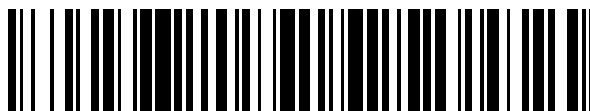


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 594 284**

51 Int. Cl.:

C12N 9/02 (2006.01)

C12N 15/82 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **10.11.2010** E 13164699 (4)

97 Fecha y número de publicación de la concesión europea: **06.07.2016** EP 2669372

54 Título: **Variantes de HPPD y procedimientos de uso**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
19.12.2016

73 Titular/es:

BAYER CROPSCIENCE AG (100.0%)
Alfred-Nobel-Strasse 50
40789 Monheim am Rhein, DE

72 Inventor/es:

LANGE, GUDRUN;
POREE, FABIEN;
LABER, BERND;
FREIGANG, JÖRG y
SCHULZ, ARNO

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 594 284 T3

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

Variantes de HPPD y procedimientos de uso

La presente invención se refiere a un ácido nucleico aislado que comprende una secuencia de nucleótidos que codifica una proteína de HPPD mutada, en la que dicha proteína de HPPD mutada tiene actividad de HPPD, en la que en dicha proteína de HPPD mutada un aminoácido ha sido reemplazado de modo tal que la secuencia de aminoácidos resultante comprende al menos un aminoácido seleccionado de ciertos aminoácidos en una posición específica importante por conferir una mayor tolerancia al inhibidor de HPPD. La presente invención también se refiere a proteínas codificadas por el ácido nucleico de la invención, a genes quiméricos, células vegetales que comprenden el ácido nucleico de la invención operativamente ligado a un promotor que se puede expresar en plantas y opcionalmente una región de terminación de la transcripción y poliadenilación, plantas que consisten esencialmente en las células vegetales de la invención y procedimientos para obtener plantas transgénicas.

En la presente memoria descriptiva, se cita una cantidad de documentos que incluyen solicitudes de patentes y manuales de fabricantes. La divulgación de estos documentos, aunque no se consideren relevantes para la patentabilidad de la presente invención, se incorpora aquí por referencia en su totalidad. Más específicamente, todos los documentos mencionados se incorporan por referencia como si cada documento individual se indicara específica e individualmente como incorporado por referencia.

Las proteínas de HPPD (hidroxifenilpiruvato dioxigenasa) son enzimas que catalizan la reacción en la que se transforma el para-hidroxifenilpiruvato (abreviado en el presente documento como HPP), un producto de degradación de la tirosina, en homogentisato (abreviado en el presente documento como HG), el precursor en plantas de tocoferol y plastoquinona (Crouch N.P. y col. (1997) Tetrahedron, 53, 20, 6993-7010, Fritze y col., (2004), Plant Physiology 134: 1388-1400). El tocoferol actúa como un antioxidante asociado a membrana. La plastoquinona actúa primeramente como un vehículo de electrones entre PSII y el complejo de citocromo b6/f y secundariamente, es un cofactor redox para la fitoeno desaturasa, que está implicada en la biosíntesis de los carotenoides.

Hasta ahora, más de 700 secuencias de ácidos nucleicos de diversos organismos presentes en la base de datos NCBI se anotaron como codificadores de una proteína teórica que tiene un dominio de HPPD. Pero para la mayoría de estas secuencias, no se probó que la proteína tuviera una actividad enzimática de HPPD ya sea en un ensayo in vitro o en un enfoque in planta, ni que tal proteína de HPPD pudiera conferir una tolerancia a herbicidas para los herbicidas inhibidores de HPPD cuando se expresa en una planta. Varias proteínas de HPPD y sus secuencias primarias fueron descritas en el estado de la técnica, en particular las HPPD de bacterias tales como *Pseudomonas* (Rüetschi y col., Eur. J. Biochem., 205, 459-466, 1992, documento WO 96/38567), de plantas tales como *Arabidopsis* (documento WO 96/38567, Genbank AF047834), zanahoria (documento WO 96/38567, Genbank 87257), *Avena sativa* (documento WO 02/046387), trigo (documento WO 02/046387), *Brachiaria platyphylla* (documento WO 02/046387), *Cenchrus echinatus* (documento WO 02/046387), *Lolium rigidum* (documento WO 02/046387), *Festuca arundinacea* (documento WO 02/046387), *Setaria faberi* (documento WO 02/046387), *Eleusine indica* (documento WO 02/046387), *Sorghum* (documento WO 02/046387), *Coccicoides* (Genbank COITRP), *Coptis japonica* (documento WO 06/132270), *Chlamydomonas reinhardtii* (documento ES 2275365), o de mamíferos tales como ratón o cerdo.

La mayoría de las plantas sintetizan tirosina por medio de arrogenato (Abou-Zeid y col. (1995), Applied Env Microb 41: 1298-1302; Bonner y col., (1995), Plant Cells Physiol. 36, 1013-1022; Byng y col., (1981), Phytochemistry 6: 1289-1292; Connely y Conn (1986), Z. Naturforsch 41c: 69-78; Gaines y col., (1982), Plants 156: 233-240). En estas plantas, el HPP se deriva solo de la degradación de tirosina. Por otro lado, en organismos tales como la levadura *Sacharomyces cerevisiae* o la bacteria *Escherichia coli*, el HPP es un precursor de tirosina y se sintetiza por la acción de una enzima, prefenato deshidrogenasa (de ahora en adelante referida como PDH), que convierte prefenato en HPP (Lingens y col., (1967) European J. Biochem 1: 363-374; Sampathkumar y Morrisson (1982), Bioch Biophys Acta 701: 204-211). En estos organismos en consecuencia la producción de HPP está conectada directamente con la vía de biosíntesis de aminoácidos aromáticos (ruta de shikimato) y no con la vía de degradación de la tirosina.

La inhibición de HPPD produce desacoplamiento de la fotosíntesis, deficiencia en pigmentos recolectores de luz accesorios y lo más importante, la destrucción de la clorofila por radiación UV y especies reactivas de oxígeno (blanqueo) debido a la falta de fotoprotección normalmente proporcionada por los carotenoides (Norris y col. (1995), Plant Cell 7: 2139-2149). El blanqueo de los tejidos fotosintéticamente activos lleva a la inhibición del crecimiento y la muerte de las plantas.

En la actualidad, los herbicidas inhibidores de HPPD más disponibles comercialmente pertenecen a una de estas cuatro familias químicas:

1) las tricetonas, por ejemplo sulcotriona [es decir 2-[2-cloro-4-(metilsulfonil)benzoil]-1,3-ciclohexandiona], mesotriona [es decir 2-[4-(metilsulfonil)-2-nitrobenzoil]-1,3-ciclohexandiona]; tembotriona [es decir 2-[2-cloro-4-(metilsulfonil)-3-[(2,2,2-trifluoroetoxi)metil]benzoil]-1,3-ciclohexandiona]; tefuriltriona [es decir 2-[2-cloro-4-(metilsulfonil)-3-[[tetrahydro-2-furanil]metoxi]metil]benzoil]-1,3-ciclohexandiona]; biciclopirona [es decir 4-hidroxi-3-

[[2-[(2-metoxietoxi)metil]-6-(trifluorometil)-3-piridinil]carbonil]biciclo[3.2.1]oct-3-en-2-ona]; benzobicyclona [es decir 3-(2-cloro-4-mesilbenzoil)-2-feniltiobiciclo[3.2.1]oct-2-en-4-ona]

2) los dicetonitrilos, por ejemplo 2-ciano-3-ciclopropil-1-(2-metilsulfonyl-4-trifluorometilfenil)-propan-1,3-diona y 2-ciano-1-[4-(metilsulfonyl)-2-trifluorometilfenil]-3-(1-metilciclopropil)propan-1,3-diona;

5 3) los isoxazoles, por ejemplo isoxaflutol [es decir (5-ciclopropil-4-isoxazolil)[2-(metilsulfonyl)-4-(trifluorometil)fenil]metanona]. En plantas, el isoxaflutol se metaboliza rápidamente en DKN, un compuesto de dicetonitrilo que exhibe la propiedad de inhibidor de HPPD; y

10 4) los pirazolinatos, por ejemplo topamezona [es decir [3-(4,5-dihidro-3-isoxazolil)-2-metil-4-(metilsulfonyl)fenil][5-hidroxi-1-metil-1H-pirazol-4-il]metanona] y pirasulfotol [(5-hidroxi-1,3-dimetilpirazol-4-il)(2-mesil-4-trifluorometilfenil)metanona]; pirazofeno [2-[4-(2,4-diclorobenzoil)-1,3-dimetilpirazol-5-iloxi]acetofenona].

15 Estos herbicidas inhibidores de HPPD se pueden usar contra pasto y/o malas hierbas de hoja ancha en plantas de cultivo que muestran tolerancia metabólica, como maíz (*Zea mays*) en las que se degradan rápidamente (Schulz y col., (1993). *FEBS letters*, 318, 162-166; Mitchell y col., (2001) *Pest Management Science*, vol. 57, 120-128; García y col., (2000) *Biochem.*, 39, 7501-7507; Pallett y col., (2001) *Pest Management Science*, vol. 57, 133-142). A fin de extender el alcance de estos herbicidas inhibidores de HPPD, se han desarrollado varios intentos con el fin de conferirles a las plantas, en particular plantas sin o con una tolerancia metabólica de bajo rendimiento, un nivel de tolerancia aceptable en condiciones de campo agronómicas.

20 Además del intento de desviar la protección mediada por HPPD del homogentisato (documento US 6.812.010), se realizó una sobreexpresión de la enzima sensible para producir cantidades de la enzima objetivo en la planta que son suficientes en relación con el herbicida (documento WO 96/38567). La sobreexpresión de HPPD dio como resultado una mejor tolerancia pre-germinación al derivado de dicetonitrilo (DKN) de isoxaflutol (IFT), pero la tolerancia no era suficiente para la tolerancia al tratamiento pre-germinación (Matringe y col., (2005), *Pest Management Science* 61: 269-276).

25 En el documento WO 04/024928, los inventores han tratado de aumentar la biosíntesis de la prenilquinona (por ejemplo, síntesis de plastoquinonas, tocoferoles) en las células vegetales aumentando el flujo del precursor de HPP en las células de estas plantas. Esto se ha realizado conectando la síntesis de dicho precursor con la ruta de "shikimato" por sobreexpresión de una enzima de PDH. También se ha observado que la transformación de plantas con un gen que codifica una enzima de PDH hace posible aumentar la tolerancia de dichas plantas a los inhibidores de HPPD.

30 Otra estrategia fue mutar la HPPD a fin de obtener una enzima objetivo que, mientras que conserva sus propiedades de catalizar la transformación de HPP en homogentisato, es menos sensible a los inhibidores de HPPD que la HPPD nativa antes de la mutación.

35 Esta estrategia se ha aplicado exitosamente para la producción de plantas tolerantes a 2-ciano-3-ciclopropil-1-(2-metilsulfonyl-4-trifluorometilfenil)-propan-1,3-diona y a 2-ciano-1-[4-(metilsulfonyl)-2-trifluorometilfenil]-3-(1-metilciclopropil)propan-1,3-diona (documento EP496630), dos herbicidas inhibidores de HPPD que pertenecen a la familia de los dicetonitrilos (documento WO 99/24585). Pro215Leu, Gly336Glu, Gly336Ile y más en particular Gly336Trp (las posiciones del aminoácido mutado se indican con referencia a la HPPD de *Pseudomonas*) se identificaron como mutaciones que son responsables de una mayor tolerancia al tratamiento de pre-germinación con estos herbicidas de dicetonitrilo sin causar una alteración de la actividad de la enzima.

40 Más recientemente, la introducción de un gen HPPD de *Pseudomonas* en el genoma de plástido de tabaco y soja se mostró más efectiva que la transformación nuclear, confiriendo incluso tolerancia a la aplicación post-germinación de isoxaflutol (Dufourmantel y col., 2007, *Plant Biotechnol J* 5(1):118-33).

45 En la solicitud de patente WO 2009/144079, se describe una secuencia de ácidos nucleicos que codifica una hidroxifenilpiruvato dioxigenasa mutada (HPPD) en la posición 336 de la proteína de HPPD de *Pseudomonas fluorescens* y su uso para obtener plantas que sean tolerantes a herbicidas del inhibidor de HPPD.

En el documento WO 2002/046387, se han identificado varios dominios de las proteínas de HPPD originados de plantas que pueden ser relevantes para conferir tolerancia a diversos herbicidas inhibidores de HPPD pero ni los datos in planta ni bioquímicos se han mostrando para confirmar el impacto de las funciones del dominio tal como se describen.

50 En el documento WO 2008/150473, la combinación de dos mecanismos distintos de tolerancia –un gen modificado de *Avena sativa* que codifica una enzima de HPPD mutante y una CYP450 monooxigenasa de maíz (gen nsf1)– se ejemplificó a fin de obtener una mayor tolerancia a herbicidas inhibidores de HPPD, pero no se divulgaron datos que demuestren los efectos sinérgicos a base de la combinación de ambas proteínas. El documento US 2010/0197503 sugiere una cantidad de mutaciones en diferentes posiciones dentro o cerca del sitio activo de la HPPD de *Avena sativa* y examinó alguna de ellas respecto de su inhibición por determinados inhibidores de HPPD como sulcotriona in vitro e in planta.

A pesar de estos éxitos obtenidos para el desarrollo de plantas que muestran tolerancia a varios herbicidas inhibidores de HPPD descritos con anterioridad, aún sigue siendo necesario desarrollar y/o mejorar la tolerancia de plantas a inhibidores de HPPD más nuevos o diferentes, en particular inhibidores de HPPD que pertenecen a las clases de las tricetonas (por ejemplo, sulcotriona, mesotriona, tembotriona, benzobiciclona y biclopirona) y los pirazolinatos (por ejemplo, topramezona y pirasulfotol).

Conforme a ello, la presente invención se refiere a un ácido nucleico aislado que comprende una secuencia de nucleótidos que codifica una proteína de HPPD mutada, en la que dicha proteína de HPPD mutada tiene actividad de HPPD, en la que en dicha proteína de HPPD mutada un aminoácido ha sido reemplazado de modo tal que la secuencia de aminoácidos resultante comprende al menos un aminoácido seleccionado de:

His, Ile, Leu, Asn o Ser en una posición en la proteína HPPD que corresponde a la posición 252 de la secuencia de aminoácidos de la SEC ID N.º: 2.

A menos que se indique otra cosa, las definiciones específicas o características específicas de determinadas formas de realización se pueden introducir en cualquier otra forma de realización de la presente invención. De acuerdo con la presente invención, un "ácido nucleico" se entiende como una secuencia de nucleótidos que puede ser del tipo ADN o ARN, con preferencia, del tipo ADN y en particular bicatenario, ya sea de origen natural o de origen sintético, en particular una secuencia de ADN en la que los codones que codifican la HPPD de acuerdo con la invención se han optimizado de acuerdo con el organismo huésped en el que se ha de expresar (por ejemplo, sustituyendo los codones con aquellos codones de mayor preferencia o de máxima preferencia en tablas de uso de codones de tal organismo huésped o del grupo al que dicho organismo huésped pertenece, en comparación con el organismo original o fuente).

Un "ácido nucleico/ADN/proteína aislado", tal como se usa en la presente solicitud, se refiere a un ácido nucleico/ADN/proteína que no se da en la naturaleza (tal como un ADN artificial o sintético con una secuencia de nucleótidos diferentes que el ADN que se da en la naturaleza, o una proteína modificada) o que ya no está en el ambiente natural en el que estaba originalmente presente, por ejemplo una secuencia que codifica el ADN asociada con un elemento de regulación heterólogo (tal como una secuencia codificadora bacteriana operativamente ligada a un promotor expresable en plantas) en un gen quimérico, un ADN transferido a otra célula huésped, tal como una célula de una planta transgénica.

La terminología relacionada con ácido nucleico o proteína "que comprende" determinada secuencia de nucleótidos o secuencia de aminoácidos, tal como se usa en todo el texto, se refiere a un ácido nucleico o proteína que incluye o contiene al menos la secuencia descrita, de modo que se pueda incluir otra secuencia de nucleótidos o aminoácidos en el extremo 5' (o N-terminal) y/o 3' (o C-terminal), por ejemplo (la secuencia de nucleótidos de) una proteína marcadora seleccionable, (la secuencia de nucleótidos de) un péptido de tránsito y/o una secuencia líder 5' o una secuencia remolque 3'. De modo similar, se debe entender que el uso del término "comprender", "que comprende" o "comprende" a lo largo de todo el texto y las reivindicaciones de esta solicitud implica la inclusión de un número entero o etapa o grupo de números enteros o etapas establecidos, pero no la exclusión de cualquier otro número entero o etapa o grupo de números enteros o etapas. La expresión "que comprende" también incluye la expresión "que consiste en".

De acuerdo con la presente invención, la expresión "proteína de HPPD mutada" que se usa indistintamente con la expresión "proteína de HPPD mutante" denota una proteína de HPPD que tiene una secuencia de aminoácidos que no se produce en la naturaleza. A diferencia del término "aislado" referido con anterioridad, el término "mutado" no se puede referir al ambiente de la secuencia (aminoacídica o proteica) en cuestión, como aislada de su ambiente natural o acoplada a una secuencia heteróloga (aminoacídica o proteica), sino que solo se refiere a la secuencia de aminoácidos que define dicha proteína de HPPD mutada que no se puede hallar en lugar alguno de la naturaleza pero que surgió a partir de una secuencia de aminoácidos no mutada o inicial de tipo silvestre. En otras palabras, al llegar al ácido nucleico de la presente invención que codifica una proteína de HPPD mutada, se debe tomar una secuencia de aminoácidos inicial de una proteína existente de forma natural y debe ser modificada por el hombre reemplazando al menos un aminoácido según se define en la presente solicitud.

La secuencia que codifica una HPPD no mutada original que será mutada de acuerdo con la invención puede tener cualquier origen. En particular, puede ser de origen bacteriano, vegetal o animal. Los ejemplos ventajosos que se pueden citar son bacterias de tipo *Pseudomonas* sp., por ejemplo, *Pseudomonas fluorescens*, o, de otro modo, cianobacterias del género *Synechocystis*. La secuencia también puede ser de origen vegetal, en particular derivada de plantas dicotiledóneas, plantas umbelíferas o, si no, plantas monocotiledóneas. Los ejemplos ventajosos que se pueden citar son plantas tales como tabaco, *Arabidopsis*, *Daucus carotta*, *Zea mays* (maíz), trigo, cebada, *Avena sativa*, trigo, *Brachiaria platyphylla*, *Cenchrus echinatus*, *Lolium rigidum*, *Festuca arundinacea*, *Setaria faberi*, *Eleusine indica* y *Sorghum*. Las secuencias codificantes y el modo de aislarlas y clonarlas, se describen en las referencias previamente citadas. En una forma de realización particular de la invención, la HPPD es de origen bacteriano, en particular de *Pseudomonas* sp., más en particular de *Pseudomonas fluorescens*, *Rhodococcus* sp., *Blepharisma japonicum*, *Synechococcus* sp., *Picrophilus torridus*, *Kordia algicida* o de un origen vegetal, en particular de *Arabidopsis thaliana* o *Avena sativa*. La HPPD para preparar la(s) mutación/mutaciones para los fines de la invención puede ser cualquier HPPD natural o cualquiera de sus fragmentos activos o cualquiera de sus

variantes en los que algunos aminoácidos (1 a 10 aminoácidos) se han reemplazado, añadido o suprimido para fines de clonación, para preparar una fusión de péptido de tránsito y similares, que conserva la actividad de HPPD, es decir la propiedad de catalizar la conversión de para-hidroxifenilpiruvato en homogentisato.

5 La proteína de HPPD mutada de acuerdo con la presente invención tiene actividad de HPPD, es decir, tal como se describió con anterioridad, cataliza la reacción en la que el para-hidroxifenilpiruvato se transforma en homogentisato. Con preferencia, la actividad catalítica de la HPPD mutada aislada de la presente invención, cuando se ensaya in vitro, no difiere de la de la proteína de HPPD de referencia no mutada en más del 70 %, con preferencia, más del 50 %, con mayor preferencia, más del 30 %, incluso con mayor preferencia, más del 20 % cuando se ensaya en idénticas condiciones y en ausencia de los herbicidas inhibidores de HPPD descritos con anterioridad. La actividad catalítica de una enzima de HPPD se puede definir por diversos procedimientos bien conocidos en la técnica. El documento WO 2009/144079 describe diversos procedimientos de cribado apropiados.

Las cribas iniciales se pueden llevar a cabo con el ácido nucleico que codifica la proteína de HPPD mutada de la invención que se expresa en las bacterias.

Ensayo de cribado colorimétrico para enzimas de HPPD activas:

15 Un medio de cultivo de tipo caldo de YT con agarosa al 1 %, L-tirosina 5 mM y succinato 42 mM, que contiene el agente de selección para el vector pSE420 (Invitrogen, Karlsruhe, Alemania), se vierte en placas de pocillos profundos. El cultivo de *E. coli* en la fase de crecimiento exponencial que contiene el vector pSE420-HPPDx (HPPDx implica cualquier gen que codifica una enzima/proteína de HPPD teórica) se aplica a cada pocillo. Después de 16 horas a 37 °C, los pocillos que no contienen el medio de cultivo, aquellos que se sembraron con un cultivo de *E. coli* con el vector pSE420 vacío son transparentes, aquellos que se sembraron con un cultivo de *E. coli* que contiene un vector pSE420-HPPDx que contiene un gen que codifica una HPPD inactiva son transparentes, si bien los pocillos sembradas con un cultivo de *E. coli* con el vector pSE420-HPPDx que codifica una HPPD activa son marrones. Previamente se demostró que este ensayo refleja la actividad de HPPD, no importa cuál sea el origen de esta actividad y permite la identificación de las actividades de HPPD (documento US 6.768.044), es decir, a un nivel cualitativo.

Se pueden llevar a cabo otros cribados y más elaborados en células vegetales o plantas que expresan la proteína de HPPD mutada de la invención.

30 Los mismos cribados también se pueden usar cuando se examina si una proteína de HPPD mutada es capaz de modular, tal como reducir o aumentar, la tolerancia de una planta a al menos un inhibidor de herbicida de HPPD que se mencionará luego más adelante, con la diferencia de que se añade al menos uno de tales inhibidores de HPPD. Los ejemplos de inhibidores de HPPD para usar en aquellos cribados incluyen tembotriona, mesotriona, pirasulfotol, biclopirona, topramezona y sulcotriona. Un procedimiento de cribado que es simple de implementar consiste en determinar la dosis de inhibidor de HPPD que inhibe completamente la HPPD no mutada original y que es letal para las células que expresan esta HPPD no mutada y someter las células mutadas a esta dosis predeterminada y después de ello, aislar las células mutadas que han resistido esta dosis letal y luego aislar y clonar el gen que codifica la HPPD mutada.

40 De modo alternativo, a nivel cuantitativo como pl_{50} (el valor pl_{50} implica el valor log de la concentración de inhibidor necesario para inhibir el 50 % de la actividad enzimática en concentración molar) se pueden obtener datos usando el polipéptido de HPPD aislado y purificado, es decir, el polipéptido de HPPD mutado frente al no mutado y en presencia o ausencia de cualquier herbicida inhibidor de HPPD respectivo.

45 Los términos “tolerancia”, “tolerante” o “menos sensible” denotan la falta de susceptibilidad de una planta que expresa la proteína de HPPD mutada de la presente invención a sustancias, en particular herbicidas, que inhiben las proteínas de HPPD, opcionalmente en comparación con la proteína de HPPD propia de la planta o con cualquier proteína conocida de HPPD. Más específicamente, dichos términos significan los niveles relativos de tolerancia inherente de la HPPD cribada de acuerdo con un fenotipo indicador visible de la cepa o la planta transformada con un ácido nucleico que comprende el gen que codifica la respectiva proteína de HPPD en presencia de diferentes concentraciones de los diversos inhibidores de HPPD. Las respuestas a las dosis y los desplazamientos relativos en las respuestas a las dosis asociadas con estos fenotipos indicadores (formación de color marrón, inhibición del crecimiento, blanqueo, efecto herbicida, etc.) se expresan de modo conveniente en términos, por ejemplo, de valores de GR50 (concentración del 50 % de reducción del crecimiento) o de MIC (concentración inhibitoria mínima) en los que los aumentos en valores corresponden a aumentos en la tolerancia inherente de la HPPD expresada, de la manera normal basada en el daño de plantas, síntomas de blanqueo meristemático, etc. en un rango de diferentes concentraciones de herbicidas. Estos datos se pueden expresar en términos de, por ejemplo, valores de GR50 derivados de curvas de dosis/respuesta que tienen la “dosis” representada en el eje x y “muerte en porcentaje”, “efecto herbicida”, “números de plantas verdes que germinan”, etc., representados en el eje y donde valores de GR50 aumentados corresponden a niveles aumentados de tolerancia inherente de la HPPD expresada. Los herbicidas se pueden aplicar apropiadamente pre-germinación o post-germinación.

Del mismo modo, el nivel de tolerancia del ácido nucleico o gen que codifica una proteína de HPPD de acuerdo con

la invención, o la proteína de HPPD mutada de la invención se somete a criba por medio de ensayos de transgénesis, regeneración, cruza y pulverización de una planta de ensayo tal como tabaco, o de una planta de cultivo tal como soja o algodón. En línea con los resultados obtenidos por tal criba, tales plantas son al menos 2-4 veces más tolerantes a los inhibidores de HPPD como tembotriona, mesotriona, dicetonitrilo y/o biciclopirona, pirasulfotol, que las plantas que no contienen ningún gen exógeno que codifique una proteína de HPPD, o que plantas que contienen un gen que comprende un ADN que codifica HPPD de *Arabidopsis thaliana*, sometido al control del mismo promotor que el ácido nucleico que codifica la proteína de HPPD mutada de la invención. Conforme a ello, la expresión "capaz de aumentar la tolerancia de una planta a al menos un herbicida que actúa sobre HPPD" denota una tolerancia incrementada en una planta en al menos el factor de 2, alternativamente al menos el factor de 3 o 4 o incluso 5 o 6 en comparación con una planta que solo expresa su HPPD endógena o una planta que expresa una HPPD de *Arabidopsis thaliana*. En este sentido, la expresión "herbicida que actúa sobre HPPD" no está limitada a sustancias que se conocen y/o que se usan como herbicidas sino a cualquier sustancia que inhibe la actividad catalítica de las proteínas de HPPD.

