gatcaacgaa tgtgcaactc gtttaggact tgtttctccc tcgcttgggc cttgaatagt
1201 gtctctgccc tcgtgagaat gggctctcag tatctggttag caggctcatgc agagcccag
1261 ttgccttttc cctgtctgta tgtacaaatg atatatatga tatatggtga tcaagttatc
1321 aggcctttgt cctaaagtct tccctggaat atccacatca ttccgaattc tgttttgagc
1381 ttcactctga ggcctttctc ctcgacaggt ggggtttctc ctgctgtcag cttggatacc
1441 atattctggt ccttcaccag caacctgtgc gccttcttca aaattgcgca gtacaaggtg
1501 gtttagattc agggcggtac ccttaaagaa tcacaggcca cattgaacaa agtctttgcc
1561 ctgtaccaga ccagcttggc tatgtgcaac gatctgaatc agtgctacca accgattatc
1621 tgcgcccagt tcttcatttc atctctgcaa cctctgcattc tgggatattc gttctccatt
1681 acttttgccc agacagaggg cgtctactac gcctcattca tagccacaat cattatataa
1741 gccatatact actgctactg cggggagaa cctgaagacg agagtgccag cttcgagtgg
1801 gccatctacg acagtccgtg gcacgagagt tgggtgctg gtggagcttc tacctcgatc
1861 tgccgatccc tgcctatcag catgatgcgg gctcatcggg gattccgcat tacgggatcc
1921 tttttcgagg caaacatgga ggccttctca tgggtgggtg aatcatttcc attgtacaat
1981 acatggattc acatgaatac tctcttctaa ctttccgttc ttagattgtc cgcacggcga
2041 tgcctacat cacaatgctg agatcattct cctaa

```

FIG. 9

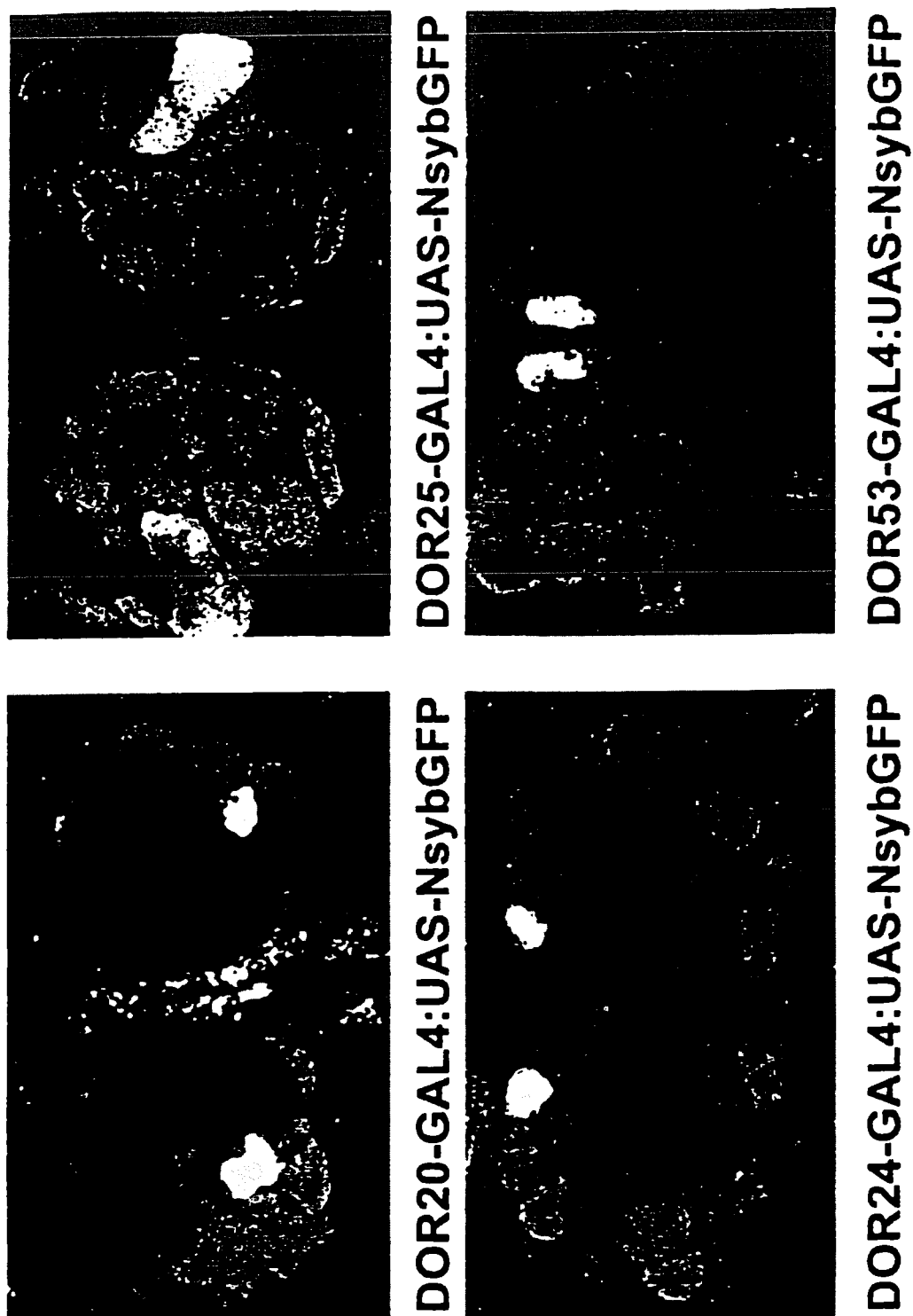
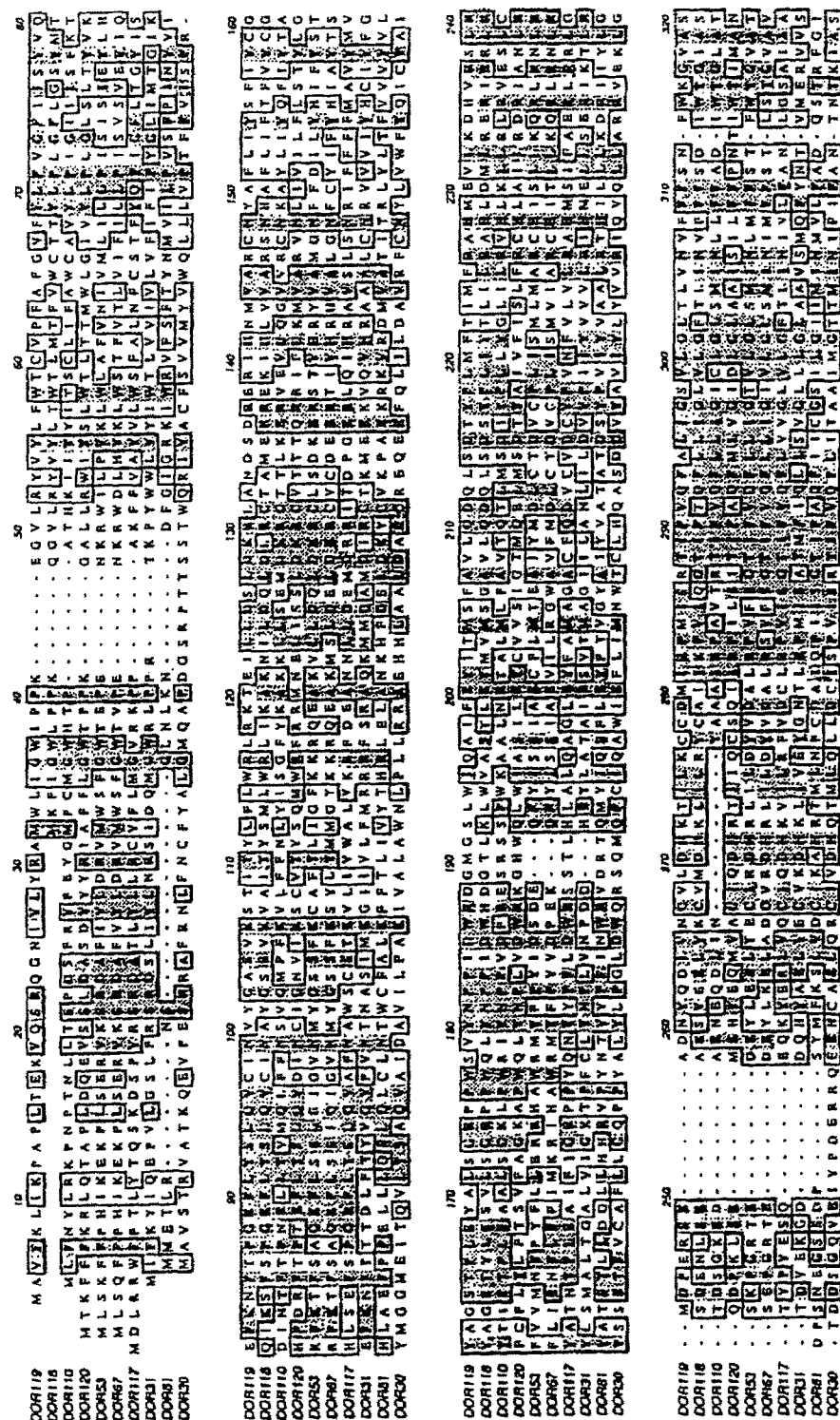


FIG. 10A-1

Alineamiento de GlustalW de la Familia DOR53



ES 2 315 227 T3

LISTA DE SECUENCIAS

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

5 <120> GENES QUE CODIFICAN RECEPTORES DE ODORIZANTES DE INSECTO Y USOS DE LOS MISMOS

<130> 0575/58715

10 <140> 09/257,706

<141> 1999-02-25

<160> 24

15

<170> PatentIn Ver. 2.0

<210> 1

20

<211> 553

<212> DNA

<213> *Drosophila melanogaster* DOR62

25

<400> 1

30 caggaaactca tcgagtgcac ccgcgatctg gcgcgggtcc atcggctgag ggagatcatt 60
cagcgggtcc tttcagtgcc ctgcatggcc cagttcgctc gctccgccgc cgtccagtgt 120
accgtcgcca tgcacttcct gtacgtagcg gatgaccacg accacaccgc catgatcatc 180
tcgattgtat ttttctcggc cgtcaccttg gaggtgtttg taatctgcta ttttggggac 240
aggatgcgga cacagagcga ggcgctgtgc gatgccttct acgattgcaa ctggatagaa 300
cagctgcca agttcaagcg cgaactgctc ttcacccttg ccaggacgca gcggccttct 360
35 cttatttacg caggcaacta catcgactc tcgctggaga ccttcgagca ggatcatgag 420
ttcacatact ctgtttttcac actcttgctg agggccaagt aagaacttta taatctcttt 480
ttggggagaa aaatttttaa gcacaatagc agaaaaatat atcagataat ataacaaaaa 540
40 aaaaaaaaaa aaa 553

40

<210> 2

<211> 153

45

<212> PRT

<213> Traducción DOR62

50

55

60

65

ES 2 315 227 T3

<400> 2