En una forma de realización alternativa del ácido nucleico que codifica un polipéptido de HPPD mutado que comprende al menos una de las mutaciones definidas con anterioridad, la proteína de HPPD comprende

a) una His en una posición en una proteína de HPPD, correspondiendo dicha posición a la posición 226 de la secuencia de aminoácidos de la SEC ID N.º 2;

b) una Ser en una posición en una proteína de HPPD, correspondiendo dicha posición a la posición 267 de la secuencia de aminoácidos de la SEC ID N.º 2;

c) una Asn en una posición en una proteína de HPPD, correspondiendo dicha posición a la posición 282 de la secuencia de aminoácidos de la SEC ID N.º 2;

d) una His en una posición en una proteína de HPPD, correspondiendo dicha posición a la posición 308 de la secuencia de aminoácidos de la SEC ID N.º 2;

e) una Tyr en una posición en una proteína de HPPD, correspondiendo dicha posición a la posición 342 de la secuencia de aminoácidos de la SEC ID N.º 2;

f) un Glu en una posición en una proteína de HPPD, correspondiendo dicha posición a la posición 394 de la secuencia de aminoácidos de la SEC ID N.º 2;

g) una Gly en una posición en una proteína de HPPD, correspondiendo dicha posición a la posición 420 de la secuencia de aminoácidos de la SEC ID N.º 2; y

h) una Asn en una posición en una proteína de HPPD, correspondiendo dicha posición a la posición 423 de la secuencia de aminoácidos de la SEC ID N.º 2.

En la proteína de HPPD mutada codificada por el ácido nucleico de la invención se ha modificado o reemplazado al menos un aminoácido según se ha definido con anterioridad.

La sustitución o la delección se pueden realizar en la secuencia de ácidos nucleicos que codifica la HPPD no mutada original, es decir que se da en la naturaleza, según se define con anterioridad por cualquier medio que sea apropiado para sustituir, en dicha secuencia, el codón que codifica el aminoácido para sustituirlo por el codón que corresponde al aminoácido que se va a sustituir o por delección de un codón, describiéndose dichos codones ampliamente en la bibliografía y siendo bien conocidos por el experto en la técnica.

Se pueden usar varios procedimientos biológicos moleculares para lograr esta sustitución o delección. Un procedimiento preferido para preparar una secuencia de ácidos nucleicos mutados de acuerdo con la invención y la correspondiente proteína comprende llevar a cabo mutagénesis dirigida a sitio sobre codones que codifican uno o más aminoácidos que se seleccionan de antemano. Los procedimientos para obtener estas mutaciones dirigidas a sitio se conocen bien por el experto en la técnica y se describen ampliamente en la bibliografía (en particular: Directed Mutagenesis: A Practical Approach, 1991, editado por M. J. McPHERSON, IRL PRESS), o son procedimientos para los que es posible emplear kits comerciales (por ejemplo, el kit de mutagénesis U. S. E. de PHARMACIA). Después de la mutagénesis dirigida a sitio, es de utilidad seleccionar las células que contienen una HPPD mutada que es menos sensible a un inhibidor de HPPD usando una ayuda de cribado apropiada. Los procedimientos de cribado apropiados para lograr esto se han descrito con anterioridad.

De acuerdo con la presente invención, la expresión "dicha posición corresponde a la posición X", siendo X cualquier número hallado en el respectivo contexto en la presente solicitud, no solo incluye la respectiva posición en la SEC ID N.º mencionada luego sino que también incluye cualquier secuencia que codifica una proteína de HPPD, en la que, después de la alineación con la SEC ID N.º de referencia, la respectiva posición puede tener un número diferente pero corresponde al indicado para la SEC ID N.º de referencia. Si bien las secuencias de HPPD pueden ser muy diversas y solo pueden mostrar una baja identidad de secuencia de aproximadamente el 30 %, las proteínas de HPPD se caracterizan por una estructura de consenso tridimensional común que se logra a pesar de una baja

identidad de secuencia. Debido a las posiciones específicas que se conservan dentro de las proteínas de HPPD, la alineación de las proteínas de HPPD se pueden efectuar aplicando diversas herramientas de alineación de una manera sensible.

5 Los procedimientos de alineación de secuencias de ácidos nucleicos o aminoácidos y conforme a ello, la determinación de la identidad de secuencia de dos o más secuencias, se conocen bien en la técnica. Incluyen realizar algoritmos matemáticos tales como el algoritmo de Myers y Miller (1988) CABIOS 4: 11-17 o el algoritmo de alineación local de Smith y Waterman (1981) Adv. Appl. Math. 2: 482-489; el algoritmo de alineación global de Needleman y Wunsch (1970) J. Mol. Biol. 48: 443-453; el algoritmo de Karlin y Altschul (1990) Proc. Natl. Acad. Sci USA 87:2264 y el de Brutlag y col. (Comp. App. Biosci. 6:237-245 (1990)).

10 Estos algoritmos se pueden implementar en programas de ordenador incluyendo pero sin limitación CLUSTALX, ALIGN, GAP, BESTFIT, BLAST, FASTDB y FASTA.

15 Por ejemplo, cuando se usa BESTFIT (Wisconsin Sequence Analysis Package, Versión 8 para Unix, Genetics Computer Group, University Research Park, 575 Science Drive, Madison, WI 53711) o cualquier otro programa de alineación de secuencias para determinar si una secuencia particular es, por ejemplo, un 95 % idéntica a una secuencia de referencia, los parámetros se fijan, por supuesto, de modo tal que el porcentaje de identidad se calcule en toda la longitud de la secuencia de nucleótidos de referencia y se permitan brechas en homología de hasta el 5 % de la cantidad total de nucleótidos en la secuencia de referencia.

20 La identidad entre una primera secuencia y una segunda secuencia, también mencionada como una alineación de secuencias global, se determina usando el programa de ordenador FASTDB basado en el algoritmo de Brutlag y col. (Comp. App. Biosci. 6:237-245 (1990)). En una alineación de secuencias la consulta y las secuencias objetivo son ambas secuencias de ADN. El resultado de dicha alineación de secuencias global es en porcentaje de identidad. Los parámetros preferidos usados en una alineación de FASTDB de secuencias de ADN para calcular el porcentaje de identidad son: Matriz = unitaria, k-tuple = 4, penalización por apareamiento incorrecto = 1, penalización por unión = 30, longitud del grupo de distribución aleatoria = 0, puntuación límite = 1, penalización por hueco = 5, penalización por tamaño de hueco 0,05, tamaño de ventana = 500 o la longitud de la secuencia de nucleótidos objetivo, que es más corta.

30 La presente invención se basa en los resultados de una combinación de una comparación de las secuencias de aminoácidos de proteínas de HPPD de diversos organismos y el análisis de la unión de sustrato y del sitio de unión del inhibidor de proteínas de HPPD seleccionadas, usando cristalografía por rayos X. Usando este enfoque combinado, era posible determinar las posiciones clave en las proteínas de HPPD, en las que un aminoácido se puede sustituir por uno de un grupo definido de otros aminoácidos a fin de modular la actividad catalítica de HPPD y la afinidad por al menos un inhibidor herbicida de HPPD con una planta que expresa la proteína de HPPD mutada.

35 La superposición de la estructura 3D de HPPD de *Arabidopsis thaliana* (1TFZ) (Yang y col., 2004, Biochemistry 43, 10414-10423) con las estructuras 3D de HPPD de otras especies tales como *Pseudomonas fluorescens* (1CJX) (Serre y col., 1999, Structure Fold Des. 7, 977-988), *Streptomyces avermitilis* (1T47) (Brownlee y col., 2004, Biochemistry 43, 6370-6377), *Homo sapiens* (3ISQ) (PDB ID: 3isq Pilka y col., Structural Genomics Consortium (SGC). Crystal structure of human 4-Hydroxyfenilpiruvato dioxygenase), *Rattus norvegicus* (1SQI) (Yang y col., 2004, Biochemistry 43, 10414-10423) muestra que tienen el mismo plegado y los aminoácidos correspondientes están en una posición equivalente en la estructura 3D de la proteína. Como las especies con estructuras 3D conocidas son muy diversas en su secuencia de aminoácidos, se puede asumir que todas las secuencias de HPPD tienen el mismo plegado básico a pesar de que la identidad de secuencia general es baja. La secuencia y la estructura 3D de *Arabidopsis thaliana* se ha usado como estructura de referencia en la presente invención. La Figura 1 muestra la superposición de la estructura de HPPD de *A. thaliana* con la estructura de (a) *Pseudomonas fluorescens*, (b) *Streptomyces avermitilis*, (c) *Homo sapiens* y (d) *Rattus norvegicus*. A fin de definir el sitio de unión del sustrato y/o inhibidores, se seleccionaron los aminoácidos que desempeñan un papel en la catálisis o unión del inhibidor. Esto incluye aminoácidos en el sitio activo y aminoácidos de la hélice C-terminal. La disposición 3D se muestra en la Figura 2 que muestra los aminoácidos definidos como sitio de unión en el caso de (a) *Arabidopsis thaliana*, (b) *Pseudomonas fluorescens*, (c) *Streptomyces avermitilis*, (d) *Homo sapiens* y (e) *Rattus norvegicus*. La numeración de los aminoácidos de la estructura de *Pseudomonas fluorescens* (1cjsx) se cambió a la numeración de acuerdo con la SEC ID N.º: 10. Los 36 aminoácidos que definen el sitio de unión incluyendo su posición se enumeran en la Tabla 1 para (a) *Arabidopsis thaliana*, (b) *Pseudomonas fluorescens*, (c) *Streptomyces avermitilis*, (d) *Homo sapiens* y (e) *Rattus norvegicus*.

Tabla 1: aminoácidos que forman el sitio de unión en *A. thaliana*, *P. fluorescens*, *S. avermitilis*, *H. sapiens*, *R. norvegicus* [no formando parte de la invención]

<i>Arabidopsis thaliana</i>		<i>Pseudomonas fluorescens</i>		<i>Streptomyces avermitilis</i>		<i>Homo sapiens</i>		<i>Rattus norvegicus</i>	
Posición	Amino-ácido	Posición	Amino-ácido	Posición	Amino-ácido	Posición	Amino-ácido	Posición	Amino-ácido
226	H	162	H	187	H	183	H	183	H
228	V	164	T	189	V	185	V	185	V
248	H	186	R	211	T	207	H	207	H
250	F	188	A	213	M	209	F	209	F
251	A	189	R	214	K	210	W	210	W
252	E	190	Y	215	E	211	S	211	S
253	F	191	F	216	F	212	V	212	V
265	L	200	L	228	L	224	L	224	L
267	S	202	S	230	S	226	S	226	S
268	A	203	K	231	K	227	I	227	I
269	V	204	A	232	V	228	V	228	V
270	L	205	M	233	V	229	V	229	V
271	A	206	S	234	A	230	A	230	A
280	P	215	P	243	P	239	P	239	P
282	N	217	N	245	N	241	N	241	N
293	Q	226	Q	255	Q	251	Q	251	Q
294	I	227	I	256	I	252	I	252	I
307	Q	240	Q	269	Q	265	Q	265	Q
308	H	241	H	270	H	266	H	266	H
335	M	264	M	293	L	289	L	289	L
342	Y	271	Y	299	Y	295	Y	295	Y
368	L	295	L	323	L	323	L	323	L
379	Q	310	Q	334	Q	334	Q	334	Q
381	F	312	F	336	F	336	F	336	F
392	F	321	F	347	F	347	F	347	F
394	E	323	E	349	E	349	E	349	E
419	F	333	F	359	F	359	F	359	F
420	G	334	G	360	G	360	G	360	G
421	K	335	E	361	K	361	A	361	A
422	G	336	G	362	G	362	G	362	G
423	N	337	N	363	N	363	N	363	N
424	F	338	F	364	F	364	F	364	F
425	S	339	K	365	K	365	N	365	N
426	E	340	A	366	A	366	S	366	S
427	L	341	L	367	L	367	L	367	L
431	I	345	I	371	I	371	F	371	F

5 Una alineación de ejemplo de las proteínas de HPPD se da en la Tabla 2a para las proteínas de HPPD con estructuras 3D conocidas. La Tabla 2a da la numeración de los aminoácidos de la secuencia de *Arabidopsis* y también los aminoácidos que son comunes dentro de estas secuencias de HPPD, designando estos aminoácidos con un asterisco. Sobre la base de tal alineación y a partir de la definición del aminoácido de *Arabidopsis* por su posición y su naturaleza, es fácil identificar la posición del correspondiente aminoácido en otra secuencia de HPPD.

La Figura 2 muestra que esto se puede llevar a cabo con la alineación de las secuencias de diferente origen vegetal, mamífero y bacteriano, lo que demuestra que este procedimiento de alineación, que es bien conocido por el experto en la técnica, puede ser generalizado a cualquier otra secuencia. Una alineación de diferentes secuencias de HPPD también se describe en la solicitud de patente WO 97/49816.

5 Tabla 2a: alineación de secuencias de HPPD con estructuras cristalinas conocidas, es decir, *A. thaliana*, *P. fluorescens*, *S. avermitilis*, *H. sapiens*, *R. norvegicus* [no formando parte de la invención]

Pos.	1TFZ <i>A. thaliana</i>	Pos.	1CJX <i>P. fluorescens</i>	Pos.	1T47 <i>S. avermitilis</i>	Pos.	3ISQ <i>H. sapiens</i>	Pos.	1SQI <i>R. norvegicus</i>	Común aminoácido
1	M		-		-		-			
2	G		-		-		-			
3	H		-		-		-			
4	Q		-		-		-			
5	N		-		-		-			
6	A		-		-		-			
7	A		-		-		-			
8	V		-		-		-			
9	S		-		-		-			
10	E		-		-		-			
11	N		-		-		-			
12	Q		-		-		-			
13	N		-		-		-			
14	H		-	1	M		-			
15	D		-	2	T		-			
16	D		-	3	Q		-			
17	G		-	4	T		-			
18	A		-	5	T		-			
19	A		-	6	H	8	G	8	G	
20	S		-	7	H	9	A	9	P	
21	S		-	8	T	10	K	10	K	
22	P		-	9	P	11	P	11	P	
23	G		-		-		-			
24	F		-		-		-			
25	K		-		-		-			
26	L		-		-		-			
27	V		-		-		-			
28	G		-		-		-			
29	F		-		-		-			
30	S		-		-		-			
31	K		-		-		-			
32	F		-		-		-			
33	V		-		-		-			
34	R		-	10	D	12	E	12	E	
35	K	2	A	11	T	13	R	13	R	

ES 2 594 284 T3

Pos.	1TFZ	Pos.	1CJX	(continuación)				Pos.	1SQI	Común aminoácido
	A. thaliana		P. fluorescens	Pos.	1T47	Pos.	3ISQ		Pos.	
				S. avermitilis		H. sapiens				
36	N	3	D	12	A	14	G	14	G	
37	P	4	L	13	R	15	R	15	R	
38	K	5	Y	14	Q	16	F	16	F	
39	S	6	E	15	A	17	L	17	L	
40	D	7	N	16	D	18	H	18	H	
41	K	8	P	17	P		-			
42	F	9	M	18	F		-			
43	K	10	G	19	P		-			
44	V	11	L	20	V		-			
45	K	12	M	21	K		-			
46	R	13	G	22	G		-			
47	F	14	F	23	M	19	F	19	F	
48	H	15	E	24	D	20	H	20	H	
49	H	16	F	25	A	21	S	21	S	
50	I	17	I	26	V	22	V	22	V	
51	E	18	E	27	V	23	T	23	T	
52	F	19	F	28	F	24	F	24	F	x
53	W	20	A	29	A	25	W	25	W	
54	C	21	S	30	V	26	V	26	V	
55	G	22	P	31	G	27	G	27	G	
56	D	23	T	32	N	28	N	28	N	
57	A	24	P	33	A	29	A	29	A	
58	T	25	G	34	K	30	K	30	K	
59	N	26	T	35	Q	31	Q	31	Q	
60	V	27	L	36	A	32	A	32	A	
61	A	28	E	37	A	33	A	33	A	
62	R	29	P	38	H	34	S	34	S	
63	R	30	I	39	Y	35	F	35	F	
64	F	31	F	40	Y	36	Y	36	Y	
65	S	32	E	41	S	37	C	37	C	
66	W	33	I	42	T	38	S	38	N	
67	G	34	M	43	A	39	K	39	K	
68	L	35	G	44	F	40	M	40	M	
69	G	36	F	45	G	41	G	41	G	
70	M	37	T	46	M	42	F	42	F	
71	R	38	K	47	Q	43	E	43	E	
72	F	39	V	48	L	44	P	44	P	
73	S	40	A	49	V	45	L	45	L	
74	A	41	T	50	A	46	A	46	A	
75	K	42	H	51	Y	47	Y	47	Y	
76	S	43	R	52	S	48	R	48	K	

ES 2 594 284 T3

(continuación)

Pos.	1TFZ	Pos.	1CJX	Pos.	1T47	Pos.	3ISQ	Pos.	1SQI	Común
	A. thaliana		P. fluorescens		S. avermitilis		H. sapiens		R. norvegicus	aminoácido
77	D	44	S	53	G	49	G	49	G	
78	L	45	K	54	P	50	L	50	L	
79	S	46	N	55	E	51	E	51	E	
80	T		-	56	N	52	T	52	T	
81	G		-	57	G	53	G	53	G	
82	N		-	58	S	54	S	54	S	
83	M		-	59	R	55	R	55	R	
84	V		-	60	E	56	E	56	E	
85	H		-	61	T	57	V	57	V	
86	A		-	62	A	58	V	58	V	
87	S	47	V	63	S	59	S	59	S	
88	Y	48	H	64	Y	60	H	60	H	
89	L	49	L	65	V	61	V	61	V	
90	L	50	Y	66	L	62	I	62	I	
91	T	51	R	67	T	63	K	63	K	
92	S	52	Q	68	N	64	Q	64	Q	
93	G	53	G	69	G	65	G	65	G	X
94	D	54	E	70	S	66	K	66	K	
95	L	55	I	71	A	67	I	67	I	
96	R	56	N	72	R	68	V	68	V	
97	F	57	L	73	F	69	F	69	F	
98	L	58	I	74	V	70	V	70	V	
99	F	59	L	75	L	71	L	71	L	
100	T	60	N	76	T	72	S	72	C	
101	A	61	N	77	S	73	S	73	S	
102	P	62	E	78	V	74	A	74	A	
103	Y	63	P	79	I	75	L	75	L	
104	S	64	N	80	K	76	N	76	N	
105	P	65	S	81	P		-			
106	S	66	I	82	A		-			
107	L	67	A	83	T		-			
108	S	68	S	84	P	77	P	77	P	
109	A		-	85	W	78	W	78	W	
110	G		-	86	G	79	N	79	N	
111	E		-	87	H	80	K	80	K	
112	I		-	88	F	81	E	81	E	
113	K		-	89	L	82	M	82	M	
114	P		-	90	A	83	G	83	G	
115	T		-		-		-			
116	T		-		-		-			
117	T		-		-		-			

ES 2 594 284 T3

(continuación)

Pos.	1TFZ	Pos.	1CJX	Pos.	1T47	Pos.	3ISQ	Pos.	1SQI	Común
	A. thaliana	P. fluorescens	S. avermitilis	H. sapiens	R. norvegicus	aminoácido				
118	A	-	-	-	-	-	-	-	-	
119	S	-	-	-	-	-	-	-	-	
120	I	-	-	-	-	-	-	-	-	
121	P	-	-	-	-	-	-	-	-	
122	S	-	-	-	-	-	-	-	-	
123	F	-	-	-	-	-	-	-	-	
124	D	-	-	-	-	-	-	-	-	
125	H	-	-	-	-	-	-	-	-	
126	G	-	-	-	-	-	-	-	-	
127	S	-	-	-	-	-	-	-	-	
128	C	-	-	-	-	-	-	-	-	
129	R	-	-	-	-	-	-	-	-	
130	S	69	Y	91	D	84	D	84	D	
131	F	70	F	92	H	85	H	85	H	
132	F	71	A	93	V	86	L	86	L	
133	S	72	A	94	A	87	V	87	V	
134	S	73	E	95	E	88	K	88	K	
135	H	74	H	96	H	89	H	89	H	X
136	G	75	G	97	G	90	G	90	G	X
137	L	76	P	98	D	91	D	91	D	
138	G	77	S	99	G	92	G	92	G	
139	V	78	V	100	V	93	V	93	V	X
140	R	79	C	101	V	94	K	94	K	
141	A	80	G	102	D	95	D	95	D	
142	V	81	M	103	L	96	I	96	I	
143	A	82	A	104	A	97	A	97	A	X
144	I	83	F	105	I	98	F	98	F	
145	E	84	R	106	E	99	E	99	E	
146	V	85	V	107	V	100	V	100	V	X
147	E	86	K	108	P	101	E	101	E	
148	D	87	D	109	D	102	D	102	D	X
149	A	88	S	110	A	103	C	103	C	
150	E	89	Q	111	R	104	D	104	E	
151	S	90	K	112	A	105	Y	105	H	
152	A	91	A	113	A	106	I	106	I	
153	F	92	Y	114	H	107	V	107	V	
154	S	93	N	115	A	108	Q	108	Q	
155	I	94	R	116	Y	109	K	109	K	
156	S	95	A	117	A	110	A	110	A	
157	V	96	L	118	I	111	R	111	R	
158	A	97	E	119	E	112	E	112	E	

ES 2 594 284 T3

(continuación)

Pos.	1TFZ	Pos.	1CJX	Pos.	1T47	Pos.	3ISQ	Pos.	1SQI	Común
	A. thaliana	P. fluorescens		S. avermitilis		H. sapiens		R. norvegicus		aminoácido
159	N	98	L	120	H	113	R	113	R	
160	G	99	G	121	G	114	G	114	G	X
161	A	100	A	122	A	115	A	115	A	
162	I	101	Q	123	R	116	K	116	K	
163	P		-	124	S	117	I	117	I	
164	S		-	125	V	118	M	118	V	
165	S		-	126	A	119	R	119	R	
166	P		-	127	E	120	E	120	E	
167	P	102	P	128	P	121	P	121	P	X
168	I	103	I	129	Y	122	W	122	W	
169	V	104	H	130	E	123	V	123	V	
170	L	105	I	131	L	124	E	124	E	
171	N	106	D	132	K	125	Q	125	E	
172	E	107	T	133	D	126	D	126	D	
173	A	108	G	134	E	127	K	127	K	
174	V	109	P	135	H	128	F	128	F	
175	T	110	M	136	G	129	G	129	G	
176	I	111	E	137	T	130	K	130	K	
177	A	112	L	138	V	131	V	131	V	
178	E	113	N	139	V	132	K	132	K	
179	V	114	L	140	L	133	F	133	F	
180	K	115	P	141	A	134	A	134	A	
181	L	116	A	142	A	135	V	135	V	
182	Y	117	I	143	I	136	L	136	L	
183	G	118	K	144	A	137	Q	137	Q	
184	D	119	G	145	T	138	T	138	T	
185	V	120	I	146	Y	139	Y	139	Y	
186	V	121	G	147	G	140	G	140	G	
187	L	122	G	148	K	141	D	141	D	
188	R	123	A	149	T	142	T	142	T	
189	Y	124	P	150	R	143	T	143	T	
190	V	125	L	151	H	144	H	144	H	
191	S	126	Y	152	T	145	T	145	T	
192	Y	127	L	153	L	146	L	146	L	
193	K	128	I	154	V	147	V	147	V	
194	A	129	D	155	D	148	E	148	E	
195	E	130	R	156	R	149	K	149	K	
196	D	131	F	157	T	150	M	150	I	
197	T	132	G	158	G	151	N	151	N	
198	E	133	E	159	Y	152	Y	152	Y	
	-	134	G		-		-			

ES 2 594 284 T3

(continuación)

Pos.	1TFZ	Pos.	1CJX	Pos.	1T47	Pos.	3ISQ	Pos.	1SQI	Común
	A. thaliana	P. fluorescens		S. avermitilis		H. sapiens		R. norvegicus		aminoácido
199	K	135	S	160	D	153	I	153	T	
200	S	136	S	161	G	154	G	154	G	
201	E	137	I	162	P	155	Q	155	R	
202	F	138	Y	163	Y	156	F	156	F	
203	L	139	D	164	L	157	L	157	L	
204	P	140	I	165	P	158	P	158	P	
205	G	141	D	166	G	159	G	159	G	
206	F	142	F	167	Y	160	Y	160	F	
207	E	143	V	168	V	161	E	161	E	
208	R	144	Y	169	A	162	A	162	A	
209	V	145	L	170	A	163	P	163	P	
210	E	146	E	171	A	164	A	164	T	
211	D	147	G		-	165	F	165	Y	
	-	-	-		-	166	M	166	K	
	-	-	-		-	167	D	167	D	
	-	-	-	172	P	168	P	168	T	
212	A	148	V	173	I	169	L	169	L	
213	S	149	E	174	V	170	L	170	L	
214	S	150	R	175	E	171	P	171	P	
215	F	151	N	176	P	172	K	172	K	
216	P	152	P	177	P	173	L	173	L	
217	L	153	V	178	A	174	P	174	P	
218	D	154	G	179	H	175	K	175	S	
219	Y	155	A	180	R	176	C	176	C	
220	G	156	G	181	T	177	S	177	N	
221	I	157	L	182	F	178	L	178	L	
222	R	158	K	183	Q	179	E	179	E	
223	R	159	V	184	A	180	M	180	I	
224	L	160	I	185	I	181	I	181	I	
225	D	161	D	186	D	182	D	182	D	x
226	H	162	H	187	H	183	H	183	H	x
227	A	163	L	188	C	184	I	184	I	
228	V	164	T	189	V	185	V	185	V	
229	G	165	H	190	G	186	G	186	G	
230	N	166	N	191	N	187	N	187	N	x
231	V	167	V	192	V	188	Q	188	Q	
232	P	168	Y	193	E	189	P	189	P	
	-	169	R	194	L	190	D	190	D	
	-	170	G	195	G	191	Q	191	Q	
233	E	171	R	196	R	192	E	192	E	
234	L	172	M	197	M	193	M	193	M	

ES 2 594 284 T3

(continuación)

Pos.	1TFZ	Pos.	1CJX	Pos.	1T47	Pos.	3ISQ	Pos.	1SQI	Común
	A. thaliana	P. fluorescens		S. avermitilis		H. sapiens		R. norvegicus		aminoácido
235	G	173	V	198	N	194	V	194	E	
236	P	174	Y	199	E	195	S	195	S	
237	A	175	W	200	W	196	A	196	A	
238	L	176	A	201	V	197	S	197	S	
239	T	177	N	202	G	198	E	198	E	
240	Y	178	F	203	F	199	W	199	W	
241	V	179	Y	204	Y	200	Y	200	Y	
242	A	180	E	205	N	201	L	201	L	
243	G	181	K	206	K	202	K	202	K	
244	F	182	L	207	V	203	N	203	N	
245	T	183	F	208	M	204	L	204	L	
246	G	184	N	209	G	205	Q	205	Q	
247	F	185	F	210	F	206	F	206	F	x
248	H	186	R	211	T	207	H	207	H	
249	Q	187	E	212	N	208	R	208	R	
250	F	188	A	213	M	209	F	209	F	
251	A	189	R	214	K	210	W	210	W	
252	E	190	Y	215	E	211	S	211	S	
253	F	191	F	216	F	212	V	212	V	
254	T	192	D	217	V	213	D	213	D	
255	A	193	I	218	G	214	D	214	D	
256	D	194	K	219	D	215	T	215	T	
257	D	195	G	220	D	216	Q	216	Q	
258	V	196	E	221	I	217	V	217	V	
259	G		-	222	A	218	H	218	H	
260	T		-	223	T	219	T	219	T	
261	A		-	224	E	220	E	220	E	
262	E	197	Y	225	Y	221	Y	221	Y	
263	S	198	T	226	S	222	S	222	S	
264	G	199	G	227	A	223	S	223	S	
265	L	200	L	228	L	224	L	224	L	x
266	N	201	T	229	M	225	R	225	R	
267	S	202	S	230	S	226	S	226	S	x
268	A	203	K	231	K	227	I	227	I	
269	V	204	A	232	V	228	V	228	V	
270	L	205	M	233	V	229	V	229	V	
271	A	206	S	234	A	230	A	230	A	x
272	S	207	A	235	D	231	N	231	N	
273	N	208	P	236	G	232	Y	232	Y	
274	D	209	D	237	T	233	E	233	E	
275	E	210	G	238	L	234	E	234	E	

ES 2 594 284 T3

(continuación)

Pos.	1TFZ	Pos.	1CJX	Pos.	1T47	Pos.	3ISQ	Pos.	1SQI	Común
	A. thaliana	P. fluorescens		S. avermitilis		H. sapiens		R. norvegicus		aminoácido
276	M	211	M	239	K	235	S	235	S	
277	V	212	I	240	V	236	I	236	I	
278	L	213	R	241	K	237	K	237	K	
279	L	214	I	242	F	238	M	238	M	
280	P	215	P	243	P	239	P	239	P	x
281	I	216	L	244	I	240	I	240	I	
282	N	217	N	245	N	241	N	241	N	x
283	E	218	E	246	E	242	E	242	E	x
284	P	219	E	247	P	243	P	243	P	
285	V	220	S	248	A	244	A	244	A	
286	H		-	249	L	245	P	245	P	
287	G		-	250	A	246	G	246	G	
288	T	221	S	251	K	247	K	247	R	
289	K	222	K	252	K	248	K	248	K	x
290	R	223	G		-		-			
291	K	224	A	253	K	249	K	249	K	
292	S	225	G	254	S	250	S	250	S	
293	Q	226	Q	255	Q	251	Q	251	Q	x
294	I	227	I	256	I	252	I	252	I	x
295	Q	228	E	257	D	253	Q	253	Q	
296	T	229	E	258	E	254	E	254	E	
297	Y	230	F	259	Y	255	Y	255	Y	
298	L	231	L	260	L	256	V	256	V	
299	E	232	M	261	E	257	D	257	D	
300	H	233	Q	262	F	258	Y	258	Y	
301	N	234	F	263	Y	259	N	259	N	
302	E	235	N	264	G	260	G	260	G	
303	G	236	G	265	G	261	G	261	G	x
304	A	237	E	266	A	262	A	262	A	
305	G	238	G	267	G	263	G	263	G	x
306	L	239	I	268	V	264	V	264	V	
307	Q	240	Q	269	Q	265	Q	265	Q	x
308	H	241	H	270	H	266	H	266	H	x
309	L	242	V	271	I	267	I	267	I	
310	A	243	A	272	A	268	A	268	A	x
311	L	244	F	273	L	269	L	269	L	
312	M	245	L	274	N	270	K	270	R	
313	S	246	T	275	T	271	T	271	T	
314	E	247	D	276	G	272	E	272	E	
315	D	248	D	277	D	273	D	273	D	x
316	I	249	L	278	I	274	I	274	I	

ES 2 594 284 T3

(continuación)

Pos.	1TFZ	Pos.	1CJX	Pos.	1T47	Pos.	3ISQ	Pos.	1SQI	Común
	A. thaliana	P. fluorescens		S. avermitilis		H. sapiens		R. norvegicus		aminoácido
317	F	250	V	279	V	275	I	275	I	
318	R	251	K	280	E	276	T	276	T	
319	T	252	T	281	T	277	A	277	T	
320	L	253	W	282	V	278	I	278	I	
321	R	254	D	283	R	279	R	279	R	
322	E	255	A	284	T	280	H	280	H	
323	M	256	L	285	M	281	L	281	L	
324	R	257	K	286	R	282	R	282	R	
325	K	258	K	287	A	283	E	283	E	
326	R	259	I	288	A	284	R	284	R	
327	S		-		-		-		-	
328	S		-		-		-		-	
329	I		-		-		-		-	
330	G		-		-		-		-	
331	G	260	G	289	G	285	G	285	G	x
332	F	261	M	290	V	286	L	286	M	
333	D	262	R	291	Q	287	E	287	E	
334	F	263	F	292	F	288	F	288	F	x
335	M	264	M	293	L	289	L	289	L	
336	P	265	T	294	D	290	S	290	A	
337	S	266	A	295	T	291	V	291	V	
338	P	267	P	296	P	292	P	292	P	x
339	P	268	P		-		-		-	
340	P	269	D	297	D	293	S	293	S	
341	T	270	T	298	S	294	T	294	S	
342	Y	271	Y	299	Y	295	Y	295	Y	x
343	Y	272	Y	300	Y	296	Y	296	Y	x
344	Q	273	E	301	D	297	K	297	R	
345	N	274	M	302	T	298	Q	298	L	
346	L	275	L	303	L	299	L	299	L	x
347	K	276	E	304	G	300	R	300	R	
348	K	277	G	305	E	301	E	301	E	
349	R	278	R	306	W	302	K	302	N	
350	V	279	L	307	V	303	L	303	L	
351	G	280	P	308	G	304	K	304	K	
352	D	281	D	309	D	305	T	305	T	
353	V	282	H	310	T	306	A	306	S	
354	L	283	G	311	R	307	K	307	K	
355	S	284	E	312	V	308	I	308	I	
356	D	285	P	313	P	309	K	309	Q	
357	D	286	V	314	V	310	V	310	V	

ES 2 594 284 T3

(continuación)

Pos.	1TFZ	Pos.	1CJX	Pos.	1T47	Pos.	3ISQ	Pos.	1SQI	Común
	A. thaliana		P. fluorescens		S. avermitilis		H. sapiens		R. norvegicus	aminoácido
358	Q	287	D	-	-	311	K	311	K	
	-	288	Q	-	-	312	E	312	E	
	-	289	L	-	-	313	N	313	N	
359	I	290	Q	-	-	314	I	314	M	
360	K	291	A	315	D	315	D	315	D	
361	E	292	R	316	T	316	A	316	V	
362	C	293	G	317	L	317	L	317	L	
363	E	294	I	318	R	318	E	318	E	
364	E	295	L	319	E	319	E	319	E	
365	L	296	L	320	L	320	L	320	L	x
366	G	297	D	321	K	321	K	321	K	
367	I	298	G	322	I	322	I	322	I	
368	L	299	S	323	L	323	L	323	L	
369	V	300	S	324	A	324	V	324	V	
370	D	301	V	325	D	325	D	325	D	
371	R	302	E	326	R	326	Y	326	Y	
372	D	303	G	327	D	327	D	327	D	
373	D	304	D	328	E	328	E	328	E	
374	Q	305	K	329	D	329	K	329	K	
375	G	306	R	330	G	330	G	330	G	
376	T	307	L	331	Y	331	Y	331	Y	
377	L	308	L	332	L	332	L	332	L	x
378	L	309	L	333	L	333	L	333	L	x
379	Q	310	Q	334	Q	334	Q	334	Q	x
380	I	311	I	335	I	335	I	335	I	x
381	F	312	F	336	F	336	F	336	F	x
382	T	313	S	337	T	337	T	337	T	
383	K	314	E	338	K	338	K	338	K	
384	P	315	T	339	P	339	P	339	P	
385	L	316	L	340	V	340	V	340	M	
386	G	317	M	341	Q	341	Q	341	Q	
387	D	318	G	342	D	342	D	342	D	
388	R	-	-	343	R	343	R	343	R	
389	P	-	-	344	P	344	P	344	P	
390	T	319	P	345	T	345	T	345	T	
391	I	320	V	346	V	346	L	346	L	
392	F	321	F	347	F	347	F	347	F	x
393	I	322	F	348	F	348	L	348	L	
394	E	323	E	349	E	349	E	349	E	x
395	I	324	F	350	I	350	V	350	V	
396	I	325	I	351	I	351	I	351	I	x

ES 2 594 284 T3

(continuación)

Pos.	1TFZ	Pos.	1CJX	Pos.	1T47	Pos.	3ISQ	Pos.	1SQI	Común
	A. thaliana		P. fluorescens		S. avermitilis		H. sapiens		R. norvegicus	aminoácido
397	Q	326	Q	352	E	352	Q	352	Q	
398	R	327	R	353	R	353	R	353	R	x
399	V	328	K	354	H	354	H	354	H	
400	G		-		-		-		-	
401	C		-		-		-		-	
402	M		-		-		-		-	
403	M		-		-		-		-	
404	K		-		-		-		-	
405	D		-		-		-		-	
406	E		-		-		-		-	
407	E		-		-		-		-	
408	G		-		-		-		-	
409	K		-		-		-		-	
410	A		-		-		-		-	
411	Y		-		-		-		-	
412	Q		-		-		-		-	
413	S		-		-		-		-	
414	G		-		-		-		-	
415	G	329	G	355	G	355	N	355	N	
416	C	330	D	356	S	356	H	356	H	
417	G	331	D	357	M	357	Q	357	Q	
418	G	332	G	358	G	358	G	358	G	x
419	F	333	F	359	F	359	F	359	F	x
420	G	334	G	360	G	360	G	360	G	x
421	K	335	E	361	K	361	A	361	A	
422	G	336	G	362	G	362	G	362	G	x
423	N	337	N	363	N	363	N	363	N	x
424	F	338	F	364	F	364	F	364	F	x
425	S	339	K	365	K	365	N	365	N	
426	E	340	A	366	A	366	S	366	S	
427	L	341	L	367	L	367	L	367	L	x
428	F	342	F	368	F	368	F	368	F	x
429	K	343	E	369	E	369	K	369	K	
430	S	344	S	370	A	370	A	370	A	
431	I	345	I	371	I	371	F	371	F	
432	E	346	E	372	E	372	E	372	E	x
433	E	347	R	373	R	373	E	373	E	
434	Y	348	D	374	E	374	E	374	E	
435	E	349	Q	375	Q	375	Q	375	Q	
436	K	350	V	376	E	376	N	376	A	
437	T	351	R	377	K	377	L	377	L	

(continuación)

Pos.	1TFZ	Pos.	1CJX	Pos.	1T47	Pos.	3ISQ	Pos.	1SQI	Común
	<i>A. thaliana</i>		<i>P. fluorescens</i>		<i>S. avermitilis</i>		<i>H. sapiens</i>		<i>R. norvegicus</i>	aminoácido
438	L	352	R	378	R	378	R	378	R	
439	E	353	G	379	G	379	G	379	G	
440	A	354	V	380	N	380	N	380	N	
441	K	355	L	381	L	381	L	381	L	
442	Q	356	A			382	T	382	T	
443	L	357	T			383	N	383	D	
444	V	358	D			384	M	384	L	
445	G					385	E	385	E	
						386	T	386	T	
						387	N	387	N	
						388	G	388	G	
						389	V	389	V	
						390	V	390	R	
						391	P	391	S	
						392	G	392	G	
						393	M	393	M	
						394	A	394		
						395	E	395		
						396	N	396		
						397	L	397		
						398	Y	398		
						399	F	399		
						400	Q	400		

5 Un análisis de secuencia de más de 700 secuencias de HPPD de bases de datos públicas incluyendo secuencias de proteínas de HPPD y proteínas de HPPD pronosticadas tales como de plantas, mamíferos, hongos y bacterias se realizó usando ClustalX. La alineación se corrigió usando la información de las estructuras 3D disponibles. Idénticas secuencias de aminoácidos con diferentes identificadores se incluyeron solo una vez y algunas secuencias con errores de secuencia obvios fueron excluidas. La alineación también incluye secuencias incompletas. La Tabla 2b muestra la alineación de las secuencias para un grupo representativo de proteínas de HPPD e incluye secuencias de plantas, bacterias, mamíferos.

10

Tabla 2b: alineación de un grupo representativo de secuencias de HPPD

CLUSTAL X (1.81) alineación de secuencias múltiples

```

RATTUS_NORVEGICUS      ---MTT-----Y--SNKGPKE
HOMO-SAPIENS           -----GAKPE
XENOPUS-LAEVIS        ---MTS-----YTDKGPKE
BLEPHARISMA-JAPONICUS ---MTY-----YDKQETRPD
MYCOSPHAERELLA-GRAMINICOLA ---MAPGALLVTSQNG-----RTSFLYDSDGYVPAAPALVVG
ASPERGILLUS-FUMIGATUS ---MAPSAISTS-----PPPTDRVSSS
MAGNAPORTHE-GRISEA    ---MSPGAITESPRN-----SVVDHTSGLQVDSLAVQGP
CANDIDA-ALBICANS      -----
PICHIA-STIPITIS       ---LLKELPFLPTSS-----DPITEPDIDELLSDGHVNSKYP
RHODOCOCCUS-SP.       ---MTIEQTLTD-----KERLAGLDLQLEQLVGLVEYDTRD
RHODOCOCCUS-ERYTHROPOLIS ---MTVEQTLTD-----DEQLAGLDLEQLRQLVGLVEYDSDHD
JANIBACTER-SP.-HTCC2649 MTDTTITNPTTDTASQPTPLDLTPQEREANLNLEQLKQLVGLVEYDESKD
STREPTOMYCES-AVERMITILIS ---MTQTTHHTF-----DTARQAD
KORDIA-ALGICIDA       ---MAAEIKNLKD-----LQNTTEYGLKKLFDEAED
LEUWENHOEKIELLA-BLANDENSIS-ME ---MSKELK-----SVDYGLEKIFDGAQD
GEMMATIMONAS-AURANTIACA ---MATLTT-----PEIGTEQD
PICROPHILUS-TORRIDUS -----MYGKNLISELRE
ARABIDOPSIS-THALIANA -----MGHQNAAVSENQNHDDGAAS--SFGFKLVGFSKFKVRKNPKSD
BRASSICA-RAPA-SUBSP.-PEKINENSI -----MGHENA AVSENQHDDAATTASPGFKLVGFSKFKVRKNPKSD
COPTIS-JAPONICA-VAR.-DISSECTA -----MVPST-----AGNLKLVGHTNFVHNNPKSD
HEVEA-BRASILIENSIS    -----MGKENDSVPS-----APGFKLLGFSNFVRTNPRSD
MEDICAGO-TRUNCATULA   -----MAIETEQTQT-----QTGFKLVGFKNFVRANPKSD
GLYCINE-MAX           -----MCNEIQQAQAQ-----AQGFKLVGFKNFVRTNPKSD
SOLANUM-LYCOPERSICUM ---MG--MGKETLS--TTDTTG-----ATPKLVGFNNFIRANPRSD
SOLANUM-TUBEROSUM     ---MG--MGKETLSTTTTDETG-----ATPKLG-FNNFIRANPRSD
NICOTIANA-BENTHAMIANA ---MGKLETVTTTTSATAADDSEL----TTNFKLVGFKNFIRTNPRSD
DAUCUS-CAROTA         ---MGKKQSEAEILSSNSNTSP-----ATPKLVGFNNFVRANPKSD
SOLENOSTEMON-SCUTELLARIOIDES ---MG-----QESTAAA AVVP-----AEPKLVGHKNFVRSNPMDS
HORDEUM-VULGARE       ---MPPTPTTPAAT--GAAA AVTPRHARP----HRMVRFNPRSD
TRITICUM-AESTIVUM     ---MPPTPTTPAATGAGAAA AVTPEHARP----RRMVRFNPRSD
AVENA-SATIVA          ---MPPTPATATGA--AAA AVTPEHAARS---FPRVVRVNRSD
ORYZA-SATIVA          ---MPPTPTTATTGAVSAAAAAAGENAGFRLVGHRRFVRANPRSD
SORGHUM-BICOLOR       ---MPPTPTTAAAG--GA AVAAASAEQAAPRLVGHRRNFVRVNRSD
ZEA-MAYS              ---MPPTPTAAAAG--AA AVAAASAEQAAPRLVGHRRNFVRFNPRSD
ABO95005_OLUCIMARINUS ---MATVPS-----KRKLVGCANFVRSNPLSD
OTAURI                ---MTSAS-----GRKLVGHANFVRCNPLSD
MICROMONAS-PUSILLA-CCMP1545 ---MASSEANAAPA-----AKRHKLVGCKNFVRNPKSD
SYNECHOCOCCUS-SP.    -----MNPDIR
VIBRIO-SP.-MED222     -----MVDTYN
MARINOMONAS-SP.-MED121 -----MNTALKIDYGIN
PSEUDOMONAS-FLUORESCENS -----ADLYEN
SULFITOBACTER-SP.-NAS-14.1 ---MGFFPHDAEKSKITDEN
OCEANICOLA-BATSENSIS-HTCC2597 ---MGFFPHDAPKSEITDEN
BDELLOVIBRIO-BACTERIOVORUS -----MAQVTEKN

RATTUS_NORVEGICUS      RGRFLHFHSVTFWVGNAK--QAASFYCNKMGFEPLAYKGLTGSREVVSH
HOMO-SAPIENS           RGRFLHFHSVTFWVGNAK--QAASFYCSKMGFEPLAYRGLTGSREVVSH
XENOPUS-LAEVIS        VGRYLAFDHI TFYVGNK--QAAAYYATRFGFNP IAYRGLTGH RDVCTH
BLEPHARISMA-JAPONICUS LGEFYGFHHVRFYVNSSE--QAASFYTSRFGFSPVAYEGLETGNQKFCFN
MYCOSPHAERELLA-GRAMINICOLA EVNYRGYHHA EWVGNK--QVAQFYITRMGFEPVAHKGLTGSRF FASH
ASPERGILLUS-FUMIGATUS LASYKGYDHVHWYVGNK--QAASYITRMGFKRIAYRGLTGC RSVCSH
MAGNAPORTHE-GRISEA    FPSFHGYDHTVWVGNK--QAASYNTLFGMKI IAYRGLTGSRYFASY
CANDIDA-ALBICANS      -----MSKYLQLAMGFKEVAYKGLTGSKLI GAH
PICHIA-STIPITIS       TDGFIKFFSLKICSSNAK--QMSKYLQLAMDFKEIAYKGLENDSRLVGAH
RHODOCOCCUS-SP.       PFPVSGWD AVVWVGNAT--QTAHYFQS AFGMTLVAYS GPTTGNRDHHSF
RHODOCOCCUS-ERYTHROPOLIS PFPVSGWDGLEWIVGNAT--QTSHFFQS AFGMELVAYS GPSTGNRDHHSF
    
```

JANIBACTER-SP.-HTCC2649
 STREPTOMYCES-AVERMITILIS
 KORDIA-ALGICIDA
 LEEUWENHOEKIELLA-BLANDENSIS-ME
 GEMMATIMONAS-AURANTIACA
 PICROPHILUS-TORRIDUS
 ARABIDOPSIS-THALIANA
 BRASSICA-RAPA-SUBSP.-PEKINENSI
 COPTIS-JAPONICA-VAR.-DISSECTA
 HEVEA-BRASILIENSIS
 MEDICAGO-TRUNCATULA
 GLYCINE-MAX
 SOLANUM-LYCOPERSICUM
 SOLANUM-TUBEROSUM
 NICOTIANA-BENTHAMIANA
 DAUCUS-CAROTA
 SOLENOSTEMON-SCUTELLARIOIDES
 HORDEUM-VULGARE
 TRITICUM-AESTIVUM
 AVENA-SATIVA
 ORYZA-SATIVA
 SORGHUM-BICOLOR
 ZEA-MAYS
 ABO95005_OLUCIMARINUS
 OTAURI
 MICROMONAS-PUSILLA-CCMP1545
 SYNECHOCOCCUS-SP.
 VIBRIO-SP.-MED222
 MARINOMONAS-SP.-MED121
 PSEUDOMONAS-FLUORESCENS
 SULFITOBACTER-SP.-NAS-14.1
 OCEANICOLA-BATSSENSIS-HTCC2597
 BDELLOVIBRIO-BACTERIOVORUS

RATTUS_NORVEGICUS
 HOMO-SAPIENS
 XENOPUS-LAEVIS
 BLEPHARISMA-JAPONICUS
 MYCOPHAERELLA-GRAMINICOLA
 ASPERGILLUS-FUMIGATUS
 MAGNAPORTHE-GRISEA
 CANDIDA-ALBICANS
 PICHIA-STIPITIS
 RHODOCOCCUS-SP.
 RHODOCOCCUS-ERYTHROPOLIS
 JANIBACTER-SP.-HTCC2649
 STREPTOMYCES-AVERMITILIS
 KORDIA-ALGICIDA
 LEEUWENHOEKIELLA-BLANDENSIS-ME
 GEMMATIMONAS-AURANTIACA
 PICROPHILUS-TORRIDUS
 ARABIDOPSIS-THALIANA
 BRASSICA-RAPA-SUBSP.-PEKINENSI
 COPTIS-JAPONICA-VAR.-DISSECTA
 HEVEA-BRASILIENSIS
 MEDICAGO-TRUNCATULA
 GLYCINE-MAX
 SOLANUM-LYCOPERSICUM
 SOLANUM-TUBEROSUM
 NICOTIANA-BENTHAMIANA
 DAUCUS-CAROTA
 SOLENOSTEMON-SCUTELLARIOIDES
 HORDEUM-VULGARE
 TRITICUM-AESTIVUM

PPFVPTGWDAIVFVVGNA--QAAAYYQGTWGMELVGYSGPENGRDHKAF
 PFPVKGMDAVFFAVGNAK--QAAHYSTAFGMQLVAYSGPENGSRETASY
 FLPLLLGTDYVELYVGNK--QSAHFYKTAFGFQSEAYAGLETGLTDRVSY
 FLPLLLGTDYVEFYVGNK--QAAHFYKTAFGFQSEAYSGLETGKKDRVSY
 APFINGTLDYVEFYVGNK--QASHYRAAFGYSLVAYRGPETGVRDRASY
 KEIFKRLHHEVEFYVSGAK--TWSYFMNRGLGPKTVAYAGPETGIRDKI SY
 KFKVKRFHHIEFWCGDAT--NVARRFWSWGLGMRFSAKSDLSTGNMVHASY
 KFKVKRFHHIEFWCGDAT--NVARRFWSWGLGMRFSAKSDLSTGNMVHASY
 KFHVKRFHHIEFWSTDAT--NTARRFSWGLGMPVAKSDLSTGNMVHASY
 LFKVKRFHHVEFWCTDAT--NTACRFWSWGLGMPVAKSDLSTGNVTHASY
 RFNVKRFHHVEFWCTDAT--NTARRFSWGLGMPVAKSDLSTGNLTHASY
 RFQVNRFFHHIEFWCTDAT--NASRRFSWGLGMPVAKSDLSTGNQIHASY
 FFSVKRFHHIEFWCGDAT--NTSRRFSWSLGMPIITAKSDLSTGNSVHASY
 FFSVKRFHHIEFWCGDAT--NTSRRFSWSLGMPIITAKSDLSTGNSVHASY
 FFSVKRFHHIEFWCGDAT--NTSRRFSWSLGMPIITAKSDLSTGNSVHASY
 HFAVKRFHHIEFWCGDAT--NTSRRFSWGLGMPVAKSDLSTGNSVHASY
 HFPVHRFFHHVEFWCGDAT--NTSRRFSWGLGMPVAKSDLSTGNSAHASY
 RFHTLSFHHVEFWCADAA--SAAGRFALGAPLAARSDDLSTGNSAHAQ
 RFHTLSFHHVEFWCADAA--SAAGRFALGAPLAARSDDLSTGNSVHASQ
 RFPVLSFHHVELWCADAA--SAAGRFALGAPLAARSDDLSTGNSAHAQL
 RFQALAFHHVELWCADAA--SAAGRFALGAPLAARSDDLSTGNSAHAQL
 RFHTLAFHHVELWCADAA--SAAGRFSGLGAAPLAARSDDLSTGNTAHASL
 RFHTLAFHHVELWCADAA--SAAGRFSGLGAAPLAARSDDLSTGNSAHAQL
 AFECEKFDHIEFWCGDAT--NAAARFVGLGMGLRCKSDATTGNGTYASY
 AFECVGFHDHVEFWCGDAT--NAAARFVGLGMGLRCKSDATTGNGTYASY
 LFTMRKFFHHVEFWCRDAT--TTAARFVGLGMQLVAKSDLTGTNARYASY
 --IVQGIHLLHFYLDLPL--RWREHFCRVWGFVAVSDAGN-----TL
 PLGTDGFEFFVEYTAVDHKGIEQLKALLVSLGFABIAKHSKE-----AW
 PLGTDGFEFFVEYTADEKGIADLKALFVSLGFTEVAKHSKE-----VW
 PMGLMGFEFFIEFASPTPG---TLEPIFEIMGFTKVATHRSKN-----VH
 PAGTDGFEFFVEFASADPQ---ELRDLFTRMGYAHVANHKTQK-----IE
 PAGTDGFEFFVEFAHPDPQ---ELRDLFSKMGYELVGRHKSQD-----VE
 PVGLNGVDFIEYSGPDAH---FFEQVFKRYAFKEVGVHVGKN-----IK