```

5      Gln Glu Leu Ile Glu Cys Ile Arg Asp Leu Ala Arg Val His Arg Leu
      1              5              10              15

      Arg Glu Ile Ile Gln Arg Val Leu Ser Val Pro Cys Met Ala Gln Phe
10      20              25              30

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

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

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

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

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

      Leu Ala Arg Thr Gln Arg Pro Ser Leu Ile Tyr Ala Gly Asn Tyr Ile
35      115             120             125

      Ala Leu Ser Leu Glu Thr Phe Glu Gln Val Met Arg Phe Thr Tyr Ser
40      130             135             140

      Val Phe Thr Leu Leu Leu Arg Ala Lys
      145             150

```

<210> 3

<211> 1493

<212> DNA

<213> *Drosophila melanogaster* DOR104

ES 2 315 227 T3

<400> 3

```

5      ggcacgagca gtcgatggcc agtcttcagt tccacggcaa cgtcgatgcg gacatcaggt 60
      atgataattag cctggatccg gctagggaaat cgaatctctt ccgtctgcta atgggactcc 120
      agttggcgaa tggcacgaag ccatacgccgc ggttacccaa atggtggcca aagcggctgg 180
      aaatgattgg taaagtgctg cccaaagcct attgttccat ggtgattttc acctccctgc 240
      atttgggtgt cctgttcacg aaaaccacac tggatgtcct gccgacgggg gagctgcagg 300
10     ccataacgga tgccctcacc atgaccataa tatacttttt cacgggctac ggcaccatct 360
      actggtgcct gcgtctcccg cgctctcttg cctacatgga gcacatgaac cgggagtatc 420
      gccatcattc gctggccggg gtgacctttg tgagtagcca tgcggccttt aggatgtcca 480
      gaaacttcac ggtgggtgtg ataattgcct gcctgctggg cgtgatttcc tggggcggtt 540
15     cgccactgat gctgggcatc cggatgctgc cgctccaatg ttggtatccc ttcgacgccc 600
      tgggtcccg caccatatac gcggtctatg ctacacaact tttcggtcag atcatggtgg 660
      gcatgacctt tggattcggg ggatcactgc ttgtcaccct gagcctgcta ctctggggac 720
      aattcgatgt gctctactgc agcctgaaga acctggatgc ccataccaag ttgctgggcg 780
20     gggagtctgt aaatggcctg agttcgctgc aagaggagt gctgctgggg gactcgaaga 840
      gggaattaaa tcagtacgtt ttgctccagg agcatccgac ggatctgctg agattgtcgg 900
      caggacgaaa atgtcctgac caaggaaatg cgtttcacia cgcttgggtg gaatgcattc 960
      gcttgcatcg ctctattctg cactgctcac aggagtgtga gaatctattc agtccatatt 1020
25     gtctgggtcaa gtcactgcag atcacctttc agctttgcct gctgggtctt gtgggcggtt 1080
      cgggtactcg agaggctctg cggattgtca accagctaca gtacttggga ctgaccatct 1140