VIKQK--IVFVLSALNPWN-----
 VIKQK--IVFVLSALNPWN-----
 VVRQNN--ATFVFQSPLNPGN-----
 VVRSNH--VVIFTSALTPED-----
 VVQNG--VRFVFTSPVRSAA-----
 VVRNGD--ITFILTSPLRS-----
 LVGKED--VRFVFTSPIRSHVH-----
 VMQNGS--ITLEIINTLETIDDDNVLKFPFFQNDLNKFRNINHEYFLENF
 VIRNGD--VTLEIINTLETIVEDDNVLFKFPYFEKDLKQFPQLNESKYLRDF
 VLESQA--VRFVIGAVNPDSP-----
 VLKSGA--VRFVVKGAVDPASP-----
 VLKSGS--IRFVLKAVSPDSP-----
 VLTNGS--ARFVLTSVVKPATPW-----
 VLKQDK--IRLVLTPLGKGE-----
 VLRQDK--IRLVLTSPGASAP-----
 LMQQGK--IRLVLTSTADTP-----
 VMSQGT--ARISFTSSMNDSDY-----
 LLTSGD--LRFLFTAPYSP---SLSAGEI-----
 LLTSGD--LRFLFTAPYSP---SLSAGEN-----
 LLRSGE--LNFLFTAPYSP---SIAGNT-----
 LLRSGD--LSFLFTAPYSP---TIASMENT-----
 LLRSGD--LNFLFSAAYSP---SISLSS-----
 LLRSGD--LSFLFSAAYSP---SLSAGSS-----
 LLRSVSGELQFVFTAPYSP---SISVPS-----
 LLRSVSGELQFVFTAPYSP---SISVPS-----
 LLRPVSGSLQFLFTAPYSP---SISTPS-----
 LVRSAN--LSFVFTAPYSP---STTSSG-----
 LLRSGE--LSFVFTAPYSP---SLAEPS-----
 LLRSGS--LAFFLFTAPYAN----GCDAA-----
 LLRSGN--LAFFLFTAPYAN----GCDAA-----

AVENA-SATIVA	LLRSGA--LAFLFTAPYAPPPQEAATAAA-----
ORYZA-SATIVA	LLRSAS--VAFLFTAPYGGDHGVDGADAAT-----
SORGHUM-BICOLOR	LLRSGA--LAFLFTAPYAH-----GADAA-----
ZEA-MAYS	LLRSGS--LSFLFTAPYAH-----GADAA-----
ABO95005_OLUCIMARINUS	AMKSND--LTFVFTAPYG-----VESGGSR-----
OTAURI	AMKSHD--LTFVFTAPYGDDERAVGCGGSS-----
MICROMONAS-PUSILLA-CCMP1545	VLQSDN--LRFVFSAPYDVPEGEENDDARS-----
SYNECHOCOCCUS-SP.	ELEQGS--LRLRLSQPARAG-----
VIBRIO-SP.-MED222	LYRQGD--INFIVNEQPHS-----
MARINOMONAS-SP.-MED121	LYRQND--INFIVNSEPNS-----
PSEUDOMONAS-FLUORESCENS	LYRQGE--INLILNNEPNS-----
SULFITOBACTER-SP.-NAS-14.1	LWQQGD--ITYVLNADPDS-----
OCEANICOLA-BATSENSIS-HTCC2597	LWQQGD--ITYIINAEPGT-----
BDELLOVIBRIO-BACTERIOVORUS	LFRQGD--INFILNCEPHT-----

RATTUS NORVEGICUS	-----
HOMO-SAPIENS	-----
XENOPUS-LAEVIS	-----
BLEPHARISMA-JAPONICUS	-----
MYCOSPHAERELLA-GRAMINICOLA	-----
ASPERGILLUS-FUMIGATUS	-----
MAGNAPORTHE-GRISEA	-----
CANDIDA-ALBICANS	KLTTDDLIFDFVNSRIEGLSFKSDYLKFGKQFYNNI IRSDDYQESMAKVS
PICHIA-STIPITIS	KITTNDLVDFVNSRIESFSVSPNAHYFRRKLYNKIVSSRAFRNMFYDN
RHODOCOCCUS-SP.	-----
RHODOCOCCUS-ERYTHROPOLIS	-----
JANIBACTER-SP.-HTCC2649	-----
STREPTOMYCES-AVERMITILIS	-----
KORDIA-ALGICIDA	-----
LEEUWENHOEKIELLA-BLANDENSIS-ME	-----
GEMMATIMONAS-AURANTIACA	-----
PICROPHILUS-TORRIDUS	-----
ARABIDOPSIS-THALIANA	-----
BRASSICA-RAPA-SUBSP.-PEKINENSI	-----
COPTIS-JAPONICA-VAR.-DISSECTA	-----
HEVEA-BRASILIENSIS	-----
MEDICAGO-TRUNCATULA	-----
GLYCINE-MAX	-----
SOLANUM-LYCOPERSICUM	-----
SOLANUM-TUBEROSUM	-----
NICOTIANA-BENTHAMIANA	-----
DAUCUS-CAROTA	-----
SOLENOSTEMON-SCUTELLARIOIDES	-----
HORDEUM-VULGARE	-----
TRITICUM-AESTIVUM	-----
AVENA-SATIVA	-----
ORYZA-SATIVA	-----
SORGHUM-BICOLOR	-----
ZEA-MAYS	-----
ABO95005_OLUCIMARINUS	-----
OTAURI	-----
MICROMONAS-PUSILLA-CCMP1545	-----
SYNECHOCOCCUS-SP.	-----
VIBRIO-SP.-MED222	-----
MARINOMONAS-SP.-MED121	-----
PSEUDOMONAS-FLUORESCENS	-----
SULFITOBACTER-SP.-NAS-14.1	-----
OCEANICOLA-BATSENSIS-HTCC2597	-----
BDELLOVIBRIO-BACTERIOVORUS	-----

RATTUS NORVEGICUS	-----KEMGDHLVKHGDGVKDI AFEVEDCEHIVQK
HOMO-SAPIENS	-----KEMGDHLVKHGDGVKDI AFEVEDCDYIVQK
XENOPUS-LAEVIS	-----HPISDHVAMHGDGVKDVAFSVEDCRGIYKR

BLEPHARISMA-JAPONICUS	-----NEVNRHVKGKSDGVQDIAFSVSDARGMYEK
MYCOSPHAERELLA-GRAMINICOLA	-----RQTLKAAPLADQARLDEMYDHLDKHGDGVKDVAFEVDDVLAVYEN
ASPERGILLUS-FUMIGATUS	-----LDQVDRFPPEEQELLKEIHAHLEKKGDKVAVFEVDSVDSVFYA
MAGNAPORTHE-GRISEA	-----LPEDEPISDEDRALLKEMHAHLEKKGDAVKDVCFEVDNVQGVYER
CANDIDA-ALBICANS	NFIVQTTINNSEEIYNDMMECTLIQKFLKLHAEGVMDIAFNVDVDTIFNR
PICHIA-STIPITIS	NLILNVINNSEVIYNDIMECTLIQKFLKTHGEGVMDISFLVEDVITIFDK
RHODOCOCCUS-SP.	-----LIDHHRTHGDGVVDIALAVPDVVKCIAH
RHODOCOCCUS-ERYTHROPOLIS	-----LIEHHSRHRGDIIRDIALSVDPVVKCIAH
JANIBACTER-SP.-HTCC2649	-----LIAHHTKHGDGVVDISLEVPDVKCIAQ
STREPTOMYCES-AVERMITILIS	-----GHFLADHVAEHGDGVVDLAI EVPDARAAHAY
KORDIA-ALGICIDA	-----INEHIDLHGDGVKVALWVEDATKAFEE
LEEUWENHOEKIELLA-BLANDENSIS-ME	-----INEHIVKHGDGVKVALWVEDARSAFEE
GEMMATIMONAS-AURANTIACA	-----IAEHVHRHGDGVRDYALWVDDARLAYET
PICROPHILUS-TORRIDUS	-----ISNHVKKHGDGVKIDIALEVDDLDEAKSL
ARABIDOPSIS-THALIANA	-----KPTTTASIPSFHDGSCRSFFSSHGLGVRVAIEVEDAESAFSI
BRASSICA-RAPA-SUBSP.-PEKINENSI	-----PPTTTASIPSFHDVTYRSFFSSHGLGVRVAIEVEDAEAAFSI
COPTIS-JAPONICA-VAR.-DISSECTA	-----LTHTASIPTYSHNLARLFASTHGLAVRAIAIEVQDAELAYNI
HEVEA-BRASILIENSIS	-----SHTATASIPTF SHEACRNFSAKHGLGVRVAIAIEVEDAEIAYNT
MEDICAGO-TRUNCATULA	-----PSSTAAIPTFSASTCFSSASHGLAVRAIEVEDAEVAFTT
GLYCINE-MAX	-----AASSASIPSFDAATCLAFAAKHGFGVRAIAIEVEDAEAAFSI
SOLANUM-LYCOPERSICUM	-----TAGIPSFSTPTYRDF TAKHGLGVRVAIEVENAYLAFSA
SOLANUM-TUBEROSUM	-----TAGIPSFSTSTHRDFTAKHGLGVRVAIEVENAYLAFSA
NICOTIANA-BENTHAMIANA	-----SAAIPSFSTSTHRSFAATHGLGVRVAIEVENAYTAFSA
DAUCUS-CAROTA	-----SAAIPSFASGFHSAFAKHGLAVRAIAIEVADVAAAFEA
SOLENOSTEMON-SCUTELLARIOIDES	-----SASIPTF SFSDHRAFTSSHGLAVRAIAIQVDSASSAYSIA
HORDEUM-VULGARE	-----TASLPFSADAAARRFSADHGLAVRSIALRVADAABEFRA
TRITICUM-AESTIVUM	-----TASLPFSADAAARRFSADHGLAVRSIALRVADAABEFRA
AVENA-SATIVA	-----TASIPSFADAAARTFAAAHGLAVRSVGVRVADAABEFRV
ORYZA-SATIVA	-----TASIPSFSPGAARRFAADHGLAVHVALRVADAADAFRA
SORGHUM-BICOLOR	-----TASLPFSAAEARFAADHGLAVRAVALRVADAEDAFRA
ZEA-MAYS	-----TAALPSFSAAAARRFAADHGLAVRAVALRVADAEDAFRA
ABO95005_OLUCIMARINUS	-----GEAPHPG---HEGRAMMRF FEKHGLAARAVGVRVKDAARAYEE
OTAURI	-----VNVPHPG---NERGAMMRF FERHGLAARAVGLRVGDARAAAYEE
MICROMONAS-PUSILLA-CCMP1545	-----SMFEKSGVLSHDP SFMRTFCERHGLAVRAVCLLVDDAAVAFYT
SYNECHOCOCCUS-SP.	-----DEVDRHLQRHGPGVVDVALAVGEQELPALA
VIBRIO-SP.-MED222	-----QAEAFARVHGPSVCGMAFRVNEATAAMEQ
MARINOMONAS-SP.-MED121	-----QSEAFAR IHGPSVCGMAFRVKDASLAMQH
PSEUDOMONAS-FLUORESCENS	-----IASYFAAEHGPSVCGMAFRVKDSQKAYNR
CULFITOBACTER-SP.-NAS-14.1	-----FAAGFVAEHGPCAPSMGWRVVDQAQALDH
OCEANICOLA-BATSENSIS-HTCC2597	-----HAATFIEEHGPCAPSMGWRVVDQAQAFDH
BDELLOVIBRIO-BACTERIOVORUS	-----FATDFAKLHGPCVNATGFRVIDADQAFKT
RATTUS_NORVEGICUS	ARERGAKIVREPWVEED-----KFGKVKFAVLQTYG-DTHTTLVEK---
HOMO-SAPIENS	ARERGAKIMREPWVEEQD-----KFGKVKFAVLQTYG-DTHTTLVEK---
XENOPUS-LAEVIS	AIERGAKSVPREPWEESD-----EFGTVVMATIQTYG-DTHTTFVER---
BLEPHARISMA-JAPONICUS	AIKAGCKSFREPQVLQD-----QFGSVIIASLQTYG-DTVHTLVQN---
MYCOSPHAERELLA-GRAMINICOLA	AVANGAESVSSPHTDSC-----DEGDVISAAIKTYG-DTHTTFIQR---
ASPERGILLUS-FUMIGATUS	ATNNGAKIVSQPRTLED-----DNGQVRVATIQTYG-ETHTTLVER---
MAGNAPORTHE-GRISEA	AVQQGAVSIAPPKTLSDK-----EHGSVTMAVIQTYG-DTHTLTSR---
CANDIDA-ALBICANS	AIKAGSGIIRLPKIIISD-----ENGVVKLATISIPNSDIQHTLIEN---
PICHIA-STIPITIS	AVAAGAGIIRLPKIIISD-----CNGSVRLGTISI PKTDIQHTLIEN---
RHODOCOCCUS-SP.	ARAQGATVLDPHDVTD-----DHGTVRLAAIATYG-DTRHTLVDR---
RHODOCOCCUS-ERYTHROPOLIS	AITQGATVLDSEPHDITD-----EHGTVRLASIAIATYG-ETRHTLVDR---
JANIBACTER-SP.-HTCC2649	AKAAGARVVQEAETVSD-----EFGSVRIGAIATYG-ETRHTLVQRTVD
STREPTOMYCES-AVERMITILIS	AIIEHGARSVAEPYELKD-----EHGTVVLAIAIATYG-KTRHTLVDR---
KORDIA-ALGICIDA	TTKRGAKPYMEPTKEED-----ENGYVIRSGIYTYG-ETVHVVFVER---
LEEUWENHOEKIELLA-BLANDENSIS-ME	AIARGAIPQEPQVYSD-----EHGEVVRSGIYTYG-ETVHVVFVER---
GEMMATIMONAS-AURANTIACA	AIARGAIPQEPQVYSD-----EHGEVVRSGIYTYG-ETVHVVFVER---
PICROPHILUS-TORRIDUS	IEKYGTKVS-KINEIKD-----GNGKIRTAIEIKTYG-ETVHTLIET---
ARABIDOPSIS-THALIANA	SVANGAIPSSPPIVLNE-----AVTIAEVKLYG-DVVLRYVSYKAE
BRASSICA-RAPA-SUBSP.-PEKINENSI	SVSNGAVPSSPPIVLND-----AVTIAEVKLYG-DVVLRYVSYKVA
COPTIS-JAPONICA-VAR.-DISSECTA	SVANGAKPSSSPIKLDE-----GVVLSEIQLYG-DVVLRYLSFKNT
HEVEA-BRASILIENSIS	SVARGALPMGGPITLDN-----RAVVVAEVHLYG-DVVLRYISYKNS
MEDICAGO-TRUNCATULA	SVNLGAI PPSPPVILEN-----NVKLAEVHLYG-DVVLRYVSYNDL
GLYCINE-MAX	SVAKGAEPASPPVLVDD-----RTGFAEVRLYG-DVVLRYVSYKDA

SOLANUM-LYCOPERSICUM SVARGAKPRFEPVTIDE-----HVAVAEVHLYG-DVVLRFVSLVKD
SOLANUM-TUBEROSUM SVSRGAKPRFEPVTIDE-----HVAVAEVHLYG-DVVLRFVSVFKD
NICOTIANA-BENTHAMIANA SVSRGAKPMFEPVTIDG-----QVAMAEVHLYG-DVVLRFVMS-LKD
DAUCUS-CAROTA SVARGARPASAPVELDD-----QAWLAEVELYG-DVVLRFVSPGRE
SOLENOSTEMON-SCUTELLARIOIDES AVSRGAKPVSPVVLADC-----ETAIAEVHLYG-DTVLRFVSCGSG
HORDEUM-VULGARE SRRRGARPAFAPVDLGR-----GFAPAEVELYG-DVVLRFVSHPDG
TRITICUM-AESTIVUM SVDGGARPAFSPVDLGR-----GFGFAEVELYG-DVVLRFVSHPPD
AVENA-SATIVA SVAGGARPAFAPADLGH-----GFGLAEVELYG-DVVLRFVSPDE
ORYZA-SATIVA SVAAGARPAFQPADLGG-----GFGLAEVELYG-DVVLRFVSHPDG
SORGHUM-BICOLOR SVAAGARPAFEPVELGL-----GFRLAEVELYG-DVVLRYVSPDD
ZEA-MAYS -----GFRLAEVELYG-DVVLRYVSPDG
ABO95005_OLUCIMARINUS AVKRGARGVLAPTTLTHTVDDGCAKGGQVIAEIELYG-DVVLRFVSNATDG
OTAURI AMKRGARGVLEPTEMRHEKHHDGCVMTQII SEVELYG-DVVLRFVSRADG
MICROMONAS-PUSILLA-CCMP1545 SGQHGR-----SPAFSSACDGFAR----VSEVELYG-DVVLRYVSPHAG
SYNECHOCOCCUS-SP. ELLRGR-----AQLAWIP-AAAALCLHTPYG
VIBRIO-SP.-MED222 AFKGGGEEYK-TEIGP-----MELSI PAIYGIG-ESLLYFVDRYGK
MARINOMONAS-SP.-MEDI21 ALANGAKEFS-GNLGA-----MELKLPVYVYIG-ESTLYFIDRYGD
PSEUDOMONAS-FLUORESCENS ALELGAQPIH-IDTGP-----MELNLPKIGIG-GAPLYLIDRFGE
SULFITOBACTER-SP.-NAS-14.1 AVSKGAEYK-GAG-----KVLDPVPAIKIG-GSLIYFVDQYYD
OCEANICOLA-BATSENSIS-HTCC2597 AVKNGATPYE-GDG-----KVMDVPAILGIG-GSLIYFIEDYEE
BDELLOVIBRIO-BACTERIOVORUS AVARGARPYEGNEHQK-----GATPPPAIYGIG-DSLIIYFMDQKNQ

RATTUS NORVEGICUS -INYTG-----RFLPGFEAPTYKDT-----LLPKLPSCNLEIIDHIVGNQ
HOMO-SAPIENS -MNYIG-----QFLPGYEAPAFMDP-----LLPKLPKCSLEMIIDHIVGNQ
XENOPUS-LAEVIS -TNYKAP--AHVFLPNYRA-CEVDP-----INNVLPTVKLLNVDHVVGNQ
BLEPHARISMA-JAPONICUS -VDYTG-----PFLPGFRAITKDDP-----LNSAFPQVNYDII DHVVGNQ
MYCOSPHAERELLA-GRAMINICOLA -TTYTG-----PFLPGYRSCITVDS-----ANKFLPVPVLEAIDHCVGNQ
ASPERGILLUS-FUMIGATUS -GSYHG-----AFPLPGYRMTGVEDP-----ISQLLPGVHLNRI DHCVGNQ
MAGNAPORTHE-GRISEA -DNFRG-----TFLPGFRDVRQPAA----YSALAP-VLQRI DHCVGNQ
CANDIDA-ALBICANS -INYTG-----PFLPGFSQPIYPLADYYQVQLNMMPPVNLTVLDHCVENY
PICHIA-STIPITIS -IDYTG-----PFLPNYSESVTQYNSKYDQMQNIPTVSEFCIDHCVENY
RHODOCOCCUS-SP. -SHYTG-----PYLPGYTARTSGHT-----KRDGAPKRLFQALDHVVGNV
RHODOCOCCUS-ERYTHROPOLIS -SRYTG-----PYLPGYVERTSSYR-----KRDGAPKRLFQALDHVVGNV
JANIBACTER-SP.-HTCC2649 GQTYSG-----PYLPGYVARSSSFV-----KRDGAPKRLFQALDHIVGNV
STREPTOMYCES-AVERMITILIS -TG YDG-----PYLPGYVAAAP-----IVEPPAHRTFQAI DHCVGNV
KORDIA-ALGICIDA -KNYNG-----VFLPGYQRWES-----HYNPEPVGLKFI DHMVGNV
LEEUVENHOEKIELLA-BLANDENSIS-ME -KNYNG-----QFLPGYRKWES-----DYNPEPTGLKYIDHVMGNV
GEMMATIMONAS-AURANTIACA -RNYNG-----VFLPGFKAVTP-----HYQPSDVGLKYIDHCVGNV
PICROPHILUS-TORRIDUS -GDYNG-----VFMPGYEESI-----NSKNTGIKKI DHIVGNV
ARABIDOPSIS-THALIANA DTE-----KSEFLPGFERVEDASSF-----P-LDYGIRRLDHAVGNV
BRASSICA-RAPA-SUBSP.-PEKINENSI TV-----FLPRFETVDDTSSF-----P-LDYGIRRLDHAVGNV
COPTIS-JAPONICA-VAR.-DISSECTA N-----QSCP-----FLPGFEEVGEVSSS-----RGLDFGIRRLDHAVGNV
HEVEA-BRASILIENSIS NP-----NQNPNLFFLPGFERVSDSSN-----SSLDFGIRRLDHAVGNV
MEDICAGO-TRUNCATULA APQAPHADPSRWFPLPGFEAAASSSF-----PELDYGI RRLDHAVGNV
GLYCINE-MAX ADTL-----IFLPGFEAMDETSSF-----KELDYGI RRLDHAVGNV
SOLANUM-LYCOPERSICUM EDL-----IFLPGFEAMDETSSF-----KELDYGI RRLDHAVGNV
SOLANUM-TUBEROSUM ADL-----VFLPGFNAMDETASY-----KELDYGI RRLDHAVGNV
NICOTIANA-BENTHAMIANA EG-----LFLPGFEAVEGTASF-----PDLDYGI RRLDHAVGNV
DAUCUS-CAROTA ADG-----WFLPGFEVVGDVSC-----QELDYGI RRLDHAVGNV
SOLENOSTEMON-SCUTELLARIOIDES TDVP-----FLPGFEGVTNP-----DAVDYGLTRFDHVVGNV
HORDEUM-VULGARE TDVP-----FLPGFEGVSNP-----DAVDYGLTRFDHVVGNV
TRITICUM-AESTIVUM TDLP-----FLPGFERVSSP-----GAVDYGLTRFDHVVGNV
AVENA-SATIVA ADAP-----FLPGFEGVSNP-----GAVDYGLTRFDHVVGNV
ORYZA-SATIVA ADAS-----FLPGFVGVTSF-----GAADYGLRRFDHIVGNV
SORGHUM-BICOLOR AAGR-----FLPGFEGVASP-----GAADYGLTRFDHIVGNV
ZEA-MAYS -----FDGDFLCNYSATRDAP-----DVSYGLQRLDHAVGNV
ABO95005_OLUCIMARINUS -----FDGDFLCNYEATRDVP-----SVSYGLRRLDHAVGNV
OTAURI EKNA----KPKTFLPGYEDVPLEPPHT-----TPLTYGLQRLDHAVGNV
MICROMONAS-PUSILLA-CCMP1545 IRHS-----LIPGLDAAAPAEAG-----LFSHWDHVVILNV
SYNECHOCOCCUS-SP. -QSIYD---VDFRFYDQAEQRMAEA-----NVGLYEIDHLLTHNV
VIBRIO-SP.-MED222 -KSIYD---VDFNFYENYQEKMLSH-----QAGLYEVDHLLTHNV
MARINOMONAS-SP.-MEDI21 GSSIYD---IDFVYLEGVERNVPV-----AGLKVIDHLLTHNV
PSEUDOMONAS-FLUORESCENS -TSPYN---EEYDWIAQS--KPA-----VGFYLLDHLTHNV
SULFITOBACTER-SP.-NAS-14.1 -TSPYN---AEFDWLAQS--KPRG-----VGFYLLDHLTHNV
OCEANICOLA-BATSENSIS-HTCC2597

BDELLOVIBRIO-BACTERIOVORUS -DKLYN----EIFQVKPED-KAPVG-----VGFTVVDHFTMNV

RATTUS_NORVEGICUS PDQEMESASEWYLKNLQFHRFWSVDDTQVHTEYSSLRISIVVANYEESIKM
HOMO-SAPIENS PDQEMVSASEWYLKNLQFHRFWSVDDTQVHTEYSSLRISIVVANYEESIKM
XENOPUS-LAEVIS PDDMMVPVAEWYKMLMFHRFWSVDDTQMHTDYSALRSIVVTDYDEVIKM
BLEPHARISMA-JAPONICUS PGGDMTPTVEWYKYLEFHRYWSADESVIHTDYSALRSVVDWDEVIKM
MYCOSPHAERELLA-GRAMINICOLA DWDEMSDACDFYERCLGFHRFWSVDDKQICTEYSALKSIVMSSPNQVVKM
ASPERGILLUS-FUMIGATUS DWDEMDCVCEYERKALGFHRFWSVDDKQICTEYSALKSIVMASPNQVVKM
MAGNAPORTHE-GRISEA DWDDMRAACDFYERCLSFHRFWSVDDNQISTDFALNSIVMASPNNVVKM
CANDIDA-ALBICANS SWNQOMMEQAKLYADMFGFHKYWSVDEDDISTGFTALRSIVMSSNGQIKM
PICHIA-STIPITIS SWNQOMMAQAKLYASLFGFHKYWSADDHDIATDNTALRSIVMASGNGKIKM
RHODOCOCCUS-SP. ELGKMDHWVDFYNRVMGFTNMAEFVGGEDIATDYSALMSKVVVSGNHRVKF
RHODOCOCCUS-ERYTHROPOLIS ELGKMDQWVDFYNRVMGFTNMAEFVGGEDIATDYSALMSKVVVSGNHRVKF
JANIBACTER-SP.-HTCC2649 ELGKMDQWVDFYNRVMGFTNMAEFVGGEDIATDYSALMSKVVVSGNHRVKF
STREPTOMYCES-AVERMITILIS ELGRMNEWVGFYKVMGFTNMAEFVGGEDIATEYSALMSKVVADGTLKVKF
KORDIA-ALGICIDA GWGEMKEWCEFYAKVMGFAQII SFTDDDI STDFALMSKVVMSNGRIKF
LEEUEWENHOEKIELLA-BLANDENSIS-ME GWGEMNTVVKWYEDVMGFVNFITFDDKQITTEYSALMSKVVMSNGRIKF
GEMMATIMONAS-AURANTIACA ELGKMNQWVGYADVLGFRNLI TFDTDINTEYSSLSKVMANGNDRIKF
PICROPHILUS-TORRIDUS YEGEMSWVNFYI EKLGFHELI TFDKDIRTDYSALRSKVVKY-NDDIVF
ARABIDOPSIS-THALIANA P--ELGPALTYVAGFTCFHQFAEFTADDVGTAEGLNSAVLANNDVLL
BRASSICA-RAPA-SUBSP.-PEKINENSIS P--ELGPALTYLRLTCFHQFAEFTADDVGTAEGLNSAVLANNDVLL
COPTIS-JAPONICA-VAR.-DISSECTA P--NLAEAIGYLKFTCFHEFAEFTAEVGTESGLNSIVLANSDEMVL
HEVEA-BRASILIENSIS P--ELAPAVSYVKEFTCFHEFAEFTAEVGTESGLNSIVLANNEDTVLL
MEDICAGO-TRUNCATULA P--ELSSAVKYVKQFTCFHEFAEFTAEVGTESGLNSIVLANNEDTVLL
GLYCINE-MAX P--ELAPAVRYLKGFSGFHEFAEFTAEVGTESGLNSIVLANNSETVLL
SOLANUM-LYCOPERSICUM P--ELGPVVDYIKFTCFHEFAEFTAEVGTAEGLNSIVLANNEDTVLL
SOLANUM-TUBEROSUM P--ELGPVVDYIKFTCFHEFAEFTAEVGTAEGLNSIVLANNEDTVLL
NICOTIANA-BENTHAMIANA P--ELGPAVDYIKRFTCFHEFAEFTSEVGTAEGLNSMVVANNDVLL
DAUCUS-CAROTA T--ELGPVVEYIKGFTCFHEFAEFTAEVGTLEGLNSIVLANNEDTVLL
SOLENOSTEMON-SCUTELLARIOIDES P--KLEPVVDYIKKFTCFHEFAEFTAEVGTAEGLNSIVLANNENNVLF
HORDEUM-VULGARE P--ELAPAAAYIAGFTCFHEFAEFTAEVGTESGLNSIVLANNSEGVLL
TRITICUM-AESTIVUM P--ELAPAAAYVAGFAGCFHEFAEFTEDVGTAEGLNSIVLANNSEGVLL
AVENA-SATIVA P--EMAPVIDYMKGRVLCGFHEFAEFTAEVGTAEGLNSIVLANNSEAVLL
ORYZA-SATIVA P--ELAPVAAAYISGFTCFHEFAEFTAEVGTAEGLNSIVLANNAAETVLL
SORGHUM-BICOLOR P--ELAPAAAYFAGFTCFHEFAEFTAEVGTESGLNSIVLANNAAENVLL
ZEA-MAYS P--ELAPAAAYFAGFTCFHEFAEFTEDVGTAEGLNSIVLANNSENVLL
ABO95005_OLUCIMARINUS H--DLIETVDYITKVTCFHEFAEFTAEIDGTTIDSGLSMVLANNMEYVLL
OTOURI H--NLLETVDYIMKITCFHEFAEFTAEIDGTTIDSGLSMVLANNMEYVLL
MICROMONAS-PUSILLA-CCMP1545 P--NLLETVDYITAMTGMHEFAEFTAEVGTVDGSLNSMVLANNDEMILL
SYNECHOCOCCUS-SP. EQGSLQAADDFYKGRVLCWRRLYRYS---IGTATSGLESVVGDPFAGIQW
VIBRIO-SP.-MED222 KQGNMVDVWSGFYERLGNFREIRYFDIEGKLTG---LVSRAMTSPCGKIRI
MARINOMONAS-SP.-MED121 MRGNMDHWAGFYENIGNFREIRYFDIEGKLTG---LVSRAMTSPCGKIRI
PSEUDOMONAS-FLUORESCENS YRGRMVYWANFYKLFNFREARYFDIKGEYTG---LTSKAMSPDGMIRI
SULFITOBACTER-SP.-NAS-14.1 FKGNMVDVWFKFYGDLNFREIRFFDIEGKFTG---LTSRALTSPCGRIRI
OCEANICOLA-BATSENSIS-HTCC2597 FKGNMDTWFRFYGDLNFREIRFFDIQKGYTG---LFSRALTSPCGRIRI
BDELLOVIBRIO-BACTERIOVORUS PKGEMDKWQHPYEDIFGFYEAKYFDIRGSKTG---LLSRAMRSPCGKFSV

RATTUS_NORVEGICUS PINEPAPGRK-KSQIQEYVDYNGGAGVQHIALRTEDIITTIRHLRER---
HOMO-SAPIENS PINEPAPGKK-KSQIQEYVDYNGGAGVQHIALKTEDIITAIRHLRER---
XENOPUS-LAEVIS PINEPAPGKK-KSQIQEFVEYGGAGVQHIALRTDDILRDVSAMRAR---
BLEPHARISMA-JAPONICUS PINEPADGLR-KSQIQEYVEYGGAGVQHIALKVNDIISVISTLRAR---
MYCOSPHAERELLA-GRAMINICOLA PINEPAHGKK-KSQIEEYVDFYNGPGVQHIALRTPNII EAVSNLRSR---
ASPERGILLUS-FUMIGATUS PINEPAKGGK-QSQIEEYVDFYNGAGVQHIALLTDDIIRDITNLKAR---
MAGNAPORTHE-GRISEA PINEPAKGGK-RSQIEEYVTFNSGAGVQHIALLTSDIITTVAMRSR---
CANDIDA-ALBICANS PINEPVKSTM-KQIEEFFNDFNGGPGTQHTAFRTNNTIETVMAIMQR---
PICHIA-STIPITIS PINEPVKSKM-RQIEEFFHDFNGGPGVQHIALRTNDIIDTVCALLAR---
RHODOCOCCUS-SP. PLNEPALAKK-RSQIDEYLDYRGPQAHLALATNDILTAVDQLTAE---
RHODOCOCCUS-ERYTHROPOLIS PLNEPAIAKK-RSQIDEYLDYRGPQAHLALATNDILRSVDELKRE---
JANIBACTER-SP.-HTCC2649 PLNEPAIAKK-RSQIDEYLDYRGPQAHLAVATNDILRSVDELKRE---
STREPTOMYCES-AVERMITILIS PINEPALAKK-KSQIDEYLDYRGGAGVQHIALNTGDIVETVRTMRAA---
KORDIA-ALGICIDA PINEPAEGKK-KSQIEEYLDYRNGSGVQHIAVATNDIIDTVSQMRER---
LEEUEWENHOEKIELLA-BLANDENSIS-ME PINEPAEGIK-KSQIEEYLDYRGGAGVQHIALVATDDIVKTVAAALKAR---
GEMMATIMONAS-AURANTIACA PINEPASGKK-KSQIEEYLDYRGGAGVQHIALATDDILATVTLALDR---
PICROPHILUS-TORRIDUS PINEPAKGLR-KSQIEEYLDYRSEGVQHIALLTDDIIKTVSMMEEN---

ARABIDOPSIS-THALIANA
 BRASSICA-RAPA-SUBSP.-PEKINENSI
 COPTIS-JAPONICA-VAR.-DISSECTA
 HEVEA-BRASILIENSIS
 MEDICAGO-TRUNCATULA
 GLYCINE-MAX
 SOLANUM-LYCOPERSICUM
 SOLANUM-TUBEROSUM
 NICOTIANA-BENTHAMIANA
 DAUCUS-CAROTA
 SOLENOSTEMON-SCUTELLARIOIDES
 HORDEUM-VULGARE
 TRITICUM-AESTIVUM
 AVENA-SATIVA
 ORYZA-SATIVA
 SORGHUM-BICOLOR
 ZEA-MAYS
 ABO95005_OLUCIMARINUS
 OTAURI
 MICROMONAS-PUSILLA-CCMP1545
 SYNECHOCOCCUS-SP.
 VIBRIO-SP.-MED222
 MARINOMONAS-SP.-MED121
 PSEUDOMONAS-FLUORESCENS
 SULFITOBACTER-SP.-NAS-14,1
 OCEANICOLA-BATSENSIS-HTCC2597
 BDELLOVIBRIO-BACTERIOVORUS

RATTUS_NORVEGICUS
 HOMO-SAPIENS
 XENOPUS-LAEVIS
 BLEPHARISMA-JAPONICUS
 MYCOSPHAERELLA-GRAMINICOLA
 ASPERGILLUS-FUMIGATUS
 MAGNAPORTHE-GRISEA
 CANDIDA-ALBICANS
 PICHIA-STIPTIS
 RHODOCOCCUS-SP.
 RHODOCOCCUS-ERYTHROPOLIS
 JANIBACTER-SP.-HTCC2649
 STREPTOMYCES-AVERMITILIS
 KORDIA-ALGICIDA
 LEEUWENHOEKIELLA-BLANDENSIS-ME
 GEMMATIMONAS-AURANTIACA
 PICROPHILUS-TORRIDUS
 ARABIDOPSIS-THALIANA
 BRASSICA-RAPA-SUBSP.-PEKINENSI
 COPTIS-JAPONICA-VAR.-DISSECTA
 HEVEA-BRASILIENSIS
 MEDICAGO-TRUNCATULA
 GLYCINE-MAX
 SOLANUM-LYCOPERSICUM
 SOLANUM-TUBEROSUM
 NICOTIANA-BENTHAMIANA
 DAUCUS-CAROTA
 SOLENOSTEMON-SCUTELLARIOIDES
 HORDEUM-VULGARE
 TRITICUM-AESTIVUM
 AVENA-SATIVA
 ORYZA-SATIVA
 SORGHUM-BICOLOR
 ZEA-MAYS
 ABO95005_OLUCIMARINUS
 OTAURI
 MICROMONAS-PUSILLA-CCMP1545

PINEPVHGTKRKRSQIQTYLEHNEGAGLQHLALMGEDIFRTLREMRKRSS-
 PVNEPVHGTKRKRSQIQTYLEHNEGAGVQHLALMGEDIFRTLREMRKRSG-
 PMNEPVYGTKRKRSQIQTYLEHNEGAGVQHLALVSEDIFFTLREMRRRSG-
 PLNEPVFGTKRKSQIQTYLEHNEGAGLQHLALVSEDIFKTLREMRRRSG-
 PMNEPVYGTKRKRSQIETYLEHNEGAGLQHLALMGADIFRTLREMRKRSG-
 PLNEPVYGTKRKRSQIETYLEHNEGAGVQHLALVTHDIFFTLREMRKRFS-
 PLNEPVYGTKRKRSQIQTYLEHNEGAGVQHLALVTEIDIFRTLREMRKRSG-
 PMNEPVYGTKRKRSQIQTYLEHNEGAGVQHLALVTEIDIFRTLREMRKRSG-
 PLNEPVYGTKRKRSQIQTYLEHNEGAGVQHLALVTEIDIFKTLKEMRKRSG-
 PLNEPVYGTKRKRSQIQTYLEHNEGAGVQHLALVSEDIFRTLREMRKRSC-
 PLNEPVYGTKRKRSQIQTYLDHNEGAGVQHLALITEDIFRTLREMRKRSE-
 PLNEPVHGTKRKRSQIQTFLEHHGGPGVQHI AVASSDVLRTLRKMFARSA-
 PLNEPVHGTKRKRSQIQTFLEHHGGSGVQHI AVASSDVLRTLRMRARSA-
 PLNEPVHGTKRKRSQIQTYLEYHGGPGVQHI ALASNDVLRTLRMRARSA-
 PLNEPVHGTKRKRSQIQTYLDHGGPGVQHI ALASDDVLGTLRMRARSA-
 PLNEPVHGTKRKRSQIQTYLDHGGPGVQHMALASDDVLRTLRMQARSA-
 PLNEPVHGTKRKRSQIQTFLDHGGPGVQHMALASDDVLRTLRMQARSA-
 PVNEPTFGTKRKSQIQTYLEQNNGPGVQHLALKTDDIFATVREMRKYSHL
 PVNEPTFGTKRKSQIQTYLEQNNGPGVQHLALKTDDIFATVREMRKYSHM
 AINEPTCAAS---QIQEFLHAHGGPGVQHI ALASNDIVASLRRLR----
 PINE-SSDDK--SQIEEFIREYNGEGIQHIALATDDIYKTVKTLRDR---
 PINE-SSDDK--SQIEEFNLQYNGEGIQHIAMSSDIYETVRQLKAG---
 PLNEESSKGA--GQIEEFMLQFNNGEGIQHVAFLTDDLVTWDLKAKKI---
 PINE-DRDEK--GQIVAYLKKYNGEGIQHIAVGARNIYDATDAIADN---
 PINE-DRGET--GQIVAYLKKYNGEGIQHIAVGARDIYAATDAIAEM---
 PINE-PTEEK--SQIQEYLDEYKSGIQHIALLTHDINYSLESLNK---

--GMEFLAVP-SSYYRIL-----RENKLTSK--IQVKEN-----MDV
 --GLEFLSVP-STYYKQL-----REKLKTAK--IKVKEN-----IDA
 --GLEFLTIP-RTYYKNL-----RARLSMSK--VQVEED-----LAE
 --GVEFLEVP-PKYDLSL-----RKRLAHS--VQIEED-----LKR
 --GVEFISVP-DTYVENM-----RLRLKAAG--MKLEES-----FDI
 --GVEFIKVP-DTYEDI-----KVRLKAG--LTLHED-----FET
 --GVEFIEVP-HTYYDTM-----RRRLKTEKRDWELQED-----FDR
 --GVEFNHTS-ENYNNL-----KQRLNNDG--IKLYED-----FDT
 --GIEFNHTS-DKYTNL-----ERLLREDD--VALFED-----FDT
 --GVEFLATP-DSYYEDP-----ELRARIG--NVRAP-----IAE
 --GIEFLSTP-ASYEDP-----ELRARIG--EVRVP-----IEE
 --GVEFLDTP-DYYNDP-----EMRARIG--EVRVP-----IEE
 --GVQFLDTP-DSYYDT-----LGEWVG--DTRVP-----VDT
 --GVEFLYVP-DTYDD-----LLERVG--DIDED-----VEE
 --GVEFLPPPQAYDD-----IPRRLGAHMDTMKED-----LNK
 --GVEFLSVP-TSYYED-----LQERVG--KIDEK-----LEE
 --GIEFLKTP-GSYYES-----LSSRIG--SIDED-----LNE
 IGGFDFMPSPPPTYQN-----LKKRVG--DVLSD-----QIKE
 VGGFDFMPSPPPTYKN-----LKNRVG--DVLSEE-----QIEE
 VGGFDFMPSPPPTYKN-----LKNRAG--DVLSD-----QIKE
 VGGFDFMPSPPPTYRN-----LKNRVG--DVLTD-----QIKE
 VGGFDFMPSPPPTYRN-----LKNRVG--DVLSD-----QIKE
 LGGFDFMPSPPPTYAN-----LHNRAA--DVLTV-----QIKQ
 VGGFDFMPPPTYKN-----LRSRAG--DVLSD-----QIQ
 VGGFDFMPSPPPTYKN-----LKSRA--DVLSD-----QIQ
 VGGFDFMPSPPPTYKN-----LKNRAG--DVLTD-----QIQ
 LGGFDFMPSPPPTYKN-----LKNRVG--DVLSD-----QIKE
 VGGFDFMPSPPPTYRN-----LKSRA--DVLSD-----QIFE
 MGGFDFLPPPLPKYYE-----VRRLAG--DVLSEA-----QIKE
 MGGFDFLPPRCRKYEG-----VRRRAG--DVLSEA-----QIKE
 MGGFDFMAPPQAKYYE-----VRRRAG--DVLSEE-----QIKE
 MGGFDFLAPPNNYDC-----VRRRAG--DVLSEE-----QINE
 MGGFDFMAPPPEYDC-----VRRRAG--DVLTEA-----QIKE
 MGGFDFMAPPPTSDYDC-----VRRRAG--DVLTEA-----QIKE
 RGGFDFQAPASDDYKQ-----LKKRIG--DALNDE-----QYAL
 HGGFDFQAPASDDYKH-----LKEKIG--DALTDE-----QYAL
 RGGFDFQKPSADYYAN-----LKVGE--DALTER-----QFKE

SYNECHOCOCCUS-SP.
 VIBRIO-SP.-MED222
 MARINOMONAS-SP.-MED121
 PSEUDOMONAS-FLUORESCENS
 SULFITOBACTER-SP.-NAS-14.1
 OCEANICOLA-BATSENSIS-HTCC2597
 BDELLOVIBRIO-BACTERIOVORUS

RATTUS_NORVEGICUS
 HOMO-SAPIENS
 XENOPUS-LAEVIS
 BLEPHARISMA-JAPONICUS
 MYCOSPHAERELLA-GRAMINICOLA
 ASPERGILLUS-FUMIGATUS
 MAGNAPORTHE-GRISEA
 CANDIDA-ALBICANS
 PICHIA-STIPITIS
 RHODOCOCCUS-SP.
 RHODOCOCCUS-ERYTHROPOLIS
 JANIBACTER-SP.-HTCC2649
 STREPTOMYCES-AVERMITILIS
 KORDIA-ALGICIDA
 LEEUWENHOEKIELLA-BLANDENSIS-ME
 GEMMATIMONAS-AURANTIACA
 PICROPHILUS-TORRIDUS
 ARABIDOPSIS-THALIANA
 BRASSICA-RAPA-SUBSP.-PEKINENSI
 COPTIS-JAPONICA-VAR.-DISSECTA
 HEVEA-BRASILIENSIS
 MEDICAGO-TRUNCATULA
 GLYCINE-MAX
 SOLANUM-LYCOBERSICUM
 SOLANUM-TUBEROSUM
 NICOTIANA-BENTHAMIANA
 DAUCUS-CAROTA
 SOLENOSTEMON-SCUTELLARIOIDES
 HORDEUM-VULGARE
 TRITICUM-AESTIVUM
 AVENA-SATIVA
 ORYZA-SATIVA
 SORGHUM-BICOLOR
 ZEA-MAYS
 ABO95005_OLUCIMARINUS
 OTAURI
 MICROMONAS-PUSILLA-CCMP1545
 SYNECHOCOCCUS-SP.
 VIBRIO-SP.-MED222
 MARINOMONAS-SP.-MED121
 PSEUDOMONAS-FLUORESCENS
 SULFITOBACTER-SP.-NAS-14.1
 OCEANICOLA-BATSENSIS-HTCC2597
 BDELLOVIBRIO-BACTERIOVORUS

OGGVDFLQVAP-QYYTS-----LERELGL--ALRSALGQAISWQD
 --GMDFMPTP-DTYYEKVDVDRVKGHG-----ED-----TDL
 --GLKFMSTP-DTYYAKVNDRVVGHG-----ED-----LEK
 --GMRFMTAPPDTYYEMLEGRLPDHG-----EP-----VDQ
 --GLKFMGPPEPTYYKMSKDRVTGHQ-----EP-----LDR
 --GVTYMPGPPDAYYDMSHDRVKDHG-----EP-----IDR
 --EIQFLTTPPHSYEMIPERVPGVT-----ED-----ISR

LEELKILVDYD-----EKGYLLQIFTKPMQDRPTLFLEVI
 LEELKILVDYD-----EKGYLLQIFTKPVQDRPTLFLEVI
 IEKLSILVDFD-----EEGYLLQIFTKPLEDRPTLFIEII
 IEDLHILVDFD-----DRGYLLQIFTKPVEDRPTLFYIEII
 IQKLNILIDFD-----EGGYLLQIFTKPLMDRPTVFIEII
 IRSLDILIDFD-----EGGYLLQIFTKHLMRPTVFIEII
 LVRNNILIDYD-----EGGYLLQIFTRPLMDRPTVFIEII
 LRSNLNILIDYDPSTKPKPKSKRKRNMKCNLLQIFSKPLHDRPTLFIEII
 LRKLNILVDYDISTR-----NKKTGICNYLLQIFTKPLHDRPTLFIEII
 LQKRGILVDRD-----EDGYLLQIFTKPLVDRPTVFFELI
 LQKRGILVDRD-----EDGYLLQIFTKPIGDRPTVFFEII
 LKSRKILVDRD-----EDGYLLQIFTKPLGDRPTVFFEII
 LRELKILADR-----EDGYLLQIFTKPVQDRPTVFFEII
 LKKHGILIDRD-----EEGYLLQIFTKTIVDRPTMFFEVI
 LQELSILVDAD-----EEGYLLQIFTKPLQDRPTLFIEII
 LAALGILVDRD-----PDGYLLQIFTKPVEDRPTLFIEII
 IEKHNILVDRD-----ENGYLLQIFTKPVDRPTMFFEVI
 CEELGILVDR-----DDQGTLLQIFTKPLGDRPTIFIEII
 CEELGILVDR-----DDQGTLLQIFTKPLGDRPTIFIEII
 CEELGILVDR-----DAQGTLLQIFTKPVGDRPTIFVEII
 CEELGILVDR-----DDQGTLLQIFTKPVGDRPTIFIEII
 CEELGILVDR-----DDQGTLLQIFTKPIGDRPTIFIEII
 CEELGILVDR-----DDQGTLLQIFTKPVGDRPTIFIXII
 CEELGILVDR-----DDQGTLLQIFTKPVGDRPTIFIEII
 CEDLILVDR-----DDQGTLLQIFTKPVGDRPTIFIEII
 CEDLILVDR-----DDQGTLLQIFTKPVGDRPTIFIEII
 CEDLILVDR-----DDQGTLLQIFTKPVGDRPTLFIEII
 CEKLGILIDR-----DDQGTLLQIFTKPVGDRPTLFIEII
 CQELGVLVDR-----DDQGVLLQIFTKPVGDRPTLFLEMI
 CQELGVLVDR-----DDQGVLLQIFTKPVGDRPTLFLEMI
 CQELGVLVDR-----DDQGVLLQIFTKPVGDRPTFFLEMI
 CQELGVLVDR-----DDQGVLLQIFTKPVGDRPTFFLEMI
 CQELGVLVDR-----DDQGVLLQIFTKPVGDRPTLFLEMI
 CQELGVLVDR-----DDQGVLLQIFTKPVGDRPTLFLEMI
 CQELGVLVDR-----DDQGVLLQIFTKPVGDRPTLFLEMI
 VEELGLLVDR-----DDQGVLLQIFTKPVGDRPTLFLEMI
 VEELGLLVDR-----DDQGVLLQIFTKPVGDRPTLFLEMI
 VEELGLLVDR-----DDQGVLLQIFTKPLGDRPTVFIEII
 LVEQQILLDATLPASDG-----QDRPLLQFTFTQPLFGRPTFFFEVI
 LRDLRVLIDGAPTKDG-----ILLQIFFTQFTVIG--PVFFEII
 LQDLNILIDGAPLKD-----TLLQIFFTDTVIG--PVFFEII
 LQARGILLDSSVVGDK-----RLLQIFSETLMG--PVFFEFI
 MKKHGILIDGEGVVDGGE-----TRILLQIFSKTVIG--PIFFEFI
 MKKHGILIDGEGVVDGGE-----TRILLQIFSKTVIG--PIFFEFI
 LEKNAILVDGD--KTG-----KYLQIFTKNTFG--PIFYELI

RATTUS_NORVEGICUS
 HOMO-SAPIENS
 XENOPUS-LAEVIS
 BLEPHARISMA-JAPONICUS
 MYCOSPHAERELLA-GRAMINICOLA
 ASPERGILLUS-FUMIGATUS
 MAGNAPORTHE-GRISEA
 CANDIDA-ALBICANS
 PICHIA-STIPITIS
 RHODOCOCCUS-SP.
 RHODOCOCCUS-ERYTHROPOLIS
 JANIBACTER-SP.-HTCC2649

QRHN-----HQQFGAGNFNSLFKAFEEEEQALRGN
 QRHN-----HQQFGAGNFNSLFKAFEEQNLRGN
 QRHN-----HQQFGAGNFKAIFESTEMRQATIRGN
 QRHN-----NNGFGIGNFKALFESLEQEQRGN
 QRNN-----FDGFGAGNFKSLFEAIEREQDLRGN
 QRHN-----FSGFGAGNFKSLFEAIEREQALRGN
 QRNE-----FDGFGAGNFKSLFEAIEREQAERGN
 QRHH-----HNGFGKGTFKGLFESIIEQQKLRGT
 QRHH-----HNGFGKGTFKGLFETIEQQRIRGT
 ERHG-----SLGFGIGNFKALFEAIEREQAARGN
 ERHG-----SLGFGIGNFKALFEAIEREQAARGN
 ERHG-----SLGFGKGNFKALFESIEREQDARGN