      tcgagctcct aatgttcacc tattgtggcg aactcctcag tcggcatagt attcgatctg 1200
      gcgacgcctt ttggaggggt gcgtgggtgga agcacgcca ttccatccgc caggacatcc 1260
30     tcactcttctt ggtcaatagt agacgtgcag ttcacgtgac tgccggcaag ttttatgtga 1320
      tggatgtgaa tcgtctaaga tcggttataa cgcaggcggt cagcttcttg actttgctgc 1380
      aaaagtggc tgccaagaag acggaatcgg agctctaacc tggtagcacg catcgatatt 1440
      tatttagcgc attaaaaaaaa agtcgagtaa aagcaaaaaa aaaaaaaaaa aaa 1493

```

<210> 4

<211> 467

<212> PRT

<213> Traducción DOR104

ES 2 315 227 T3

<400> 4

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

ES 2 315 227 T3

	Phe Asp Ala Leu Gly Pro Gly Thr Tyr Thr Ala Val Tyr Ala Thr Gln	
	195	200 205
5	Leu Phe Gly Gln Ile Met Val Gly Met Thr Phe Gly Phe Gly Gly Ser	
	210	215 220
10	Leu Phe Val Thr Leu Ser Leu Leu Leu Leu Gly Gln Phe Asp Val Leu	
	225	230 235 240
15	Tyr Cys Ser Leu Lys Asn Leu Asp Ala His Thr Lys Leu Leu Gly Gly	
	245	250 255
20	Glu Ser Val Asn Gly Leu Ser Ser Leu Gln Glu Glu Leu Leu Leu Gly	
	260	265 270
25	Asp Ser Lys Arg Glu Leu Asn Gln Tyr Val Leu Leu Gln Glu His Pro	
	275	280 285
30	Thr Asp Leu Leu Arg Leu Ser Ala Gly Arg Lys Cys Pro Asp Gln Gly	
	290	295 300
35	Asn Ala Phe His Asn Ala Leu Val Glu Cys Ile Arg Leu His Arg Phe	
	305	310 315 320
40	Ile Leu His Cys Ser Gln Glu Leu Glu Asn Leu Phe Ser Pro Tyr Cys	
	325	330 335
45	Leu Val Lys Ser Leu Gln Ile Thr Phe Gln Leu Cys Leu Leu Val Phe	
	340	345 350
50	Val Gly Val Ser Gly Thr Arg Glu Val Leu Arg Ile Val Asn Gln Leu	
	355	360 365
55	Gln Tyr Leu Gly Leu Thr Ile Phe Glu Leu Leu Met Phe Thr Tyr Cys	
	370	375 380
60	Gly Glu Leu Leu Ser Arg His Ser Ile Arg Ser Gly Asp Ala Phe Trp	
	385	390 395 400
65	Arg Gly Ala Trp Trp Lys His Ala His Phe Ile Arg Gln Asp Ile Leu	
	405	410 415
	Ile Phe Leu Val Asn Ser Arg Arg Ala Val His Val Thr Ala Gly Lys	
	420	425 430
	Phe Tyr Val Met Asp Val Asn Arg Leu Arg Ser Val Ile Thr Gln Ala	
	435	440 445

ES 2 315 227 T3

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

5 Ser Glu Leu
465

<210> 5

<211> 1556

<212> DNA

<213> *Drosophila melanogaster* DOR87

<400> 5

```

ggcacgaggc ttatagaaag tgccgagcaa tgacaatcga ggatatcggc ctggtgggca 60
tcaacgtgcg gatgtggcga cacttggccg tgcgtaccc cactccgggc tccagctggc 120
gcaagttcgc cttcgtgctg ccggtgactg cgatgaatct gatgcagttc gtctacctgc 180
tgcggtatgt gggcgacctg cccgccttca ttctgaacat gttcttcttc tcggccattt 240
tcaacgcccc gatgcgcacg tggctggtea taatcaagcg gcgccagttc gaggagtttc 300
tcggccaact ggccactctg ttccattcga ttctcgactc caccgacgag tgggggctgt 360
gcatcctgcg gagggcgga cgggaggctc ggaacctggc catccttaat ttgagtgcct 420
ccttcctgga cattgtcggg gctctgggat cggcgctttt cagggaggag agagctcatc 480
ccttcggcgt agctctacca ggagtgaaca tgaccagttc acccgctctc gaggttatct 540
acttggtcca actgcctacg cccctgctgc tgtccatgat gtacatgcct ttcgtcagcc 600
tttttgcgg cctggccatc tttgggaagg ccattgctga gatcctggta cacaggctgg 660
gccagattgg cggagaagag cagtcggagg aggagcgctt ccaaaggctg gcctcctgca 720
ttgcgtacca cagcgagggt atgcgctatg tgtggcagct caacaaactg gtggccaaca 780
ttgtggcggg ggaagcaatt atttttggct cgataatctg ctactgctc ttctgtctga 840
atattataac ctcaccacac caggtgatct cgatagtgat gtacattctg accatgctgt 900
acgtttctct cactactac aatcgggcca atgaaatatg cctcgagaac aaccgggtgg 960
cggaggctgt ttacaatgtg ccctggtagc aggcaggaa cgggtttcgc aaaaccctcc 1020
tgatcttctt gatgcaaaca caacaccoga tggagataag agtcggcaac gtttacccca 1080
tgacattggc catgttccag agtctgttga atgcgtccta ctctacttt accatgctgc 1140
gtggcgctac cggcaaatga gctgaaagac cgaaaaaacc ggagtatccc ctcccatatt 1200
ccccctgctc ctttattttt ctttcctttt ccttttccgt tttcccatc gcttttccag 1260
caatccgggt aatgcaaaaa gttgttgctg gctgtggctc tggctgcttg tttggcattt 1320
gcatatgctt gtcgtttgaa aggatttaac cggactgctg gcacggagtc ggcatcctgg 1380
ctcctggatc ctggcatgca aatagttggc ttcttagatt gttacacaaa atagattgta 1440
gattgcagct gaatgttgtg cttggaataa agtcaaaagg atgtggagtc ggcccaaggc 1500
tctgccatt ctgtttgctc gggatgcccg aaagtatgaa aaaaaaaaaa aaaaaa 1556

```

<210> 6

<211> 376

<212> PRT

<213> Traducción DOR87

ES 2 315 227 T3

<400> 6

5	Met Thr Ile Glu Asp Ile Gly Leu Val Gly Ile Asn Val Arg Met Trp	
	1 5 10 15	
10	Arg His Leu Ala Val Leu Tyr Pro Thr Pro Gly Ser Ser Trp Arg Lys	
	20 25 30	
15	Phe Ala Phe Val Leu Pro Val Thr Ala Met Asn Leu Met Gln Phe Val	
	35 40 45	
20	Tyr Leu Leu Arg Met Trp Gly Asp Leu Pro Ala Phe Ile Leu Asn Met	
	50 55 60	
25	Phe Phe Phe Ser Ala Ile Phe Asn Ala Leu Met Arg Thr Trp Leu Val	
	65 70 75 80	
30	Ile Ile Lys Arg Arg Gln Phe Glu Glu Phe Leu Gly Gln Leu Ala Thr	
	85 90 95	
35	Leu Phe His Ser Ile Leu Asp Ser Thr Asp Glu Trp Gly Arg Gly Ile	
	100 105 110	
40	Leu Arg Arg Ala Glu Arg Glu Ala Arg Asn Leu Ala Ile Leu Asn Leu	
	115 120 125	
45	Ser Ala Ser Phe Leu Asp Ile Val Gly Ala Leu Val Ser Pro Leu Phe	
	130 135 140	
50	Arg Glu Glu Arg Ala His Pro Phe Gly Val Ala Leu Pro Gly Val Ser	
	145 150 155 160	
55	Met Thr Ser Ser Pro Val Tyr Glu Val Ile Tyr Leu Ala Gln Leu Pro	
	165 170 175	
60	Thr Pro Leu Leu Leu Ser Met Met Tyr Met Pro Phe Val Ser Leu Phe	
	180 185 190	
65	Ala Gly Leu Ala Ile Phe Gly Lys Ala Met Leu Gln Ile Leu Val His	
	195 200 205	
	Arg Leu Gly Gln Ile Gly Gly Glu Glu Gln Ser Glu Glu Glu Arg Phe	
	210 215 220	
	Gln Arg Leu Ala Ser Cys Ile Ala Tyr His Thr Gln Val Met Arg Tyr	
	225 230 235 240	
	Val Trp Gln Leu Asn Lys Leu Val Ala Asn Ile Val Ala Val Glu Ala	
	245 250 255	
	Ile Ile Phe Gly Ser Ile Ile Cys Ser Leu Leu Phe Cys Leu Asn Ile	

ES 2 315 227 T3

	260	265	270
5	Ile Thr Ser Pro Thr Gln Val	Ile Ser Ile Val Met Tyr Ile Leu Thr	
	275	280	285
10	Met Leu Tyr Val Leu Phe Thr Tyr Tyr Asn Arg Ala Asn Glu Ile Cys		
	290	295	300
15	Leu Glu Asn Asn Arg Val Ala Glu Ala Val Tyr Asn Val Pro Trp Tyr		
	305	310	315 320
20	Glu Ala Gly Thr Arg Phe Arg Lys Thr Leu Leu Ile Phe Leu Met Gln		
	325	330	335
25	Thr Gln His Pro Met Glu Ile Arg Val Gly Asn Val Tyr Pro Met Thr		
	340	345	350
30	Leu Ala Met Phe Gln Ser Leu Leu Asn Ala Ser Tyr Ser Tyr Phe Thr		
	355	360	365
35	Met Leu Arg Gly Val Thr Gly Lys		
	370	375	
40	<210> 7		
	<211> 1305		
	<212> DNA		
	<213> <i>Drosophila melanogaster</i> DOR53		
45	<400> 7		
50	ttttttcccc acataaaaaga aaagccattg agcgagcggg ttaagtcccg agatgccttc 60 atctacttgg atcgggtgat gtgggtccttt ggctggacag agcctgaaaa caaaagggtgg 120 atccttcctt ataaactgtg gtttagcgttc gtgaacatag taatgctcat ccttctgccc 180 atctcgataa gcacgcagta cctccaccga tttaaaacct tctcggcggg ggagttcctt 240 agttccctcg agattggagt caacatgtac ggaagctctt ttaagtgcgc cttcaccttg 300 attggattca agaaaagaca ggaagctaag gttttactgg atcagctgga caagagatgc 360 cttagcgata aggagaggtc cactgttcat cgctatgtcg ccatgggaaa cttttctgat 420 atcttctatc acatttttta ctccaccttc gtggtaatga acttcccgta ttttctgctt 480 gagagacgcc atgcttggeg catgtacttt ccataatctg attccgacga acagttttac 540 atctccagca tcgccgagtg ttttctgatg acggaggcca tctacatgga tctctgtacg 600 gacgtgtgtc ccttgatctc catgcttatg gctcgatgcc acatcagcct cctgaaacag 660 cgactgagaa atctccgacg gaagccagga aggaccgaag atgagtactt ggaggagctc 720 accgagtgca ttcgggatca tcgattgcta ttggactatg ttgacgcatt gcgacccgtc 780 ttttcgggaa ccatttttgt gcagttcctc ctgacggta ctgtactggg tctctcaatg 840 ataaatctaa tgttcttctc gacatttttg actggtgtcg ccacttgccct ttttatgttc 900 gacgtgtcca tggagacgtt ccccttttgc tatttgtgca acatgattat cgatgactgc 960 caggaaatgt ccaattgcct ctttcaatcg gactggacct ctgccgatcg tcgctacaaa 1020 tccactttgg tatactttct tcacaatctt cagcaaccga ttactctcac ggctgggtga 1080		
55	gtgtttccta tttccatgca aacaaatttg gctatggtga agctggcatt ttctgtggtt 1140 acggtaatta agcaatttaa cttggccgaa aggtttcaat aagttgagag ggacgagctc 1200 tgctactatt atattatata ttatattata ttatatatat attattttat attatatatt 1260 gctgtaccct aataaatatt tagtaataaa aaaaaaaaaa aaaaa 1305		
60			

ES 2 315 227 T3

<210> 8

<211> 367

<212> PRT

5 <213> Traducción DOR53

<400> 8

```

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

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

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

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

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

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

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

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

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

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

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

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

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

```

ES 2 315 227 T3

	195	200	205
5	Glu Tyr Leu Glu Glu Leu Thr Glu Cys Ile Arg Asp His Arg Leu Leu 210	215	220
10	Leu Asp Tyr Val Asp Ala Leu Arg Pro Val Phe Ser Gly Thr Ile Phe 225	230	235 240
15	Val Gln Phe Leu Leu Ile Gly Thr Val Leu Gly Leu Ser Met Ile Asn 245	250	255
20	Leu Met Phe Phe Ser Thr Phe Trp Thr Gly Val Ala Thr Cys Leu Phe 260	265	270
25	Met Phe Asp Val Ser Met Glu Thr Phe Pro Phe Cys Tyr Leu Cys Asn 275	280	285
30	Met Ile Ile Asp Asp Cys Gln Glu Met Ser Asn Cys Leu Phe Gln Ser 290	295	300
35	Asp Trp Thr Ser Ala Asp Arg Arg Tyr Lys Ser Thr Leu Val Tyr Phe 305	310	315 320
40	Leu His Asn Leu Gln Gln Pro Ile Thr Leu Thr Ala Gly Gly Val Phe 325	330	335
45	Pro Ile Ser Met Gln Thr Asn Leu Ala Met Val Lys Leu Ala Phe Ser 340	345	350
50	Val Val Thr Val Ile Lys Gln Phe Asn Leu Ala Glu Arg Phe Gln 355	360	365

<210> 9

<211> 1321

<212> DNA

<213> *Drosophila melanogaster* DOR67

<400> 9

55	ggcacgagga aatgttaagc cagttctttc cccacattaa agaaaagcca ttgagcgagc 60
	gggttaagtc ccgagatgcc ttcgtttact tagatcgggt gatgtggtcc tttggctgga 120
	cagtgcctga aaacaaaagg tgggatctac attacaaact gtggtcaact ttcgtgacat 180
	tggatgatac tatccttctg ccgatatcgg taagcggtga gtatattcag cggttcaaga 240
60	ccttctcggc ggggtgagttt cttagctcaa tccagattgg cgtaaactg tacggaagca 300
	gcttttaaaag ttatttgacc atgatgggat ataagaagag acaggaggct aagatgtcac 360
	tggatgagct ggacaagaga tgcgtttgtg atgaggagag gaccattgta catcgacatg 420
	tcgccctggg aaacttttgc tatattttct atcacattgc gtacactagc tttttgattt 480
65	caaacttttt gtcattttata atgaagagaa tccatgcctg gcgcattgtac tttccctacg 540

ES 2 315 227 T3

```

tcgaccccgga aaagcaattt tacatctcta gcatcgccga agtcattctt aggggggtggg 600
ccgtcttcat ggatctctgc acggatgtgt gtcctttgat ctccatggta atagcacgat 660
gccacatcac ccttctgaaa cagcgccgtc gaaatctacg atcggaacca ggaaggacgg 720
5 aagatgagta cttgaaggag ctcgccgact gcgttcgaga tcacegcttg atattggact 780
atgtcgacgc attgcatcc gtccttttcgg ggacaatttt tgtgcagttc ctcttgatcg 840
gtattgtact ggggtctgtca atgataaata taatgttttt ctcaacactt tcgactgggtg 900
tcgccgttgt ccttttttatg tcctgctgat ctatgcagac gttccccctt tgctatttgt 960
10 gtaacatgat tatggatgac tgccaagaga tggccgactc cctttttcaa tcggactgga 1020
catctgccga tcgtcgctac aaatccactt tgggtatactt tcttcacaat cttcagcagc 1080
ccattattct tacggctggg ggagtccttc ctatttccat gcaaacaatt ttaaataatgg 1140
tgaagctggc ctttactgtg gttacaatag tgaacaatt taacttggca gaaaagtctc 1200
15 aataagttaa gatatgcaag ctctgctatt ataaacctac actcgagaaa atattttcttc 1260
acattaataa accttcagta cttactgctt gtggcgcccc cggaxaaaaa aaaaaaaaaa 1320
a

```