STREPTOMYCES-AVERMITILIS ERHG-----SMGFGKGNFKALFEAIEREQEKRG
 KORDIA-ALGICIDA QRKG-----AQSGVGVGNFKALFEAIEREQAARGT
 LEEUWENHOEKIELLA-BLANDENSIS-ME QRMG-----AKGFGAGNFKALFESIEREQARGT
 GEMMATIMONAS-AURANTIACA QRKG-----ATSGFGKGNFRALFEAIEREQELRGN
 PICROPHILUS-TORRIDUS QRKG-----ARSGFGNGNFKALFEAIEREQAKRGN
 ARABIDOPSIS-THALIANA QRVGCMMKDEEGKA-----YQSGGCGGFGKGNFSELFKSIEEYKTLA
 BRASSICA-RAPA-SUBSP.-PEKINENSI QRIGCMKKDEEGRV-----YQSGGCGGFGKGNFSELFKSIEEYKTLA
 COPTIS-JAPONICA-VAR.-DISSECTA QRLGCMMLKDEEGKT-----YQKAGCGGFGKGNFSELFKSIEEYKTLA
 HEVEA-BRASILIENSIS QRVGCMIKDETGKE-----YQKGGCGGFGKGNFSELFKSIEEYKTLA
 MEDICAGO-TRUNCATULA QRVGCMMLKDEEGKE-----YQKGGCGGFGKGNFSELFKSIEEYKTLA
 GLYCINE-MAX QRIGCMVEDEEGKV-----YQKAGCGGFGKGNFSELFKSIEEYKTLA
 SOLANUM-LYCOPERSICUM QRIGCMMLKDEKGQV-----YQKGGCGGFGKGNFSELFRSIEEYKMLEA
 SOLANUM-TUBEROSUM QRIGCMMLKDENGQV-----YQKGGCGGFGKGNFSELFRSIEEYKMLEA
 NICOTIANA-BENTHAMIANA QRIGCMMLKDEKGQV-----YQKGGCGGFGKGNFSELFRSIEEYKTLA
 DAUCUS-CAROTA QRVGCMMLKDDAGQM-----YQKGGCGGFGKGNFSELFKSIEEYKTLA
 SOLENOSTEMON-SCUTELLARIOIDES QRVGCMMKDEEGKM-----YQKGGCGGFGKGNFSELFKSIEEYKMLEA
 HORDEUM-VULGARE QRIGCMKDEBERGEE-----YQKGGCGGFGKGNFSELFKSIEDYEKSLEA
 TRITICUM-AESTIVUM QRIGCMKDEBERGEE-----YQKGGCGGFGKGNFSELFKSIEDYEKSLEA
 AVENA-SATIVA QRIGCMKDEVEGQE-----YQKGGCGGFGKGNFSELFKSIEDYEKSLEA
 ORYZA-SATIVA QRIGCMKDESEGQE-----YQKGGCGGFGKGNFSELFKSIEEYKSLEA
 SORGHUM-BICOLOR QRIGCMKDEKGQE-----YQKGGCGGFGKGNFSELFKSIEDYEKSLEA
 ZEA-MAYS QRIGCMKDEKGQE-----YQKGGCGGFGKGNFSELFKSIEDYEKSLEA
 ABO95005_OLUCIMARINUS QRIGCMRKKADSES-----FEQAAGCGGFGKGNFSELFKSIEAYEATLQI
 OTAURI QRVGCMRKKADTDE-----LEQVAGCGGFGKGNFSELFKSIEAYEATLNI
 MICROMONAS-PUSILLA-CCMP1545 QRIGCLREVKSADANAPPRIEQAGGCGGFGKGNFSELFKSIEYERTLKM
 SYNECHOCOCCUS-SP. QRLG-----GATGFGGANFQALFEALERQQRQRHQ
 VIBRIO-SP.-MED222 QRKG-----NEGFGEGNFKALFESIEEDQIRRGV
 MARINOMONAS-SP.-MED121 QRKG-----NEGFGEGNFKALFESIEEDQIRRGV
 PSEUDOMONAS-FLUORESCENS QRKG-----DDGFGEGNFKALFESIERDQVRRGV
 SULFITOBACTER-SP.-NAS-14.1 ERKG-----DDGFGEGNFKALFESIEEQIDSGE
 OCEANICOLA-BATSENSIS-HTCC2597 QRKG-----DDGFGEGNFKALFESIEEQIDNGE
 BDELLOVIBRIO-BACTERIOVORUS QRKG-----HDGFGDGNFQALFDIAIERDQRRGY

RATTUS_NORVEGICUS LTDLETNGVRSKM-----
 HOMO-SAPIENS LTNMETNGVVPMAENLYFQ
 XENOPUS-LAEVIS L-----
 BLEPHARISMA-JAPONICUS LI-----
 MYCOSPHAERELLA-GRAMINICOLA L-----
 ASPERGILLUS-FUMIGATUS LV-----
 MAGNAPORTHE-GRISEA L-----
 CANDIDA-ALBICANS FVKSQNN-----
 PICHIA-STIPITIS LVQVDEDDDSQQST-----
 RHODOCOCCUS-SP. F-----
 RHODOCOCCUS-ERYTHROPOLIS F-----
 JANIBACTER-SP.-HTCC2649 L-----
 STREPTOMYCES-AVERMITILIS L-----
 KORDIA-ALGICIDA L-----
 LEEUWENHOEKIELLA-BLANDENSIS-ME L-----
 GEMMATIMONAS-AURANTIACA L-----
 PICROPHILUS-TORRIDUS L-----
 ARABIDOPSIS-THALIANA KQLVG-----
 BRASSICA-RAPA-SUBSP.-PEKINENSI KQLVG-----
 COPTIS-JAPONICA-VAR.-DISSECTA KANVVAA-----
 HEVEA-BRASILIENSIS KRNAEAR-----
 MEDICAGO-TRUNCATULA RRRA-----
 GLYCINE-MAX KRRA-----
 SOLANUM-LYCOPERSICUM KHVNQVAAVE-----
 SOLANUM-TUBEROSUM KHVNQVAAA-----
 NICOTIANA-BENTHAMIANA -----
 DAUCUS-CAROTA KQITGSAAA-----
 SOLENOSTEMON-SCUTELLARIOIDES KLVTKTAMA-----
 HORDEUM-VULGARE KQSAAVQGS-----
 TRITICUM-AESTIVUM KQSAAVQGS-----
 AVENA-SATIVA KQSVVAQKS-----

ORYZA-SATIVA	KQAPTVQGS----
SORGHUM-BICOLOR	KQAAAAQGS----
ZEA-MAYS	KQAAAAAAAQGS-
ABO95005_OLUCIMARINUS	-----
OTARI	-----
MICROMONAS-PUSILLA-CCMP1545	-----
SYNECHOCOCCUS-SP.	ALTP-----
VIBRIO-SP.-MED222	LDDA-----
MARINOMONAS-SP.-MED121	LNDA-----
PSEUDOMONAS-FLUORESCENS	LATD-----
SULFITOBACTER-SP.-NAS-14.1	LSEA-----
OCEANICOLA-BATSENSIS-HTCC2597	IAAE-----
BDELLOVIBRIO-BACTERIOVORUS	LT-----

La identidad de secuencia general entre secuencias individuales de HPPD de longitud total es, en general, bastante baja y se muestra para las proteínas de HPPD representativas en la Tabla 3. La Tabla 4a muestra la alineación de secuencias de la bolsa de unión. Por el contrario, la identidad de secuencia de los 36 aminoácidos que forman el sitio de unión es significativamente mayor que la que se muestra para las proteínas de HPPD representativas en la Tabla 4b. En particular, los aminoácidos en 8 posiciones se conservan de forma estricta en todas las especies e ilustran que estos aminoácidos tienen un papel clave (por ejemplo, His226, His308, Glu394 que se unen con el hierro requerido para la catálisis). Estas posiciones, con referencia a la HPPD de Arabidopsis (SEC ID N.º: 2) son His226, Ser267, Asn282, His308, Tyr342, Glu394, Gly420, Asn423 (Tabla 5a). Una mutación de cualquiera de los aminoácidos en cualquiera de estas posiciones llevará más probablemente a una proteína inactiva. La variabilidad en otras posiciones dentro del sitio de unión es más alta. La Tabla 5b muestra las 28 posiciones variables en el sitio de unión y los aminoácidos que se identificaron en estas posiciones usando la alineación de secuencias. Algunas posiciones tienen solo una variabilidad limitada que refleja su papel en el ambiente 3D. Un ejemplo de ello está representado por la posición 269. Todas las proteínas de HPPD tienen en esta posición ya sea Val, Ala o Thr. Observando las estructuras 3D, parece que en esta posición se requiere un aminoácido apolar pequeño y una mutación en un aminoácido polar como Arg, His o Lys perturbará la estructura proteica de forma local. Otro ejemplo está representado por la posición 379. La mayoría de las secuencias tienen una glutamina en la posición 379. Sin embargo, también hay algunas secuencias bacterianas que tienen una histidina en esta posición. Observando la estructura 3D, parece que solo algunos aminoácidos se toleran en esta posición. Gln379 en *A. thaliana* estabiliza a través de su donante de enlace de H la conformación de la cadena lateral de Glu394 estrictamente conservado que a su vez interactúa con el hierro catalítico. Además, con sus aceptores de enlace de H, Gln379 estabiliza la conformación de la cadena lateral de Asn423 estrictamente conservada que a su vez interactúa con la Tyr342 estrictamente conservada. Solo glutamina, asparagina e histidina tienen un donante de enlace H y un aceptor requerido para la estabilización de esta disposición 3D particular que muy probablemente desempeña un papel clave en la interacción de la hélice C-terminal con el núcleo de la proteína de HPPD. También se observó una variabilidad limitada en la posición 381 bien con una fenilalanina o bien con una tirosina en todas las secuencias de HPPD. El anillo aromático estabiliza la unión de los inhibidores de HPPD y muy probablemente también la unión del sustrato con la unión de HPPD. Sin embargo, la presencia del grupo hidroxilo adicional en la tirosina en comparación con fenilalanina no perturba la actividad catalítica. La tercera categoría de posiciones incluye aquellas posiciones que muestran una variabilidad natural muy alta. Estas posiciones pueden no ser cruciales para la unión con el sustrato y la catálisis pero influyen sobre la unión con el inhibidor. Estas posiciones incluyen posiciones adyacentes en la cadena 248 a 255 y posiciones en la hélice C-terminal 419-427. Se puede asumir que la interacción de esta cadena particular y la hélice C-terminal con el núcleo de la proteína desempeña un papel crucial en la unión con el inhibidor. La Tabla 5b incluye para cada posición variable aquellos aminoácidos que se han identificado en las alineaciones de secuencias usando todas las secuencias de HPPD conocidas.

Tabla 3: identidad de secuencia a pares de las secuencias de HPPD a partir del grupo representativo de proteínas de HPPD [no formando parte de la invención]

Tabla 4a: alineación de secuencias de aminoácidos que forman la bolsa de unión en el grupo representativo de secuencias de HPPD [no formando parte de la invención]

HOMO-SAPIENS	HVHFWSVL SIVVAPNQQHLYLQFFEF GAGNFNSLF
RATTUS_NORVEGICUS	HVHFWSVL SIVVAPNQQHLYLQFFEF GAGNFNSLF
XENOPUS-LAEVIS	HVHFWSVL SIVVTPNQQHLYLQFFEF GAGNFKALI
ASPERGILLUS-FUMIGATUS	HVHFWSVL SIVMAPNQQHLYLQFFEF GAGNFKSLI
MAGNAPORTHE-GRISEA	HVHFWSVL SIVMAPNQQHLYLQFFEF GAGNFKSLI
MYCOSPHAERELLA-GRAMINICOLA	HVHFWSVL SIVMSPNQQHLYLQFFEF GAGNFKSLI
CANDIDA-ALBICANS	HVHYWSVL SIVMSPNQQHNYLQFFEF GKGTFKGLI
PICHA-STIPITIS	HVHYWSALSIVMAPNQQHNYLQFFEF GKGTFKGLI
ARABIDOPSIS-THALIANA	HVHFAEFLSAVLAPNQQHMYLQFFEF GKGNFSELI
BRASSICA-RAPA-SUBSP.-PEKINENSI	HVHFAEFLSAVLAPNQQHMYLQFFEF GKGNFSELI
HEVEA-BRASILIENSIS	HVHFAEFLSLV LAPNQQHMYLQFFEF GKGNFSELI
COPTIS-JAPONICA-VAR.-DISSECTA	HVHFAEFLSIVLAPNQQHMYLQFFEF GKGNFSELI
SOLENOSTEMON-SCUTELLARIOIDES	HVHFAEFLSVLAPNQQHMYLQFFEF GKGNFSELI
AVENA-SATIVA	HVHFAEFLSVLAPNQQHMYLQFFEF GKGNFSELI
DAUCUS-CAROTA	HVHFAEFLSVLAPNQQHMYLQFFEF GKGNFSELI
SOLANUM-TUBEROSUM	HVHFAEFLSVLAPNQQHMYLQFFEF GKGNFSELI
SOLANUM-LYCOPERSICUM	HVHFAEFLSVLAPNQQHMYLQFFEF GKGNFSELI
MEDICAGO-TRUNCATULA	HVHFAEFLSVLAPNQQHMYLQFFEF GKGNFSELI
GLYCINE-MAX	HVHFAEFLSVLAPNQQHMYLQFFEF GKGNFSELI
HORDEUM-VULGARE	HVHFAEFLSVLAPNQQHLYLQFFEF GKGNFSELI
ORYZA-SATIVA	HVHFAEFLSVLAPNQQHLYLQFFEF GKGNFSELI
SORGHUM-BICOLOR	HVHFAEFLSMVLAPNQQHMYLQFFEF GKGNFSQLI
ZEA-MAYS	HVHFAEFLSMVLAPNQQHMYLQFFEF GKGNFSQLI
ABO95005_OLUCIMARINUS	HVHFAEFLSMVLAPNQQHQYLQFFEF GKGNFSELI
OTAURI	HVHFAEFLSMVLAPNQQHQYLQFFEF GKGNFSELI
MICROMONAS-PUSILLA-CCMP1545	HVHFAEFLSMVLAPNQQHQYLQFFEF GKGNFSELI
IRIICUM-AESTIVUM	HVHFAEFLSMVLAPNQQHLYLQFFEF GKGNFSELI
NICOIANA-BENITHAMIANA	HVHFAEFLSMVVAPNQQHMYLQFFEF GKGNFWELEI
RHODOCOCCUS-SP.	HVIMAEFLSKVVSPNQQHLYLQFFEF GIGNFKALI
RHODOCOCCUS-ERYTHROPOLIS	HVIMAEFLSKVVSPNQQHLYLQFFEF GLGNFKALI
STREPTOMYCES-AVERMILIS	HVIMKEFLSKVVAPNQQHLYLQFFEF GKGNFKALI
JANIBACTER-SP.-HTCC2649	HVVMAEFLSKVVAPNQQHLYLQFFEF GKGNFKALI
KORDIA-ALGICIDA	HVAIISFLSKVMSPNQQHLYLQFFEF GVGNFKALI
LEEUVENHOEKIELLA-BLANDENSIS-ME	HVVFLTFLSKVMSPNQQHLYLQFFEF GAGNFKALI
PICROPHILUS-TORRIDUS	HVELITFLSKVKPNQQHLYLQFFEF GNGNFKALI
GEMMATIMONAS-AURANTIACA	HVRLITFLSKVMAPNQQHLYLQFFEF GKGNFRAI
VIBRIO-SP.-MED222	HTRIRYFLSRAMTPNQQHMYLQFFEF GEGNFKALI
MARINOMONAS-SP.-MED121	HTRIRYFLSRAMTPNQQHMYLQFFEF GEGNFKALI
SULFITOBACTER-SP.-NAS-14.1	HTRIRFFLSRALTPNQQHMYLQFFEF GEGNFKALI
OCEANICOLA-BATSENSIS-HTCC2597	HIRIRFFLSRALTPNQQHMYLQFFEF GEGNFKALI
PSEUDOMONAS-FLUORESCENS	HTRARYFLSKAMSPNQQHMYLQFFEF GEGNFKALI
BDELLOVIBRIO-BACTERIOVORUS	HTYAKYFLSRAMRPNQQHLYLQFFEF GDGNFQALI
BLEPHARISMA-JAPONICUS	HVHYWSALSVVAPNQQHLYLQFFEF GIGNFKALI
SYNECHOCOCCUS-SP.	HVRLYRYLSVVGANQQHLYLQFFEF GEANFQALL

Tabla 4b: identidad de secuencia a pares de aminoácidos que forman el sitio de unión de HPPD en el grupo representativo [no formando parte de la invención]

100	100	HOMO-SAPIENS	100
100	100	RATTUS NORVIGICUS	100
88	88	XENOPUS-LAEVIS	88
89	89	ASPERGILLUS-FUMIGATUS	89
89	89	MAGNAPORTE-GRISEA	89
86	86	MYCOSPHAERELLA-GRAMMNICOLA	86
75	75	CANDIDA-ALBICANS	75
75	75	PICHA-STRIPIIS	75
72	72	ARABIDOPSIS-THALIANA	72
72	72	BRASSICA-PAPA-SUBSP.-PEKINENSIS	72
72	72	HEVEA-BRASILIENSIS	72
75	75	COPTIS-JAPONICA-VAR.-DISSECTA	75
72	72	SOLENOSTEMON-SCUTELLARIOIDES	72
72	72	AYENA-SATIVA	72
72	72	DALCIS-CAROTA	72
72	72	SOLANUM-TUBEROSUM	72
72	72	SOLANUM-LYCOPERSICUM	72
72	72	MEDICAGO-TRUNCATULA	72
88	88	GLYCINE-MAX	88
75	75	HORDEUM-VULGARE	75
72	72	ORYZA-SATIVA	72
72	72	SORGHUM-BICOLOR	72
72	72	ZEA-MAYS	72
72	72	ABO85005_OLLICIMARINUS	72
72	72	OTAUARI	72
72	72	MICROMONAS-PUSILLA-COMP1545	72
75	75	TRITICUM-AESTIVUM	75
75	75	NICOTIANA-BENTHAMIANA	75
69	69	RHODOCOCCLUS-SP.	69
69	69	RHODOCOCCLUS-ERYTHROPOLIS	69
72	72	STREPTOMYCES-AVERMITILIS	72
72	72	JANIBACTER-SP-HTCC2649	72
69	69	KORDIA-ALGIDA	69
72	72	LEUWENHOEKIELLA-BLANDENSIS-ME	72
69	69	PIOROPHILUS-TORRIDUS	69
69	69	GEMMATOMONAS-AURANTIACA	69
58	58	VIBRIO-SP.-MED222	58
58	58	MARINOMONAS-SP.-MEDI121	58
58	58	TRITICOBIACTER-SP.-NAS-14.1	58
58	58	OCEANICOLA-BATSSENSIS-HTCC2597	58
58	58	PSEUDIMONAS-FLUORESCENS	58
61	61	DELLOVERBIO-BACTERIOVORUS	61
81	81	BLEPHARISMA-JAPONICUS	81
64	64	SYNECHOCOCCUS-SP.	64

Tabla 5a: posiciones de aminoácidos que se conservan estrictamente, mostradas para las estructuras cristalinas

<i>Arabidopsis thaliana</i>		<i>Pseudomonas fluorescens</i>		<i>Streptomyces avermitilis</i>		<i>Homo sapiens</i>		<i>Rattus norvegicus</i>	
Posición	Amino-ácido	Posición	Amino-ácido	Posición	Amino-ácido	Posición	Amino-ácido	Posición	Amino-ácido
226	H	162	H	187	H	183	H	183	H
267	S	202	S	230	S	226	S	226	S
282	N	217	N	245	N	241	N	241	N
308	H	241	H	270	H	266	H	266	H
342	Y	271	Y	299	Y	295	Y	295	Y
394	E	323	E	349	E	349	E	349	E
420	G	334	G	360	G	360	G	360	G
423	N	337	N	363	N	363	N	363	N

Tabla 5b. posición de aminoácidos con los aminoácidos en estas posiciones en las estructuras cristalinas conocidas y los aminoácidos en estas posiciones dentro de todas las secuencias de HPPD

<i>Arabidopsis thaliana</i>		<i>Pseudomonas fluorescens</i>		<i>Streptomyces avermitilis</i>		<i>Homo sapiens</i>		<i>Rattus norvegicus</i>		Variabilidad natural	
Pos.	Aminoácido	Pos.	Aminoácido	Pos.	Aminoácido	Pos.	Aminoácido	Pos.	Aminoácido	Aminoácido	
228	V	164	T	189	V	185	V	185	V	VTCAG	
248	H	186	R	211	T	207	H	207	H	HRQTKKEYLAGSN	
250	F	188	A	213	M	209	F	209	F	FYILVAQEDGTSMRK	
251	A	189	R	214	K	210	W	210	W	AWILSRKHDEPGNY	
252	E	190	Y	215	E	211	S	211	S	ESTYFHQNGLMVILIR	
253	F	191	F	216	F	212	V	212	V	FVIALWMQHY	
265	L	200	L	228	L	224	L	224	L	LMVIA	
268	A	203	K	231	K	227	I	227	I	AVLMIKRQY	
269	V	204	A	232	V	228	V	228	V	VAT	
270	L	205	M	233	V	229	V	229	V	LVMIA	
271	A	206	S	234	A	230	A	230	A	ASTVRKELIMHG	
280	P	215	P	243	P	239	P	239	P	PAVINI	
293	Q	226	Q	255	Q	251	Q	251	Q	QLAVRSGVFM	
294	I	227	I	256	I	252	I	252	I	IMVIASP	
307	Q	240	Q	269	Q	265	Q	265	Q	QHN	
335	M	264	M	293	L	289	L	289	L	MLIN	
368	L	295	L	323	L	323	L	323	L	LM	
379	Q	310	Q	334	Q	334	Q	334	Q	QH	
381	F	312	F	336	F	336	F	336	F	FY	
392	F	321	F	347	F	347	F	347	F	FS	
419	F	333	F	359	F	359	F	359	F	FY	
421	K	335	E	361	K	361	A	361	A	KQAILVNDDEGS	
422	G	336	G	362	G	362	G	362	G	GAPVTM	
424	F	338	F	364	F	364	F	364	F	FAVIL	
425	S	339	K	365	K	365	N	365	N	SNKGRAPSIKQR	
426	E	340	A	366	A	366	S	366	S	EQSAIVFT	
427	L	341	L	367	L	367	L	367	L	LR	
431	I	345	I	371	I	371	F	371	F	IFVMLQR	

Tabla 5c: posición de aminoácidos con los aminoácidos en estas posiciones en las estructuras cristalinas conocidas y aminoácidos más comunes en estas posiciones dentro de todas las secuencias de HPPD

Arabidopsis thaliana		Pseudomonas fluorescens		Streptomyces avermitilis		Homo sapiens		Rattus norvegicus		Más común
Pos.	Aminoácido	Pos.	Aminoácido	Pos.	Aminoácido	Pos.	Aminoácido	Pos.	Aminoácido	Aminoácido
228	V	164	T	189	V	185	V	185	V	VTCA
248	H	186	R	211	T	207	H	207	H	HRQTKEYA
250	F	188	A	213	M	209	F	209	F	FYLVAEDTM
251	A	189	R	214	K	210	W	210	W	AWILRKHDEPY
252	E	190	Y	215	E	211	S	211	S	ESTRY
253	F	191	F	216	F	212	V	212	V	FVIALWY
265	L	200	L	228	L	224	L	224	L	LMVA
268	A	203	K	231	K	227	I	227	I	AVLMIKRQY
269	V	204	A	232	V	228	V	228	V	VA
270	L	205	M	233	V	229	V	229	V	LVMI
271	A	206	S	234	A	230	A	230	A	ASTVREGK
280	P	215	P	243	P	239	P	239	P	PAVI
293	Q	226	Q	255	Q	251	Q	251	Q	QLA
294	I	227	I	256	I	252	I	252	I	IMVA
307	Q	240	Q	269	Q	265	Q	265	Q	QHN
335	M	264	M	293	L	289	L	289	L	MLI
368	L	295	L	323	L	323	L	323	L	LM
379	Q	310	Q	334	Q	334	Q	334	Q	QH
381	F	312	F	336	F	336	F	336	F	FY
392	F	321	F	347	F	347	F	347	F	F
419	F	333	F	359	F	359	F	359	F	FY
421	K	335	E	361	K	361	A	361	A	KQAVLDE
422	G	336	G	362	G	362	G	362	G	GA
424	F	338	F	364	F	364	F	364	F	FAVL
425	S	339	K	365	K	365	N	365	N	SNKRAIQ
426	E	340	A	366	A	366	S	366	S	ESAIVF
427	L	341	L	367	L	367	L	367	L	LR
431	I	345	I	371	I	371	F	371	F	IFVMQRL

5 No todos los aminoácidos que se producen en las posiciones variables tienen la misma probabilidad de estar presentes en una proteína activa. En algunos casos la mayoría de las secuencias tienen los mismos aminoácidos en una posición particular mientras que otros aminoácidos están presentes en la posición en solo algunas pocas secuencias de HPPD. Un ejemplo es la posición 392. La mayoría de las secuencias tiene en una correspondiente posición una fenilalanina mientras que algunas, es decir, las secuencia de *Burkholderia* tienen una serina. Si bien, en algunos casos raros, los aminoácidos raros pueden ser el resultado de un error de secuenciación, en la mayoría de los otros casos, la proteína resultante es activa. La Tabla 5c muestra una lista con los aminoácidos más comunes en las posiciones variables.