```

20 <210> 10
    <211> 367
    <212> PRT
25 <213> Traducción DOR67

    <400> 10

```

```

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

His Tyr Lys Leu Trp Ser Thr Phe Val Thr Leu Val Ile Phe Ile Leu
35          20          25          30

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

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

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

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

Asp Glu Glu Arg Thr Ile Val His Arg His Val Ala Leu Gly Asn Phe
55          100          105          110

Cys Tyr Ile Phe Tyr His Ile Ala Tyr Thr Ser Phe Leu Ile Ser Asn
60          115          120          125

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

```

ES 2 315 227 T3

	Pro Tyr Val Asp Pro Glu Lys Gln Phe Tyr Ile Ser Ser Ile Ala Glu	
	145	150 155 160
5	Val Ile Leu Arg Gly Trp Ala Val Phe Met Asp Leu Cys Thr Asp Val	
	165	170 175
10	Cys Pro Leu Ile Ser Met Val Ile Ala Arg Cys His Ile Thr Leu Leu	
	180	185 190
15	Lys Gln Arg Leu Arg Asn Leu Arg Ser Glu Pro Gly Arg Thr Glu Asp	
	195	200 205
20	Glu Tyr Leu Lys Glu Leu Ala Asp Cys Val Arg Asp His Arg Leu Ile	
	210	215 220
25	Leu Asp Tyr Val Asp Ala Leu Arg Ser Val Phe Ser Gly Thr Ile Phe	
	225	230 235 240
30	Val Gln Phe Leu Leu Ile Gly Ile Val Leu Gly Leu Ser Met Ile Asn	
	245	250 255
35	Ile Met Phe Phe Ser Thr Leu Ser Thr Gly Val Ala Val Val Leu Phe	
	260	265 270
40	Met Ser Cys Val Ser Met Gln Thr Phe Pro Phe Cys Tyr Leu Cys Asn	
	275	280 285
45	Met Ile Met Asp Asp Cys Gln Glu Met Ala Asp Ser Leu Phe Gln Ser	
	290	295 300
50	Asp Trp Thr Ser Ala Asp Arg Arg Tyr Lys Ser Thr Leu Val Tyr Phe	
	305	310 315 320
55	Leu His Asn Leu Gln Gln Pro Ile Ile Leu Thr Ala Gly Gly Val Phe	
	325	330 335
60	Pro Ile Ser Met Gln Thr Asn Leu Asn Met Val Lys Leu Ala Phe Thr	
	340	345 350
65	Val Val Thr Ile Val Lys Gln Phe Asn Leu Ala Glu Lys Phe Gln	
	355	360 365

<210> 11

<211> 1308

<212> DNA

<213> *Drosophila melanogaster* DOR64

ES 2 315 227 T3

<400> 11

```

5  ggcacgagcc aagaattcaa aatgaaactc agcgaaaccc taaaaatcga ctatttttcga 60
   gtccagttga atgcctggcg aattttgtgt gccttggatc tcagcgaggg taggtactgg 120
   agttgggtcga tgctattgtg catcttgggtg tacctgccga caccatgct actgagagga 180
   gtatacagtt tcgaggatcc ggtggaaaat aatttcagct tgagcctgac ggtcacatcg 240
   ctgtccaatc tcatgaagtt ctgcatgtac gtggcccaac taacaaagat ggtcgaggtc 300
10  cagagtctta ttggtcagct ggatgcccgg gtttctggcg agagccagtc tgagcgtcat 360
   agaaatatga ccgagcacct gctaaggatg tccaagctgt tccagatcac ctacgctgta 420
   gtcttcatca ttgtgcagt tcccttcggt ttcgaaactg agctaagctt acccatgccc 480
   atgtgggtttc ctttcgactg gaagaactcg atgggtggcct acatcggagc tctgggttttc 540
15  caggagattg gctatgtctt tcaaattatg caatgctttg cagctgactc gtttcccccg 600
   ctctgactgt acctgatctc cgagcaatgt caattgctga tcctgagaat ctctgaaatc 660
   ggatatgggt acaagactct ggaggagaac gaacaggatc tgggtcaactg catcagggat 720
   caaaacgcgc tgtatagatt actcgatgtg accaagagtc tcgtttcgtg tcccatgatg 780
20  gtgcagttta tggttattgg catcaacatc gccatcacc tatttgcct gatattttac 840
   gtggagacct tgtacgatcg catctattat ctttgccttc tcttgggcac caccgtgcag 900
   acatatccat tgtgtacta tggaaccatg gtgcaggaga gttttgctga gcttcactat 960
   gcggtattct gcagcaactg ggtggatcaa agtgccagct atcgtgggca catgctcacc 1020
25  ctggcggagc gcactaagcg gatgcagctt ctctcgcg gcaacctggg gcccatccac 1080
   ctgagcacct acgtggcctg ttggaaggga gcctactccc tcttcaccct gatggccgat 1140
   cgagatggcc tgggttctta gtagccagc catttcactc acattctaca tcaagtagta 1200
   ctaccactga acacgaacac gaatatttca aaagtaaaca cataatattc acaatagtgt 1260
30  atcactttta taaaattttt ggttaccatg aaaaaaaaa aaaaaaaa 1308

```

<210> 12

<211> 379

35 <212> PRT

<213> Traducción DOR64

<400> 12

```

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

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

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

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

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

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

```


ES 2 315 227 T3

	Ile Gly Gln Leu Asp Ala Arg Val Ser Gly Glu Ser Gln Ser Glu Arg	
	100	105 110
5	His Arg Asn Met Thr Glu His Leu Leu Arg Met Ser Lys Leu Phe Gln	
	115	120 125
10	Ile Thr Tyr Ala Val Val Phe Ile Ile Ala Ala Val Pro Phe Val Phe	
	130	135 140
15	Glu Thr Glu Leu Ser Leu Pro Met Pro Met Trp Phe Pro Phe Asp Trp	
	145	150 155 160
20	Lys Asn Ser Met Val Ala Tyr Ile Gly Ala Leu Val Phe Gln Glu Ile	
	165	170 175
25	Gly Tyr Val Phe Gln Ile Met Gln Cys Phe Ala Ala Asp Ser Phe Pro	
	180	185 190
30	Pro Leu Val Leu Tyr Leu Ile Ser Glu Gln Cys Gln Leu Leu Ile Leu	
	195	200 205
35	Arg Ile Ser Glu Ile Gly Tyr Gly Tyr Lys Thr Leu Glu Glu Asn Glu	
	210	215 220
40	Gln Asp Leu Val Asn Cys Ile Arg Asp Gln Asn Ala Leu Tyr Arg Leu	
	225	230 235 240
45	Leu Asp Val Thr Lys Ser Leu Val Ser Tyr Pro Met Met Val Gln Phe	
	245	250 255
50	Met Val Ile Gly Ile Asn Ile Ala Ile Thr Leu Phe Val Leu Ile Phe	
	260	265 270
55	Tyr Val Glu Thr Leu Tyr Asp Arg Ile Tyr Tyr Leu Cys Phe Leu Leu	
	275	280 285
60	Gly Ile Thr Val Gln Thr Tyr Pro Leu Cys Tyr Tyr Gly Thr Met Val	
	290	295 300
65	Gln Glu Ser Phe Ala Glu Leu His Tyr Ala Val Phe Cys Ser Asn Trp	
	305	310 315 320
	Val Asp Gln Ser Ala Ser Tyr Arg Gly His Met Leu Ile Leu Ala Glu	
	325	330 335
	Arg Thr Lys Arg Met Gln Leu Leu Leu Ala Gly Asn Leu Val Pro Ile	
	340	345 350

ES 2 315 227 T3

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

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

<210> 13

<211> 1252

<212> DNA

<213> *Drosophila melanogaster* DOR71g

<400> 13

```

atgggtcatta tgcacagtct tagtttttat cgtccattct ggatctgcat gcgattgctg 60
gtaccgactt tcttcaagga ttccctcacgt cctgtccagc tgtacgtggc gttgctgcac 120
atccctggta ccttgtgggt tccactgcat ctgctgctgc atcttctgct acttccatct 180
accgctgagt tctttaagaa cctgaccatg tctctgactt gtgtggcctg cagtctgaag 240
catgtggccc acttgtatca cttgccgcag attgtggaaa tcgaatcact gatcgagcaa 300
ttagacacat ttattgccag cgaacaggag catcgttact atcgggatca cgtacattgc 360
catgctaggc gctttacaag atgtctctat attagctttg gcatgatcta tgcgcttttc 420
ctgttcggcg tcttcgttca gggtatttagc ggaaattggg aacttctcta tccagcctat 480
ttcccattcg acttggagag caatcgcttt ctccggcgag tagccttggg ctatcaggta 540
ttcagcatgt tagttgaagg cttccagggg ctgggcaacg atacctatac cccactgacc 600
ctatgccttc tggccggaca tgtccatttg tggccatac gaatgggtca actgggatac 660
ttcgtatgac agacggtggt gaatcatcag cgtttgctgg attacattga gcagcataaa 720
ctcttggctg ggtaagcttt gattaactaa cttttgacaa gaagtttatt cactttaact 780
ggttccaaaa acgatgcact caatgtgcag attccacaac ctggtgagcc ggaccatcag 840
cgaagtgcac ctggtgcagc tgggcggatg tggagccact ctgtgcatca ttgtctccta 900
catgctcttc tttgtggcg acacaatctc gctggtctac tacttgggtg tctttggagt 960
ggctctgcgt cagctctttc ccagctgcta ttttgccagc gaagtagccg aggagttgga 1020
acggctgcca tatgcgatct tctccagcag atggtacgat caatcgcggg atcatcgatt 1080
cgatttgctc atctttacac aattaacact gggaaaccgg gggtaggata tcaaggcagg 1140
aggctctatc gagctgaatt tgaatgcctt tttcgccacc ctgaagatgg cctattccct 1200
ttttgcagtt gtggctgcggg caaaggggat atagagagtc tgtttaatta aa 1252

```

<210> 14

<211> 384

<212> PRT

<213> Traducción DOR71g

<400> 14

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

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

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

ES 2 315 227 T3

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

ES 2 315 227 T3

	290	295	300
5	Phe Ala Ser Glu Val Ala Glu Glu Leu Glu Arg Leu Pro Tyr Ala Ile		
	305	310	315 320
10	Phe Ser Ser Arg Trp Tyr Asp Gln Ser Arg Asp His Arg Phe Asp Leu		
		325	330 335
15	Leu Ile Phe Thr Gln Leu Thr Leu Gly Asn Arg Gly Trp Ile Ile Lys		
		340	345 350
20	Ala Gly Gly Leu Ile Glu Leu Asn Leu Asn Ala Phe Phe Ala Thr Leu		
		355	360 365
25	Lys Met Ala Tyr Ser Leu Phe Ala Val Val Val Arg Ala Lys Gly Ile		
		370	375 380

<210> 15

25 <211> 1321

<212> DNA

<213> *Drosophila melanogaster* DOR72g

30 <400> 15

	atggacttaa aaccgcgagt cattcgaagt gaagatatct acagaacctt ttggttatat 60
	tggcatcttt tgggcctgga aagcaatttc tttctgaatc gcttggttga tttggtgatt 120
35	acaatttttcg taaccatttg gtatccaatt cacctgattc tgggactgtt tatggaaaga 180
	tctttggggg atgtctgcaa gggctctacca attacggcag catgcttttt cgccagcttt 240
	aaatttatatt gttttcgctt caagctatct gaaattaaag aaatcgaaat attattttaa 300
	gagctggatc agcgagcttt aagtcgagag gaatgcgagt ttttcaatca aaatcagaga 360
40	cgtgaggcga atttcatttg gaaaagtttc atttgggcct atggactgtc gaatatctcg 420
	gctattgcat cagttctttt cggcggttga cataagctat tatatcccgc ctggtttcca 480
	tacgatgtgc aggccacgga actaatattt tggctaagt taacatacca aattgccgga 540
	gtaagtcttg ccatacttca gaatttggcc aatgattcct atccaccgat gacattttgc 600
45	gtggttgccg gtcattgtaag acttttggcg atgcgcttga gtagaatttg ccaaggtcca 660
	gaggaaacaa tatacttaac cggaaagcaa ttaatcgaaa gcatcgagga tcaccgaaaa 720
	ctaagtgaagt aatgtacata tatagaatgg tttttagtta ttatcattaa atgaacgtgt 780
	tgtaggaaaa ccattctgtt tgcgggtgt caggaaatc gatttttcctt aatttacata 840
50	tgatattaaa tacttctctg caaacaatta tcatattagt aatttagaat ctttattatt 900
	tattttccaga atagtggat tactgcccag caccatgaat atttcgcagc tcggccagtt 960
	tattttcaagt ggtgttaata ttccataaac actagtcaac attctcttct ttgcggtata 1020
	taatttcgct ataacctact acggagtgt cttcctatcg atggtgttgg aattattccc 1080
55	gtgctgctat tacggcaccc tgatatccgt ggagatgaac cagctgacct atgcgattta 1140
	ctcaagtaac tggatgagta tgaatcggag ctacagccgc atcctactga tcttcattga 1200
	actcaccctg gcggaagtgc agatcaaggc cgggtgggatg attggcatcg gaatgaacgc 1260
60	cttctttgce accgtgcgat tggcctactc cttcttcact ttggccatgt cgctgcgtta 1320
	a 1321

<210> 16

65 <211> 379

<212> PRT

<213> Traducción DOR72g

ES 2 315 227 T3

<400> 16

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

ES 2 315 227 T3

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

45 <210> 17
 <211> 1212
 <212> DNA
 50 <213> *Drosophila melanogaster* DOR73g
 <400> 17

55 atggattcaa gaaggaaagt ccgaagtga aatctttaca aaacctattg gctttactgg 60
 cgacttcttg gagtcgaggg cgattatcct ttctgacggc tagtggattt tacaatcacg 120
 tctttcatta cgattttatt tcccgatgcat cttatactgg gaatgtataa aaagccccag 180
 attcaagtct tcaggagtct gcatttcaca tcggaatgcc ttttctgcag ctataagttt 240
 60 ttctgttttc gttggaaact taaagaaata aagaccatcg aaggattgct ccaggatctc 300
 gatagtcgag ttgaaagtga agaagaacgc aactacttta atcaaaatcc aagtcgtgtg 360
 gctcgaatgc ttccgaaaag ttacttggtg gctgctatat cggccataat cactgcaact 420
 gtagctgggt tatttagtac tggtcgaaat ttaatgtatc tgggttgggt tccctacgat 480
 65

ES 2 315 227 T3