10 La influencia de los aminoácidos en las posiciones variables es diferente. Algunas de estas posiciones son cruciales para la catálisis y/o la interacción de HPPD con un inhibidor mientras que otras pueden tener un menor impacto. Por ejemplo los cambios en las posiciones 269 y 280 que están en contacto directo con el inhibidor y el sustrato tendrán muy probablemente un gran impacto sobre la catálisis y la unión con el inhibidor. Además las modificaciones en las posiciones implicadas en el movimiento de la hélice inducido por el inhibidor o la unión con el sustrato tales como las posiciones 252, 421 y 422 probablemente tengan un alto impacto sobre la unión con el inhibidor. Por el contrario las modificaciones en las posiciones como 293 influyen menos probablemente en la unión con el inhibidor porque esta posición está bien alejada del sitio activo. La Tabla 6a muestra la posición con muy alto impacto y la Tabla 6b con alto impacto sobre la actividad y la unión con el inhibidor.

20 A partir de estas observaciones, los autores de la presente invención sacaron la conclusión de que los aminoácidos se prefieren en los sitios de unión que se producen de forma natural en las posiciones correspondientes. Esto significa, que esas modificaciones que intercambian un aminoácido que se da en la naturaleza en otro proporcionan probablemente una proteína de HPPD catalíticamente activa que puede ejercer una tolerancia modificada o incluso incrementada a los inhibidores de herbicidas de HPPD. Son aún más promisorias aquellas proteínas mutantes que tienen el aminoácido hallado con mayor frecuencia en la naturaleza en la posición variable seleccionada de la Tabla 5c.

25 Tabla 6a: posiciones de aminoácidos con alta prioridad mostrada para las estructuras de rayos X

<i>Arabidopsis thaliana</i>		<i>Pseudomonas fluorescens</i>		<i>Streptomyces avermitilis</i>		<i>Homo sapiens</i>		<i>Rattus norvegicus</i>	
Pos.	Amino-ácido	Pos.	Amino-ácido	Pos.	Amino-ácido	Pos.	Amino-ácido	Pos.	Amino-ácido
228	V	164	T	189	V	185	V	185	V
250	F	188	A	213	M	209	F	209	F
251	A	189	R	214	K	210	W	210	W
252	E	190	Y	215	E	211	S	211	S
253	F	191	F	216	F	212	V	212	V
265	L	200	L	228	L	224	L	224	L
268	A	203	K	231	K	227	I	227	I
269	V	204	A	232	V	228	V	228	V
270	L	205	M	233	V	229	V	229	V
271	A	206	S	234	A	230	A	230	A
280	P	215	P	243	P	239	P	239	P
307	Q	240	Q	269	Q	265	Q	265	Q
335	M	264	M	293	L	289	L	289	L
368	L	295	L	323	L	323	L	323	L
379	Q	310	Q	334	Q	334	Q	334	Q
392	F	321	F	347	F	347	F	347	F

(continuación)

<i>Arabidopsis thaliana</i>		<i>Pseudomonas fluorescens</i>		<i>Streptomyces avermitilis</i>		<i>Homo sapiens</i>		<i>Rattus norvegicus</i>	
Pos.	Amino-ácido	Pos.	Amino-ácido	Pos.	Amino-ácido	Pos.	Amino-ácido	Pos.	Amino-ácido
421	K	335	E	361	K	361	A	361	A
422	G	336	G	362	G	362	G	362	G
426	E	340	A	366	A	366	S	366	S
427	L	341	L	367	L	367	L	367	L

Tabla 6b: posiciones de aminoácidos con muy alta prioridad mostrada para las estructuras de rayos X

<i>Arabidopsis thaliana</i>		<i>Pseudomonas fluorescens</i>		<i>Streptomyces avermitilis</i>		<i>Homo sapiens</i>		<i>Rattus norvegicus</i>	
Pos.	Amino-ácido	Pos.	Amino-ácido	Pos.	Amino-ácido	Pos.	Amino-ácido	Pos.	Amino-ácido
252	E	190	Y	215	E	211	S	211	S
269	V	204	A	232	V	228	V	228	V
280	P	215	P	243	P	239	P	239	P
335	M	264	M	293	L	289	L	289	L
368	L	295	L	323	L	323	L	323	L
421	K	335	E	361	K	361	A	361	A
422	G	336	G	362	G	362	G	362	G

5 En otro aspecto, en el ácido nucleico aislado de la invención según se ha definido con anterioridad, dicho aminoácido está seleccionado de Glu, Thr, Arg, Ser o Tyr en una posición en una proteína de HPPD, correspondiendo dicha posición a la posición 252 de la secuencia de aminoácidos de la SEC ID N.º: 2. Diversas secuencias de proteínas HPPD o de proteínas HPPD pronosticadas se conocen en la técnica. Ellas incluyen las secuencias de HPPD de *Streptomyces avermitilis* (Genebank SAV11864), *Daucus carota* (Genebank DCU 87257),

10 *Arabidopsis thaliana* (Genebank AF047834), *Mycosphaerella graminicola* (Genebank AF038152), *Oryza sativa* / arroz [BAD26248], *Zea mays* / maíz [ACN36372], *Avena sativa* [ABZ23427], *Pseudomonas fluorescens* [ABF50055], *Synechococcus* sp. [YP_473959], *Blepharisma japonicum* [BAF91881], *Rhodococcus* RHA1 sp. ro0240 [YP_702005], *Rhodococcus* RHA1 sp. ro0341 [YP_703002], *Picrophilus torridus* [YP_024147], *Kordia algicida* [ZP_02161490], *Sorghum bicolor* [XP_002453359], *Triticum aestivum* / trigo [AAZ67144], o *Hordeum vulgare* / cebada [O48604].

15 La secuencia de la proteína de HPPD tomada como punto de partida puede ser cualquier secuencia de aminoácidos que codifica una proteína de HPPD catalíticamente activa. En una forma de realización del ácido nucleico aislado de la invención, dicha proteína de HPPD comprende la secuencia de aminoácidos de la SEC ID N.º: 4 [*Oryza sativa*], en la que la secuencia de aminoácidos resultante comprende al menos un aminoácido seleccionado de His, Ile, Leu, Asn o Ser en la posición 249 de la secuencia de aminoácidos de la SEC ID N.º: 4 (correspondiente a la posición 252

20 de la SEC ID N.º: 2).

En otra forma de realización del ácido nucleico aislado de la invención, dicha proteína de HPPD comprende la secuencia de aminoácidos de la SEC ID N.º: 6 [*Zea mays*], en la que la secuencia de aminoácidos resultante comprende al menos un aminoácido seleccionado de His, Ile, Leu, Asn o Ser en la posición 226 de la secuencia de aminoácidos de SEC ID N.º: 6 (correspondiente a la posición 252 de la SEC ID N.º: 2).

25 En otra forma de realización del ácido nucleico aislado de la invención, dicha proteína de HPPD comprende la secuencia de aminoácidos de la SEC ID N.º: 8 [*Avena sativa*], en la que la secuencia de aminoácidos resultante comprende al menos un aminoácido seleccionado de His, Ile, Leu, Asn o Ser en la posición 243 de la secuencia de aminoácidos de la SEC ID N.º: 8 (correspondiente a la posición 252 de la SEC ID N.º: 2).

En otra forma de realización del ácido nucleico aislado de la invención, dicha proteína de HPPD comprende la

secuencia de aminoácidos de la SEC ID N.º: 10 [*Pseudomonas fluorescens*], en la que la secuencia de aminoácidos resultante comprende al menos un aminoácido seleccionado de His, Ile, Leu, Asn o Ser en la posición 190 de la secuencia de aminoácidos de la SEC ID N.º: 10 (correspondiente a la posición 252 de la SEC ID N.º: 2), preferentemente Ser.

5 En otra forma de realización del ácido nucleico aislado de la invención, dicha proteína de HPPD comprende la secuencia de aminoácidos de la SEC ID N.º: 14 [*Synechococcus* sp.], en la que la secuencia de aminoácidos resultante comprende al menos un aminoácido seleccionado de His, Ile, Leu, Asn o Ser en la posición 171 de la secuencia de aminoácidos de la SEC ID N.º: 14 (correspondiente a la posición 252 de la SEC ID N.º: 2), preferentemente Ser.

10 En otra forma de realización del ácido nucleico aislado de la invención, dicha proteína de HPPD comprende la secuencia de aminoácidos de la SEC ID N.º: 16 [*Blepharisma japonicum*], en la que la secuencia de aminoácidos resultante comprende al menos un aminoácido seleccionado de His, Ile, Leu, Asn o Ser en la posición 211 de la secuencia de aminoácidos de la SEC ID N.º: 16 (correspondiente a la posición 252 de la SEC ID N.º: 2).

15 En otra forma de realización, del ácido nucleico aislado de la invención, dicha proteína de HPPD comprende la secuencia de aminoácidos de la SEC ID N.º: 18 [*Rhodococcus* RHA1 sp. ro0240], en la que la secuencia de aminoácidos resultante comprende al menos un aminoácido seleccionado de His, Ile, Leu, Asn o Ser en la posición 261 de la secuencia de aminoácidos de la SEC ID N.º: 18 (correspondiente a la posición 252 de la SEC ID N.º: 2), preferentemente Ser.

20 En otra forma de realización, del ácido nucleico aislado de la invención, dicha proteína de HPPD comprende la secuencia de aminoácidos de la SEC ID N.º: 30 [*Rhodococcus* RHA1 sp. 0341], en la que la secuencia de aminoácidos resultante comprende al menos un aminoácido seleccionado de His, Ile, Leu, Asn o Ser en la posición 233 de la secuencia de aminoácidos de la SEC ID N.º: 18 (correspondiente a la posición 252 de la SEC ID N.º: 2), preferentemente Ser.

25 En otra forma de realización, del ácido nucleico aislado de la invención, dicha proteína de HPPD comprende la secuencia de aminoácidos de la SEC ID N.º: 20 [*Picrophilus torridus*], en la que la secuencia de aminoácidos resultante comprende al menos un aminoácido seleccionado de His, Ile, Leu, Asn o Ser en la posición 20 de la secuencia de aminoácidos de la SEC ID N.º: 20 (correspondiente a la posición 252 de la SEC ID N.º: 2), preferentemente Ser.

30 En otra forma de realización, del ácido nucleico aislado de la invención, dicha proteína de HPPD comprende la secuencia de aminoácidos de la SEC ID N.º: 22 [*Kordia algicida*], en la que la secuencia de aminoácidos resultante comprende al menos un aminoácido seleccionado de His, Ile, Leu, Asn o Ser en la posición 22 de la secuencia de aminoácidos de la SEC ID N.º: 22 (correspondiente a la posición 252 de la SEC ID N.º: 2).

35 En otra forma de realización, del ácido nucleico aislado de la invención, dicha proteína de HPPD comprende la secuencia de aminoácidos de la SEC ID N.º: 24 [*Sorghum bicolor*], en la que la secuencia de aminoácidos resultante comprende al menos un aminoácido seleccionado de His, Ile, Leu, Asn o Ser en la posición 24 de la secuencia de aminoácidos de la SEC ID N.º: 24 (correspondiente a la posición 252 de la SEC ID N.º: 2).

40 En otra forma de realización, del ácido nucleico aislado de la invención, dicha proteína de HPPD comprende la secuencia de aminoácidos de la SEC ID N.º: 26 [*Triticum aestivum* / trigo], en la que la secuencia de aminoácidos resultante comprende al menos un aminoácido seleccionado de His, Ile, Leu, Asn o Ser en la posición 26 de la secuencia de aminoácidos de la SEC ID N.º: 26 (correspondiente a la posición 252 de la SEC ID N.º: 2).

En otra forma de realización, del ácido nucleico aislado de la invención, dicha proteína de HPPD comprende la secuencia de aminoácidos de la SEC ID N.º: 2 [*Arabidopsis thaliana*], en la que la secuencia de aminoácidos resultante comprende al menos un aminoácido seleccionado de His, Ile, Leu, Asn o Ser en la posición 252 de la secuencia de aminoácidos de la SEC ID N.º: 2.

45 En una forma de realización adicional del ácido nucleico de la invención, en dicha proteína de HPPD mutada se han reemplazado al menos dos aminoácidos.

Un ácido nucleico aislado que comprende una secuencia de nucleótidos que codifica una proteína de HPPD mutada, en la que dicha proteína de HPPD mutada tiene actividad de HPPD,

50 en la que en dicha proteína de HPPD mutada al menos un aminoácido en la posición 228, 248, 270, 271, 379 y/o 427 se reemplazó por otro aminoácido.

En una forma de realización alternativa del ácido nucleico de la invención que tiene al menos un aminoácido en la posición 228, 248, 270, 271, 379 y/o 427 suprimido o reemplazado por otro aminoácido según se define con anterioridad, dicha proteína de HPPD mutada comprende

a) una His en una posición en una proteína de HPPD, en la que dicha posición corresponde a la posición 226 de

la secuencia de aminoácidos de la SEC ID N.º: 2;

- b) una Ser en una posición en una proteína de HPPD, en la que dicha posición corresponde a la posición 267 que comprende la secuencia de aminoácidos de la SEC ID N.º: 2 o en una posición correspondiente en una enzima de HPPD diferente;
- 5 c) una Asn en una posición en una proteína de HPPD, en la que dicha posición corresponde a la posición 282 de la secuencia de aminoácidos de la SEC ID N.º: 2;
- d) una His en una posición en una proteína de HPPD, en la que dicha posición corresponde a la posición 308 de la secuencia de aminoácidos de la SEC ID N.º: 2;
- 10 e) una Tyr en una posición en una proteína de HPPD, en la que dicha posición corresponde a la posición 342 de la secuencia de aminoácidos de la SEC ID N.º: 2;
- f) un Glu en una posición en una proteína de HPPD, en la que dicha posición corresponde a la posición 394 de la secuencia de aminoácidos de la SEC ID N.º: 2;
- g) una Gly en una posición en una proteína de HPPD, en la que dicha posición corresponde a la posición 420 de la secuencia de aminoácidos de la SEC ID N.º: 2; y
- 15 h) una Asn en una posición en una proteína de HPPD, en la que dicha posición corresponde a la posición 423 de la secuencia de aminoácidos de la SEC ID N.º: 2.

En otra forma de realización del ácido nucleico aislado según se define con anterioridad, en dicha proteína de HPPD mutada, al menos un aminoácido se reemplazó de modo tal que la secuencia de aminoácidos resultante tenía al menos uno seleccionado de

- 20 a. Ala, Cys, Gly, Thr o Val en una posición en una proteína de HPPD, en la que dicha posición corresponde a la posición 228 de la secuencia de aminoácidos de la SEC ID N.º: 2;
- b. Ala, Glu, Gly, His, Lys, Leu, Asn, Gln, Arg, Ser, Thr o Tyr en una posición en una proteína de HPPD, en la que dicha posición corresponde a la posición 248 de la secuencia de aminoácidos de la SEC ID N.º: 2;
- 25 c. Ala, Ile, Leu, Met o Val en una posición en una proteína de HPPD, en la que dicha posición corresponde a la posición 270 de la secuencia de aminoácidos de la SEC ID N.º: 2;
- d. Ala, Glu, His, Ile, Lys, Leu, Met, Arg, Ser, Thr o Val en una posición en una proteína de HPPD, en la que dicha posición corresponde a la posición 271 de la secuencia de aminoácidos de la SEC ID N.º: 2;
- e. His o Gln en una posición en una proteína de HPPD, en la que dicha posición corresponde a la posición 379 de la secuencia de aminoácidos de la SEC ID N.º: 2; y
- 30 f. Leu o Arg en una posición en una proteína de HPPD, en la que dicha posición corresponde a la posición 427 de la secuencia de aminoácidos de la SEC ID N.º: 2.

En otra forma de realización del ácido nucleico aislado según se define con anterioridad, en dicha proteína de HPPD mutada, al menos un aminoácido se reemplazó de modo tal que la secuencia de aminoácidos resultante tenía al menos uno seleccionado de

- 35 a. Val o Thr en una posición en una proteína de HPPD, en la que dicha posición corresponde a la posición 228 de la secuencia de aminoácidos de la SEC ID N.º: 2;
- b. Leu, Met o Val en una posición en una proteína de HPPD, en la que dicha posición corresponde a la posición 270 de la secuencia de aminoácidos de la SEC ID N.º: 2;
- 40 c. Ala o Ser en una posición en una proteína de HPPD, en la que dicha posición corresponde a la posición 271 de la secuencia de aminoácidos de la SEC ID N.º: 2;
- d. Gln en una posición en una proteína de HPPD, en la que dicha posición corresponde a la posición 379 de la secuencia de aminoácidos de la SEC ID N.º: 2; y
- e. Leu en una posición en una proteína de HPPD, en la que dicha posición corresponde a la posición 427 de la secuencia de aminoácidos de la SEC ID N.º: 2.

- 45 En otra forma de realización del ácido nucleico según se define con anterioridad, dicha proteína de HPPD mutada es capaz de incrementar la tolerancia de una planta a al menos un herbicida que actúa sobre HPPD (también denominado herbicida inhibidor de HPPD).

En otra forma de realización, la presente invención se refiere a una proteína codificada por el ácido nucleico aislado

de la invención.

En otra forma de realización, la presente invención se refiere a un gen quimérico que comprende una secuencia codificadora que comprende el ácido nucleico de la invención operativamente ligado a un promotor que se puede expresar en plantas y opcionalmente una región de terminación de la transcripción y poliadenilación.

5 Como una secuencia de regulación que funciona como un promotor en células vegetales y plantas, se puede hacer uso de cualquier secuencia promotora de un gen que se expresa naturalmente en plantas, en particular un promotor que se expresa en especial en las hojas de plantas como, por ejemplo, promotores "constitutivos" de origen bacteriano, vírico o vegetal, o promotores "dependientes de la luz", tales como aquel de un gen de subunidad pequeña de ribulosa-bis-carboxilasa/oxigenasa (RuBisCO) de planta, o cualquier promotor expresable conocido
10 apropiado que se puede usar. Entre los promotores de origen vegetal, se pueden mencionar los promotores de histonas como se describen en el documento EP 0 507 698 A1, el promotor de actina de arroz (documento US 5.641.876), o un promotor de ubiquitina vegetal (documento US 5.510.474). Entre los promotores de un gen viral de planta, se mencionará aquel del virus mosaico de la coliflor (CaMV 19S o 35S, Sanders y col. (1987), Nucleic Acids Res. 15(4):1543-58), el circovirus (AU 689 311) o el virus mosaico de la vena Cassava (CsVMV, documento US 7.053.205).

En una forma de realización de esta invención, una secuencia promotora específica de regiones o tejidos vegetales particulares se puede usar para expresar las proteínas de HPPD de la invención, como promotores específicos de semillas (Datla, R. y col., 1997, Biotechnology Ann. Rev. 3, 269-296), en especial el promotor de nabo (documento EP 255 378 A1), el promotor de faseolina, el promotor de glutenina, el promotor de heliantina (documento WO 92/17580), el promotor de albúmina (documento WO 98/45460), el promotor de oleosina (documento WO 98/45461), el promotor SAT1 o el promotor de SAT3 (documento PCT/US98/06978).

También se puede usar un promotor inducible elegido ventajosamente de los promotores de fenilalanina amoníaco liasa (PAL), HMG-CoA reductasa (HMG), quitinasa, glucanasa, inhibidor de proteinasa (PI), gen de la familia PR1, nopalina sintasa (nos) y vspB (documento US 5 670 349, Tabla 3), el promotor HMG2 (documento US 5 670 349), el
25 promotor de beta-galactosidasa (ABG1) de manzana y el promotor de aminociclopropan-carboxilato (ACC sintasa) de manzana (documento WO 98/45445).

De acuerdo con la invención, se pueden usar también, en combinación con el promotor, otras secuencias de regulación que se ubican entre el promotor y la secuencia codificadora, como activadores de la transcripción ("potenciadores"), por ejemplo, el activador de la traducción del virus del mosaico del tabaco (TMV) descritos en la
30 solicitud WO 87/07644, o del virus del grabado de tabaco (TEV) descrito por Carrington y Freed 1990, J. Virol. 64: 1590-1597, por ejemplo, o intrones tales como el intrón adh1 de maíz o el intrón 1 de la actina del arroz.

Como un terminador de regulación o secuencia de poliadenilación, se puede usar cualquier secuencia correspondiente de origen bacteriano tal como por ejemplo el terminador nos de *Agrobacterium tumefaciens*, de origen viral tal como por ejemplo, el terminador CaMV 35S, o de origen vegetal tal como por ejemplo un terminador de histonas como se describió en la solicitud de patente publicada EP 0 633 317 A1.

Un procedimiento de obtención de una proteína de HPPD mutada capaz de modular la tolerancia de una planta a al menos un herbicida que actúa sobre HPPD, en el que dicha proteína de HPPD mutada tiene actividad de HPPD, que comprende

40 i. proporcionar una proteína de HPPD, comprendiendo opcionalmente dicha HPPD una secuencia de aminoácidos, en la que

a) una His está presente en una posición en una proteína de HPPD, correspondiendo dicha posición a la posición 226 de la secuencia de aminoácidos de la SEC ID N.º 2;

b) una Ser está presente en una posición en una proteína de HPPD, correspondiendo dicha posición a la posición 267 de la secuencia de aminoácidos de la SEC ID N.º 2;

45 c) una Asn está presente en una posición en una proteína de HPPD, correspondiendo dicha posición a la posición 282 de la secuencia de aminoácidos de la SEC ID N.º 2;

d) una His está presente en una posición en una proteína de HPPD, correspondiendo dicha posición a la posición 308 de la secuencia de aminoácidos de la SEC ID N.º 2;

50 e) una Tyr está presente en una posición en una proteína de HPPD, correspondiendo dicha posición a la posición 342 de la secuencia de aminoácidos de la SEC ID N.º 2;

f) un Glu está presente en una posición en una proteína de HPPD, correspondiendo dicha posición a la posición 394 de la secuencia de aminoácidos de la SEC ID N.º 2;

g) una Gly está presente en una posición en una proteína de HPPD, correspondiendo dicha posición a la posición 420 de la secuencia de aminoácidos de la SEC ID N.º 2; y

h) un Asn está presente en una posición en una proteína de HPPD, correspondiendo dicha posición a la posición 423 de la secuencia de aminoácidos de la SEC ID N.º: 2

ii) sustituir un aminoácido en dicha enzima de HPPD de modo tal que la secuencia de aminoácidos resultante tenga al menos uno seleccionado de His, Ile, Leu, Asn o Ser en una posición en una proteína de HPPD, correspondiendo dicha posición a la posición 252 de la secuencia de aminoácidos de la SEC ID N.º: 2;

a. y una deleción o sustitución aminoacídica de al menos una posición en una proteína de HPPD, correspondiendo dicha posición a al menos una de las posiciones 228, 248, 270, 271, 379 y 427 de la secuencia de aminoácidos de la SEC ID N.º: 2;

iii) determinar la inhibición de la proteína de HPPD resultante por al menos un herbicida que actúa sobre HPPD;

en la que una inhibición de la proteína resultante de menos o más que la observada con una proteína de HPPD de referencia es indicativa de que la proteína resultante es capaz de modular la tolerancia de una planta a dicho herbicida.

Se ha de entender que también los aminoácidos (más específicos) y posiciones enumerados con anterioridad para otras formas de realización, tales como el ácido nucleico de la invención, se pueden aplicar al procedimiento de obtención de una proteína de HPPD mutada tal como se describió con anterioridad.

En una forma de realización alternativa del procedimiento de obtención de una proteína de HPPD mutada tal como se describió con anterioridad, dicha proteína de HPPD mutada es capaz de incrementar la tolerancia de una planta a al menos un herbicida que actúa sobre HPPD.

Dentro del procedimiento de obtención anterior de una proteína de HPPD mutada, se pueden seleccionar diferentes herbicidas que actúan sobre la HPPD. Conforme a ello, en otra forma de realización del procedimiento de obtención de una proteína de HPPD mutada tal como se describió con anterioridad, en donde dicha proteína de HPPD mutada es capaz de incrementar la tolerancia de una planta a al menos un herbicida que actúa sobre HPPD, el herbicida que actúa sobre HPPD está seleccionado de tricetonas, o pirazolinatos, con preferencia, tembotriona, mesotriona, topamezona o sulcotriona, biciclopirona, pirasulfotol, pirazolato, benzofenap y tefulitriona, en particular tembotriona y tales plantas que contienen la HPPD de la invención tienen una tolerancia agrónomicamente aceptable a un herbicida inhibidor de HPPD en particular a tricetonas, o pirazolinatos, con preferencia, tembotriona, mesotriona, topamezona o sulcotriona, biciclopirona, pirasulfotol, pirazolato, benzofenap y tefulitriona, en particular tembotriona.

En otra forma de realización, la presente invención se refiere a un procedimiento de producción de una planta transgénica que comprende introducir en dicho genoma vegetal el ácido nucleico de la presente invención operativamente ligado a un promotor expresable en plantas, el gen quimérico de la invención o un ácido nucleico que codifica la enzima de HPPD identificada por el procedimiento de la reivindicación 19 o 20.

En una forma de realización alternativa del procedimiento de producción de una planta transgénica tal como se describe con anterioridad, el ácido nucleico de la invención, en el que dicha proteína de HPPD mutada es capaz de incrementar la tolerancia de una planta a al menos un herbicida que actúa sobre HPPD, o un ácido nucleico identificado por el procedimiento de obtención de una proteína de HPPD mutada, en el que dicha proteína de HPPD mutada es capaz de incrementar la tolerancia de una planta a al menos un herbicida que actúa sobre HPPD, ambos operativamente ligados a un promotor expresable en plantas, o el gen quimérico de la invención que comprende un ácido nucleico, en el que dicha proteína de HPPD mutada es capaz de incrementar la tolerancia de una planta a al menos un herbicida que actúa sobre la HPPD, se introduce en dicha planta.

En otra forma de realización, la presente invención se refiere a una célula vegetal que comprende el ácido nucleico aislado de la invención o el gen quimérico de la invención en su información genética.

La presente invención también se refiere a una planta, una parte de una planta o tejido vegetal que consiste esencialmente en las células vegetales de la invención.