```

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

```

<210> 18

<211> 378

<212> PRT

<213> Traducción DOR73g

<400> 18

```

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

```

```

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

```

```

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

```

```

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

```

```

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

```

```

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

```

```

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

```

```

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

```

```

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

```

```

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

```

ES 2 315 227 T3

	145		150		155		160										
5	Phe	Gln	Ala	Thr	Ala	Ala	Ile	Tyr	Trp	Ile	Ser	Phe	Ser	Tyr	Gln	Ala	
					165					170					175		
10	Ile	Gly	Ser	Ser	Leu	Leu	Ile	Leu	Glu	Asn	Leu	Ala	Asn	Asp	Ser	Tyr	
					180				185					190			
15	Pro	Pro	Ile	Thr	Phe	Cys	Val	Val	Ser	Gly	His	Val	Arg	Leu	Leu	Ile	
					195				200					205			
20	Met	Arg	Leu	Ser	Arg	Ile	Gly	His	Asp	Val	Lys	Leu	Ser	Ser	Ser	Glu	
		210					215					220					
25	Asn	Thr	Arg	Lys	Leu	Ile	Glu	Gly	Ile	Gln	Asp	His	Arg	Lys	Leu	Met	
		225				230					235				240		
30	Lys	Ile	Ile	Arg	Leu	Leu	Arg	Ser	Thr	Leu	His	Leu	Ser	Gln	Leu	Gly	
					245					250					255		
35	Gln	Phe	Leu	Ser	Ser	Gly	Ile	Asn	Ile	Ser	Ile	Thr	Leu	Ile	Asn	Ile	
					260					265					270		
40	Leu	Phe	Phe	Ala	Glu	Asn	Asn	Phe	Ala	Met	Leu	Tyr	Tyr	Ala	Val	Phe	
					275					280					285		
45	Phe	Ala	Ala	Met	Leu	Ile	Glu	Leu	Phe	Pro	Ser	Cys	Tyr	Tyr	Gly	Ile	
					290					295					300		
50	Leu	Met	Thr	Met	Glu	Phe	Asp	Lys	Leu	Pro	Tyr	Ala	Ile	Phe	Ser	Ser	
					305					310					315		320
55	Asn	Trp	Leu	Lys	Met	Asp	Lys	Arg	Tyr	Asn	Arg	Ser	Leu	Ile	Ile	Leu	
					325					330					335		
60	Met	Gln	Leu	Thr	Leu	Val	Pro	Val	Asn	Ile	Lys	Ala	Gly	Gly	Ile	Val	
					340					345					350		
65	Gly	Ile	Asp	Met	Ser	Ala	Phe	Phe	Ala	Thr	Val	Arg	Met	Ala	Tyr	Ser	
					355					360					365		
70	Phe	Tyr	Thr	Leu	Ala	Leu	Ser	Phe	Arg	Val							
					370					375							

<210> 19

<211> 1198

<212> DNA

<213> *Drosophila melanogaster* DOR46

ES 2 315 227 T3

<400> 19

```

5   atggcagagg tcagagtgga cagtctggag tttttcaaga gccattggac cgcctggcgg 60
    tacttgggag tggctcattt tcgggtcgag aactggaaga acctttacgt gttttacagc 120
    attgtgtcga atcttctcgt gaccctgtgc taccctcgtc acctgggaat atccctcttt 180
    cgcaaccgca ccatcaccga ggacatcctc aacctgacca cttttgcgac ctgcacagcc 240
    tgttcggtga agtgcctgct ctacgcctac aacatcaagg atgtgctgga gatggagcgg 300
10  ctggtgagggc ttttggatga acgcgtcgtg ggtccggagc aacgcagcat ctacggacaa 360
    gtgaggggtcc agctgcgaaa tgtgctatac gtgttcacgc gcattctacat gccgtgtgcc 420
    ctgttcgcgg agctatcctt tctgttcaag gaggagcgcg gtctgatgta tcccgccctgg 480
    tttcccttcg actggctgca ctccaccagg aactattaca tagcgaacgc ctatcagata 540
15  gtgggcatct cgtttcagct gctgcaaaac tatgttagcg actgctttcc ggcggtgggtg 600
    ctgtgcctga tctcatccca catcaaatg ttgtacaaca gattcgagga ggtgggcctg 660
    gatccagcca gagatgcgga gaaggacctg gaggcctgca tcaccgatca caagcatatt 720
    ctagagtggg caggcggctc attgtaacct tcgtgttcta ttcactttcc aacttttttc 780
20  cagactattc cgacgcacgc aggccttcat tccctgccc atgctaattc agttcacagt 840
    gaccgccttg aatgtgtgca tcgggttagc agccctggtg tttttcgtca gcgagcccat 900
    ggcacggatg tacttcatct tctactccct ggccatgccg ctgcagatct ttccgtcctg 960
    ctttttcggc accgacaacg agtactgggt cggacgcctc cactacgcgg ccttcagttg 1020
25  caattggcac acacagaaca ggagctttta gcggaatatg atgctgttcg ttgagcaatc 1080
    gttgaagaag agcaccgctg tggctggcgg aatgatgcgt atccacctgg acacgttctt 1140
    tccacccta aagggggcct actccctctt taccatcatt attcggatga gaaagtag 1198

```

30

<210> 20

<211> 379

<212> PRT

35

<213> Traducción DOR46

<400> 20

40

```

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

```

45

```

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

```

50

```

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

```

55

```

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

```

60

```

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

```

65

```

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

```

ES 2 315 227 T3

	Glu Met Glu Arg Leu Leu Arg Leu Leu Asp Glu Arg Val Val Gly Pro	
	100	105 110
5	Glu Gln Arg Ser Ile Tyr Gly Gln Val Arg Val Gln Leu Arg Asn Val	
	115	120 125
10	Leu Tyr Val Phe Ile Gly Ile Tyr Met Pro Cys Ala Leu Phe Ala Glu	
	130	135 140
15	Leu Ser Phe Leu Phe Lys Glu Glu Arg Gly Leu Met Tyr Pro Ala Trp	
	145	150 155 160
20	Phe Pro Phe Asp Trp Leu His Ser Thr Arg Asn Tyr Tyr Ile Ala Asn	
	165	170 175
25	Ala Tyr Gln Ile Val Gly Ile Ser Phe Gln Leu Leu Gln Asn Tyr Val	
	180	185 190
30	Ser Asp Cys Phe Pro Ala Val Val Leu Cys Leu Ile Ser Ser His Ile	
	195	200 205
35	Lys Met Leu Tyr Asn Arg Phe Glu Glu Val Gly Leu Asp Pro Ala Arg	
	210	215 220
40	Asp Ala Glu Lys Asp Leu Glu Ala Cys Ile Thr Asp His Lys His Ile	
	225	230 235 240
45	Leu Glu Leu Phe Arg Arg Ile Glu Ala Phe Ile Ser Leu Pro Met Leu	
	245	250 255
50	Ile Gln Phe Thr Val Thr Ala Leu Asn Val Cys Ile Gly Leu Ala Ala	
	260	265 270
55	Leu Val Phe Phe Val Ser Glu Pro Met Ala Arg Met Tyr Phe Ile Phe	
	275	280 285
60	Tyr Ser Leu Ala Met Pro Leu Gln Ile Phe Pro Ser Cys Phe Phe Gly	
	290	295 300
65	Thr Asp Asn Glu Tyr Trp Phe Gly Arg Leu His Tyr Ala Ala Phe Ser	
	305	310 315 320
	Cys Asn Trp His Thr Gln Asn Arg Ser Phe Lys Arg Lys Met Met Leu	
	325	330 335
	Phe Val Glu Gln Ser Leu Lys Lys Ser Thr Ala Val Ala Gly Gly Met	
	340	345 350

ES 2 315 227 T3

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

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

<210> 21

<211> 1293

<212> DNA

<213> *Drosophila melanogaster* DOR19g

<400> 21

```

atgggttacgg aggactttta taagtaccag gtgtggtact tccaaatcct tgggtgttgg 60
cagctcccca cttgggcgcg agaccaccag cgtcgttttc agtccatgag gtttggtctc 120
atcctgggtca tcctgttcat catgctgctg cttttctcct tcgaaatgtt gaacaacatt 180
tcccaagtta gggagatcct aaagggtattc ttcatgttcg ccacggaaat atcctgcatg 240
gccaaattat tgcatattgaa gttgaagagc cgcaaaactcg ctggcttggg tgatgcatg 300
ttgtccccag agttcggcgt taaaagttaa caggaaatgc agatgctgga attggataga 360
gtggcggttg tccgcatgag gaactcctac ggcacatcatg ccttgggcgc ggcttccctg 420
atccttatag ttccctgttt cgacaacttt ggcgagctac cactggccat gttggaggta 480
tgcagcatcg agggatggat ctgctattgg tcgcagtacc ttttccactc gatttgcctg 540
ctgcccactt gtgtgctgaa tataacctac gactcgggtg cctactcgtt gctctgtttc 600
ttgaagggtc agctacaaat gctggtcctg cgattagaaa agttgggtcc tgtgatcgaa 660
ccccaggata atgagaaaaa cgcaatggaa ctgctgtgag gtgccgccta ctacaacagg 720
attgttctgt tcaaggacct ggtggagctg ttcataaagg ggccaggatc tgtgcagetc 780
atgtgttctg ttctggtgct ggtgtccaac ctgtacgaca tgccaccat gtccattgca 840
aacggcgatg ccattcttat gctcaagacc tgtatctatc agctggtgat gctctggcag 900
atcttcatca tttgctacgc ctccaacgag gtaactgtcc agagctctag gttgtgtcac 960
agcatctaca gctcccaatg gacgggatgg aacagggcaa accgccggat tgtccttctc 1020
atgatgcagc gctttaattc cccgatgctc ctgagcacct ttaacccac ctttgctttc 1080
agcttggagg cctttggttc tgtagggcag cagaaattcc tttatatatc atttattact 1140
ggttatgctc ttctcctttc agatcgtcaa ctgctcctac agctacttcg cactgctgaa 1200
gcgcgtcaac agttaaattt cgaaacaccg cagcacctaa agattttcaa gccgattttt 1260
aaaagcactc aaaacgttat gcacgtacat taa 1293

```

<210> 22

<211> 430

<212> PRT

<213> Traducción DOR19g

<400> 22

Met Val Thr Glu Asp Phe Tyr Lys Tyr Gln Val Trp Tyr Phe Gln Ile
1 5 10 15

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

ES 2 315 227 T3

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

ES 2 315 227 T3

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

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

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

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

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

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

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

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

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

<210> 23

<211> 2075

<212> DNA

<213> *Drosophila melanogaster* DOR24g

<400> 23

atgcgcccaa cactccaggt cctcatcagt gtcctttgtc tggctctctgc ctacgcctgg 60
gatcataccg actgtaacga tcactatata gaattcatgg actatcccga tgagagagct 120
accgcctata gcaatgaatc ctccagaatgg gatttccttg aattttggag gcaagtgttt 180
ggcctgtagc aattctgaca tcgctgacca cgcctctggg caaagcaatc acttggcgag 240
acatgtgacc ctccgatgtc agggcggtgaa taaagaacat taaaatttaa ccggagttct 300
catttactat gcctaaagag gtgtcaaacg gaggtattat ttgggcgtga tatattatac 360
cctctaaagg agtatcccca ccagggtgaaa ccctataaaa cccttgacgt cgttaatgga 420
aagtactgaa ccggaaatga aagtgggcac tttctaattg tagaaatttg tcgggatgaa 480
atatataagt agatatatat aatattatat cttgattatc cagcatcaat cacaagataa 540
aaaaaaaaata taattcataa ttcataattc atatgtaata ggcatttgta aatgttgtaa 600
acaaaggaag ggtattaaag agctgtactt tcgcatctgt atttcattta ttcctactca 660
attgaaagggt ttaaaacccc tgaactgaac acacttgact tagtgtgagg ccgaattaac 720
ccttgctcgc atggacagtt ttctgcaagt acagaagagc accattgccc ttctgggctt 780
tgatctcttt agtgaaaatc gagaaatgtg gaaacgcccc tatagagcaa tgaatgtgtt 840

ES 2 315 227 T3

```

tagcatagct gccatgtttc cctttatcct ggcagctgtg ctccataatt ggaagaatgt 900
attgctgctg gccgatgccg tgggtggccct actaataacc attctgggccc tattcaagtt 960
tagcatgata ctttacttac gtcgcgattt caagcgactg attgacaaat ttctgtttgct 1020
5 catgtcgaat ggtgagttgt aatccatttc ggccagaatg tgtatcattt cattttattat 1080
tttatagagg cggaacaggg cgaggaatac gccgagattc tcaacgcagc aaacaagcag 1140
gatcaacgaa tgtgcactct gtttaggact tgtttcctcc tcgcctgggc cttgaatagt 1200
gttctgcccc tcgtgagaat ggggtctcagc tattgggttag cagggtcatgc agagccccgag 1260
10 ttgccttttc cctgtctgtg tgtacaaatg atatatatga tatatgggtga tcaagttatc 1320
aggcctttgtt cctaaagtgt tccctgggaat atccacatca ttcgcaatta tgttttgagc 1380
ttcatctgga gcgcttttcgc ctcgacaggt gtgggttttac ctgctgtcag cttggatacc 1440
atattctgtt ccttcaccag caacctgtgc gccttcttca aaattgcgca gtacaagggtg 1500
15 gttagattta agggcggatc ccttaaagaa tcacaggcca cattgaacaa agtctttgcc 1560
ctgtaccaga ccagcttggg tatgtgcaac gatctgaatc agtgctacca accgattatc 1620
tgcccccagt tcttcatttc atctctgcaa ctctgcatgc tgggatatct gttctccatt 1680
acttttgccc agacagaggg cgtctactat gcctcattca tagccacaat cattatacaa 1740
20 gcctatatct actgctactg cgggggagaac ctgaagacgg agagtgccag cttcgagtgg 1800
gccatctacg acagtccgtg gcacgagagt ttgggtgctg gtggagcctc tacctcgatc 1860
tgccgatcct tgctgatcag catgatgcgg gctcatcggg gattccgcat tacgggatac 1920
tttttcgagg caaacatgga ggccttctca tcgggtgggtg aatcatttcc attgtacaat 1980
25 acatggattt acatgaatac tcttttctaa ctttccggtt ttagattgtt cgcacggcga 2040
tgtcctacat cacaatgctg agatcattct cctaa 2075

```

```

30 <210> 24
    <211> 383
    <212> PRT
    <213> Drosophila melanogaster DOR24g
35 <400> 24

```

```

40 Met Arg Pro Thr Leu Gln Val Leu Ile Ser Val Leu Cys Leu Val Ser
    1 5 10 15

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

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

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

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

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

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

```

ES 2 315 227 T3

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