Además, la presente invención se refiere a una planta que se puede obtener a partir del procedimiento de obtención de una proteína de HPPD mutada capaz de modular o aumentar la tolerancia de una planta a al menos un herbicida que actúa sobre HPPD en todos los aspectos alternativos descritos con anterioridad.

La planta de la presente invención puede ser cualquier planta. Los ejemplos no limitativos de plantas de la invención incluyen trigo, algodón, canola, arroz, maíz, soja, sorgo, canola, girasol, tabaco, remolacha, algodón, maíz, trigo, cebada, arroz, sorgo, tomate, mango, melocotón, manzana, pera, fresa, plátano, melón, patata, zanahoria, lechuga, repollo, cebolla, Soya spp., caña de azúcar, guisante, judía, álamo, uva, cítricos, alfalfa, centeno, avena, césped y gramíneas forrajeras, lino y colza oleaginosa y plantas productoras de nueces. La presente invención también se refiere a una semilla de la planta de la invención.

En otra forma de realización, la presente invención se refiere a un procedimiento de modulación de la tolerancia de

una planta a al menos un herbicida que actúa sobre HPPD que comprende introducir el ácido nucleico aislado de la invención operativamente ligado a un promotor expresable en plantas o el gen quimérico de la invención en el genoma de una planta.

5 En una forma de realización alternativa, la presente invención se refiere a un procedimiento para aumentar la tolerancia de una planta a al menos un herbicida que actúa sobre HPPD u obtener una planta tolerante a un herbicida inhibidor de HPPD que comprende introducir el ácido nucleico aislado de la invención, en la que dicho ácido nucleico codifica una proteína de HPPD mutada que es capaz de incrementar la tolerancia de una planta a al menos un herbicida que actúa sobre HPPD, operativamente ligado a un promotor expresable en plantas o el gen quimérico de la invención que comprende un ácido nucleico de la invención, en donde dicho ácido nucleico codifica una proteína de HPPD mutada que es capaz de incrementar la tolerancia de una planta a al menos un herbicida que actúa sobre HPPD, en el genoma de una planta.

10 Además, la presente invención se refiere a un procedimiento de control de malas hierbas que comprende pulverizar al menos un herbicida que actúa sobre HPPD en o alrededor de una planta de cultivo, en la que dicha planta de cultivo comprende el ácido nucleico de la presente invención, en la que dicha proteína de HPPD mutada es capaz de incrementar la tolerancia de una planta a al menos un herbicida que actúa sobre HPPD, operativamente ligado a un promotor expresable en plantas o el gen quimérico de la invención que comprende el ácido nucleico de la invención, en la que dicha proteína de HPPD mutada es capaz de incrementar la tolerancia de una planta a al menos un herbicida que actúa sobre la HPPD. En una forma de realización alternativa del procedimiento de control de malas hierbas, se aumenta la tolerancia de dicha planta a al menos un herbicida que actúa sobre la HPPD.

15 Además, la presente invención se refiere al uso de un gen quimérico de la invención o el ácido nucleico de la invención operativamente ligado a un promotor expresable en plantas para modular la tolerancia de una planta a al menos un herbicida que actúa sobre la HPPD.

20 En una forma de realización alternativa, la presente invención se refiere al uso de un gen quimérico de la invención o el ácido nucleico de la invención operativamente ligado a un promotor expresable en plantas para incrementar la tolerancia de una planta a al menos un herbicida que actúa sobre la HPPD. En esta forma de realización de la invención, el gen quimérico usado comprende el ácido nucleico de la invención, en la que la proteína de HPPD mutada codificada es capaz de incrementar la tolerancia de una planta a al menos un herbicida que actúa sobre HPPD. De modo alternativo, si se usa un ácido nucleico operativamente ligado a un promotor expresable en plantas, se selecciona dicho ácido nucleico de modo tal que la proteína de HPPD mutada codificada sea capaz de incrementar la tolerancia de una planta a al menos un herbicida que actúa sobre HPPD.

25 La presente invención también se refiere a la célula vegetal de la invención y la planta de la invención que puede comprender otro rasgo útil tal como se describe más adelante.

30 Si bien una cantidad de plantas de cultivos tolerantes a herbicidas son asequibles en comercios en la actualidad, un tema que surgió para muchos herbicidas comerciales y combinaciones de herbicida/cultivo es que los herbicidas individuales tienen un espectro incompleto de actividad contra especies de malas hierbas comunes. Para la mayoría de los herbicidas individuales que se usaron durante algún tiempo, poblaciones de especies de malas hierbas resistentes a herbicidas y biotipos se han vuelto más prevalentes (ver, por ejemplo, Tranel y Wright (2002) *Weed Science* 50: 700-712; Owen y Zelaya (2005) *Pest Manag. Sci.* 61: 301-311). Se han descrito plantas transgénicas que son resistentes a más de un herbicida (ver, por ejemplo, documento WO 2005/012515). Sin embargo, hay una demanda continua de mejoras en todo aspecto de la producción de cultivos, opciones de control de malas hierbas, extensión del control de malas hierbas residual y mejora en el rendimiento de cultivos.

35 La proteína o gen de HPPD de la invención se combina ventajosamente en plantas con otros genes que codifican proteínas o ARN que confieren propiedades agronómicas útiles a estas plantas. Entre los genes que codifican proteínas o ARN que confieren propiedades agronómicas útiles en plantas transformadas, se pueden mencionar secuencias de ADN que codifican proteínas que confieren tolerancia a uno o varios herbicidas que, de acuerdo con su estructura química, difieren de los herbicidas inhibidores de HPPD y otros que confieren tolerancia a ciertos insectos, aquellos que confieren tolerancia a ciertas enfermedades, ADN que codifican ARN que proporcionan un control de nematodos o insectos, etc.

40 Estos genes se describen en particular en las solicitudes de patentes PCT publicadas WO 91/02071 y WO 95/06128.

45 Entre las secuencias de ADN que codifican proteínas que confieren tolerancia a ciertos herbicidas en células vegetales transformadas y plantas, se pueden mencionar un gen bar o PAT o el gen de *Streptomyces coelicolor* descrito en el documento WO 2009/152359 que confiere tolerancia a herbicidas de glufosinato, un gen que codifica un EPSPS apropiado que confiere tolerancia a herbicidas que tienen EPSPS como un blanco, como glifosato y sus sales (documentos US 4.535.060, US 4.769.061, US 5.094.945, US 4.940.835, US 5.188.642, US 4.971.908, US 5.145.783, US 5.310.667, US 5.312.910, US 5.627.061, US 5.633.435), o un gen que codifica glifosato oxidoreductasa (documento US 5.463.175).

50 Entre las secuencias de ADN que codifican un EPSPS apropiado que confiere tolerancia a los herbicidas que tienen EPSPS como un blanco, se mencionará en particular el gen que codifica un EPSPS de planta, en particular EPSPS

de maíz, en particular un EPSPS de maíz que comprende dos mutaciones, en particular una mutación en la posición de aminoácido 102 y una mutación en la posición de aminoácido 106 (documento WO 2004/074443) y que se describe en la solicitud de patente US 6566587, en adelante mencionado EPSPS de maíz doble mutante o 2mEPSPS, o el gen que codifica un EPSPS aislado de *Agrobacterium* y que se describe por la secuencia ID N.º: 2 y la secuencia ID N.º: 3 de la patente US 5.633.435, también denominado CP4.

Entre las secuencias de ADN que codifican un EPSPS apropiado que confieren tolerancia a los herbicidas que tienen EPSPS como un blanco, se mencionará más en particular el gen que codifica un EPSPS GRG23 de *Arthrobacter globiformis*, pero también los mutantes GRG23 ACE1, GRG23 ACE2, o GRG23 ACE3, en particular los mutantes o variantes de GRG23 tal como se describen en el documento WO 2008/100353, como GRG23(ace3)R173K de la SEC ID N.º: 29 en el documento WO 2008/100353.

En el caso de las secuencias de ADN que codifican EPSPS y más en particular que codifican los genes anteriores, la secuencia que codifica estas enzimas está precedida ventajosamente por una secuencia que codifica un péptido de tránsito, en particular el "péptido de tránsito optimizado" descrito en la patente US 5.510.471 o 5.633.448.

En el documento WO 2007/024782, se divulgan plantas tolerantes a glifosato y al menos un inhibidor de ALS (acetolactato sintasa). Más específicamente se divulgan plantas que contienen genes que codifican un polipéptido de GAT (glifosato-N-acetiltransferasa) y un polipéptido que confiere resistencia a inhibidores de ALS.

En el documento US 6855533, se divulgaron plantas de tabaco transgénicas con genes ALS/AHAS de *Arabidopsis* mutados.

En el documento US 6.153.401, se divulgan plantas que contienen genes que codifican 2,4-D-monooxigenasas que confieren tolerancia a 2,4-D (ácido 2,4-diclorofenoxiacético) por metabolización.

En el documento US 2008/0119361 y el documento US 2008/0120739, se divulgan plantas que contienen genes que codifican dicamba monooxigenasas que confieren tolerancia a dicamba (ácido 3,6-dicloro-2-metoxibenzoico) por metabolización.

Todos los rasgos de tolerancia a herbicidas antes mencionados se pueden combinar con aquellos que tienen tolerancia a HPPD que son materia objeto de esta invención.

Entre las secuencias de ADN que codifican proteínas que se refieren a las propiedades de tolerancia a insectos, se mencionarán más en particular las proteínas Bt ampliamente descritas en la bibliografía y bien conocidas por los expertos en la técnica. También se hará mención de las proteínas extraídas de bacterias tales como *Photorhabdus* (documento WO 97/17432 y documento WO 98/08932).

Entre tales secuencias de ADN que codifican proteínas de interés que se refieren a propiedades de tolerancia a insectos, se hará mención más en particular a las proteínas Bt Cry o VIP descritas ampliamente en la bibliografía y bien conocidas por aquellos expertos en la técnica. Éstas incluyen la proteína Cry1F o híbridos derivados de una proteína Cry1F (por ejemplo, las proteínas híbridas Cry1A-Cry1F descritas en los documentos US 6.326.169; US 6.281.016; US 6.218.188, o sus fragmentos tóxicos), las proteínas de tipo Cry1A o sus fragmentos tóxicos, con preferencia la proteína Cry1Ac o híbridos derivados de la proteína Cry1Ac (por ejemplo, la proteína híbrida Cry1Ab-Cry1Ac descrita en el documento US 5.880.275) o la proteína Cry1Ab o Bt2 o sus fragmentos insecticidas tal como se describen en el documento EP451878, las proteínas Cry2Ae, Cry2Af o Cry2Ag tal como se describen en el documento WO 02/057664 o sus fragmentos tóxicos, la proteína Cry1A.105 descrita en el documento WO 2007/140256 (SEC ID N.º: 7) o uno de sus fragmentos tóxicos, la proteína VIP3Aa19 de N.º de acceso de NCBI ABG20428, la proteína VIP3Aa20 de N.º de acceso de NCBI ABG20429 (SEC ID N.º: 2 en el documento WO 2007/142840), las proteínas VIP3A producidas en los eventos de algodón COT202 o COT203 (documentos WO 2005/054479 y WO 2005/054480, respectivamente), las proteínas Cry tal como se describen en el documento WO 01/47952, la proteína VIP3Aa o uno de sus fragmentos tóxicos tal como se describe en Estruch y col. (1996), Proc Natl Acad Sci U S A. 28; 93(11):5389-94 y US 6.291.156, las proteínas insecticidas de *Xenorhabdus* (tal como se describe en el documento WO 98/50427), *Serratia* (en particular de *S. entomophila*) o cepas de especies de *Photorhabdus*, como proteínas Tc de *Photorhabdus* tal como se describen en el documento WO 98/08932 (por ejemplo, Waterfield y col., 2001, Appl Environ Microbiol. 67(11): 5017-24; French-Constant y Bowen, 2000, Cell Mol Life Sci.; 57(5): 828-33). También se incluyen en el presente documento cualquier variante o mutante de cualquiera de estas proteínas que difieren en algunos (1-10, con preferencia, 1-5) aminoácidos de cualquiera de las secuencias anteriores, en particular la secuencia de su fragmento tóxico, o que se fusionan con un péptido de tránsito, como un péptido de tránsito de plástido u otra proteína o péptido.

Las FIGURAS muestran:

FIGURA 1: Superposición de la estructura de rayos X de HPPD de *Arabidopsis thaliana* (gris oscuro) y (a) *Pseudomonas fluorescens* (gris claro), (b) *Streptomyces avermitilis*, (c) *Homo sapiens*, (d) *Rattus norvegicus*. Las estructuras se muestran como gráfico de cintas.

FIGURA 2: Aminoácidos que forman el sitio de unión de (a) *Arabidopsis thaliana*, (b) *Pseudomonas*

fluorescens (c) *Streptomyces avermitilis* (d) *Homo sapiens* y (e) *Rattus norvegicus* (representado por líneas en negrita). El núcleo de la proteína de HPPD se muestra como traza de C_{alfa} y el hierro está marcado.

FIGURA 3: Color de ensayo marrón en *Escherichia coli*

Listado de secuencias

- 5 SEC ID N.º: 1: secuencia de ácidos nucleicos que codifica HPPD de *Arabidopsis thaliana*
 SEC ID N.º: 2: proteína codificada por SEC ID N.º: 1
 SEC ID N.º: 3: secuencia de ácidos nucleicos que codifica HPPD de *Oryza sativa*
 SEC ID N.º: 4: proteína codificada por SEC ID N.º: 3
 SEC ID N.º: 5: secuencia de ácidos nucleicos que codifica HPPD de *Zea mays*
- 10 SEC ID N.º: 6: proteína codificada por SEC ID N.º: 5
 SEC ID N.º: 7: secuencia de ácidos nucleicos que codifica HPPD de *Avena sativa*
 SEC ID N.º: 8: proteína codificada por SEC ID N.º: 7
 SEC ID N.º: 9: secuencia de ácidos nucleicos que codifica HPPD de *Pseudomonas fluorescens*
 SEC ID N.º: 10: proteína codificada por SEC ID N.º: 9
- 15 SEC ID N.º: 13: secuencia de ácidos nucleicos que codifica HPPD de *Synechococcus* sp.
 SEC ID N.º: 14: proteína codificada por SEC ID N.º: 13
 SEC ID N.º: 15: secuencia de ácidos nucleicos que codifica HPPD de *Blepharsima japonicum*
 SEC ID N.º: 16: proteína codificada por SEC ID N.º: 15
 SEC ID N.º: 17: secuencia de ácidos nucleicos que codifica HPPD ro0341 aislada de *Rhodococcus* RHA1
- 20 SEC ID N.º: 18: proteína codificada por SEC ID N.º: 17
 SEC ID N.º: 19: secuencia de ácidos nucleicos que codifica HPPD de *Picrophilus torridus*
 SEC ID N.º: 20: proteína codificada por SEC ID N.º: 19
 SEC ID N.º: 21: secuencia de ácidos nucleicos que codifica HPPD de *Kordia algicida*
 SEC ID N.º: 22: proteína codificada por SEC ID N.º: 21
- 25 SEC ID N.º: 23: secuencia de ácidos nucleicos que codifica HPPD de *Sorghum bicolor*
 SEC ID N.º: 24: proteína codificada por SEC ID N.º: 23
 SEC ID N.º: 25: secuencia de ácidos nucleicos que codifica HPPD de *Triticum aestivum*
 SEC ID N.º: 26: proteína codificada por SEC ID N.º: 25
 SEC ID N.º: 27: secuencia de ácidos nucleicos que codifica HPPD de *Hordeum vulgare*
- 30 SEC ID N.º: 28: proteína codificada por SEC ID N.º: 27
 SEC ID N.º: 29: secuencia de ácidos nucleicos que codifica HPPD ro0240 aislado de *Rhodococcus* RHA1
 SEC ID N.º: 30: proteína codificada por SEC ID N.º: 29
 SEC ID N.º: 31: secuencia de ácidos nucleicos que codifica polipéptido de HPPD de tipo silvestre de *Arabidopsis thaliana*, que también contiene en el extremo 5' un ácido nucleico que codifica una alanina y 6 aminoácidos de histidina
- 35 SEC ID N.º: 32: proteína codificada por SEC ID N.º: 31
 SEC ID N.º: 33: secuencia de ácidos nucleicos que codifica polipéptido de HPPD mutante, que también contiene en el extremo 5' un ácido nucleico que codifica una alanina y 6 aminoácidos de histidina

SEC ID N.º: 34: proteína codificada por SEC ID N.º: 33,

SEC ID N.º: 35: secuencia de ácidos nucleicos que codifica un polipéptido de HPPD mutante, que también contiene en el extremo 5' un ácido nucleico que codifica una alanina y 6 aminoácidos de histidina

SEC ID N.º: 36: proteína codificada por SEC ID N.º: 35,

5 SEC ID N.º: 37: secuencia de ácidos nucleicos que codifica un polipéptido de HPPD mutante, que también contiene en el extremo 5' un ácido nucleico que codifica una alanina y 6 aminoácidos de histidina

SEC ID N.º: 38: proteína codificada por SEC ID N.º: 37,

SEC ID N.º: 39: secuencia de ácidos nucleicos que codifica un polipéptido de HPPD mutante, que también contiene en el extremo 5' un ácido nucleico que codifica una alanina y 6 aminoácidos de histidina

10 SEC ID N.º: 40: proteína codificada por SEC ID N.º: 39,

SEC ID N.º: 41: secuencia de ácidos nucleicos que codifica un polipéptido de HPPD mutante, que también contiene en el extremo 5' un ácido nucleico que codifica una alanina y 6 aminoácidos de histidina

SEC ID N.º: 42: proteína codificada por SEC ID N.º: 41,

15 SEC ID N.º: 43: secuencia de ácidos nucleicos que codifica un polipéptido de HPPD mutante, que también contiene en el extremo 5' un ácido nucleico que codifica una alanina y 6 aminoácidos de histidina

SEC ID N.º: 44: proteína codificada por SEC ID N.º: 43,

SEC ID N.º: 45: secuencia de ácidos nucleicos que codifica un polipéptido de HPPD mutante, que también contiene en el extremo 5' un ácido nucleico que codifica una alanina y 6 aminoácidos de histidina

SEC ID N.º: 46: proteína codificada por SEC ID N.º: 45

20 SEC ID N.º: 47: secuencia de ácidos nucleicos que codifica un polipéptido de HPPD mutante, que también contiene en el extremo 5' un ácido nucleico que codifica una alanina y 6 aminoácidos de histidina

SEC ID N.º: 48: proteína codificada por SEC ID N.º: 47

SEC ID N.º: 49: secuencia de ácidos nucleicos que codifica un polipéptido de HPPD mutante, que también contiene en el extremo 5' un ácido nucleico que codifica una alanina y 6 aminoácidos de histidina

25 SEC ID N.º: 50: proteína codificada por SEC ID N.º: 49.

Los ejemplos ilustran la invención.

Ejemplo 1:

30 Una mutación de aminoácidos estrictamente conservados lleva a una proteína inactiva. Los mutantes puntuales S267A (SEC ID N.º: 34) en *A. thaliana* y N282A (SEC ID N.º: 36) en *A. thaliana* son inactivos. Las posiciones de aminoácido dadas se refieren a la posición de la SEC ID N.º: 2.

Los polipéptidos de HPPD mutantes de la presente invención tienen cambios de aminoácidos en una o más posiciones respecto de la secuencia inicial de tipo silvestre de la que se derivan.

35 La secuencia de ADN SEC ID N.º: 2 que codifica la proteína de HPPD de tipo silvestre de *Arabidopsis thaliana* (1335 pares de bases; Genebank AF047834; documento WO 96/38567) se clonó en el vector pSE420(RI)NX (modificado del vector de clonación y expresión pSE420(RI)NX (5261 pares de bases) se basa en el plásmido pSE420 por Invitrogen (Karlsruhe, Alemania)). En el extremo 5', se insertó directamente cadena abajo de ATG una secuencia de ácidos nucleicos que codifica un aminoácido de alanina y una secuencia de ácidos nucleicos que codifica un HIS6-Tag N-terminal (6x HIS, codificado por: cat cat cat cac cat cat). La secuencia resultante se presenta como SEC ID N.º: 31. Corriente arriba de ATG, se añadieron dos pares de bases de cisteína adicionales a fin de obtener una
40 secuencia correspondiente al sitio de reconocimiento de la enzima de restricción NcoI y cadena abajo del codón de detención se añadieron las secuencias correspondientes al sitio de reconocimiento de la enzima de restricción XbaI. El plásmido resultante se usó luego para transformar células de *E. coli* BL21 (DE3) con 50 µg/ml de kanamicina o 100 µg/ml de selección de carbenicilina tal como se describe en las solicitudes de patentes europeas EP09015984.9; EP09015985.6; EP 09015986.4, EP 09015987.2 y EP09015988.0, todas presentadas el 23 de diciembre de 2009.

45 Ensayo rápido de la actividad de la proteína de HPPD: producción de color marrón

Ensayo de control colorimétrico para enzimas activas de HPPD

Un medio de cultivo de tipo caldo de YT con el 1 % de agarosa, 5 mM de L-tirosina y 42 mM de succinato, que contiene el agente de selección para el vector pSE420 se vierte en placas de pocillos profundos. El cultivo de *E. coli* en la fase de crecimiento exponencial que contiene el vector pSE420-HPPDx (cualquier gen que codifica una enzima / proteína de HPPD teórica) se aplica a cada pocillo. Después de 16 horas a 37 °C, los pocillos que no contienen el medio de cultivo, aquellos que habían sido sembrados con un cultivo de *E. coli* con el vector vacío pSE420, son transparentes, o aquellos que habían sido sembrados con un cultivo de *E. coli* con un vector pSE420 que contiene un gen que codifica una HPPD inactiva, son transparentes, mientras que los pocillos sembrados con un cultivo de *E. coli* con el vector pSE420-HPPD que codifica la HPPD activa son marrones. Previamente se había demostrado que este ensayo refleja la actividad de HPPD, no importa cuál sea el origen de esta actividad y permite la identificación de actividades de la HPPD (documento US 6.768.044). Como puede verse en la Figura 3, el cultivo de bacterias que contiene el gen que codifica la HPPD mutante no desarrolló un color marrón mientras que uno que contiene el gen que codifica la HPPD de tipo silvestre desarrolló un color marrón fuerte que reflejaba la actividad de las enzimas de HPPD. Se puede concluir que los dos mutantes de HPPD no son capaces de convertir HPP en homogentisato. Los dos mutantes de HPPD son inactivos. Se puede concluir que las posiciones de aminoácido 267 y 282 (referidas a la posición en la SEC ID N.º: 2) son esenciales para la actividad de HPPD.

La expresión de la proteína de HPPD se realizó de la siguiente manera.

Se usaron cultivos cultivados durante la noche a 37 °C para inocular el medio LB en una relación 1:100. Se dejaron crecer las células hasta que la DO alcanzara 0,5, luego se inició la expresión a partir del promotor *trp-lac* (*trc*) por inducción con 1 mM de IPTG que se une con el represor *lac* y causa su disociación del operón *lac*. La expresión se llevó a cabo durante 15 h a 28 °C.

Para preparar el cultivo de prearranque, se inocularon 2 ml de medio TB (100 µg*ml⁻¹ de carbenicilina) con 50 µl de una solución madre de K-12 BL21 glicerol de *E. coli*. El cultivo de prearranque se incubó a 37 °C con agitación a 140 rpm durante 15 h. Se usaron 200 µl del cultivo de prearranque para iniciar el cultivo de inicio (5 ml de suplemento de TB con 100 µg*l⁻¹), que se incubó durante 3 h a 37 °C.

Para preparar el cultivo principal, se inocularon 400 ml de medio TB (100 µg*ml⁻¹ de carbenicilina) con 4 ml del cultivo de inicio. Este cultivo de inicio se incubó a 37 °C con agitación a 140 rpm hasta que se alcanzó la DO₆₀₀ de 0,5. Luego se indujo la expresión de la proteína recombinante con 400 µl de 1 M de solución IPTG. Las células se dejaron en cultivo durante una hora adicional en estas condiciones, luego la temperatura se redujo a 28 °C y el cultivo se agitó a 140 rpm durante 15 h. Las células se cultivaron por centrifugación a 6000 x g durante 15 min a 4 °C. Luego se almacenaron los sedimentos celulares a -80 °C.

Aislamiento y purificación de His₆-AtHPPD en forma nativa

Lisis de células

Se lisaron células usando lisozima, una enzima que escinde las ligaciones 1,4-β entre el ácido N-acetilmurámico y residuos de N-acetil-D-glucosamina en peptidoglicano que forma la pared celular bacteriana. Las membranas celulares se desbarataron luego por la presión interna de la célula bacteriana. Además, el tampón de lisis contenía Benzonase[®] Nucleasa, una endonucleasa que hidroliza todas las formas de ADN y ARN sin dañar las proteínas y así reduce ampliamente la viscosidad del lisado celular. La lisis en condiciones nativas se llevó a cabo en hielo.

Para la purificación de proteínas rotuladas con His₆, se usó el kit QIAexpress[®] Ni-NTA Fast Start Kit según la instrucción del manual del usuario.

40 Purificación de proteínas rotuladas con His₆ por cromatografía por afinidad de ion metálico inmovilizado (IMAC)

El lisado celular clarificado (10 ml) obtenido después de la centrifugación de la reacción de lisis se cargó en una columna Ni-NTA Fast Start Column con el kit QIAexpress[®] Ni-NTA Fast Start Kit (Qiagen, Hilden, Alemania) y la purificación se llevó a cabo de acuerdo con el manual de instrucciones. La proteína rotulada con His₆ se eluyó con 2,5 ml de tampón de elución.

Desalinización de soluciones de HPPD por filtración en gel

Se aplicaron soluciones de HPPD eluidas de una columna Ni-NTA Fast Start con 2,5 ml de tampón de elución a una columna Sephadex G-25 PD-10 (GE Healthcare, Friburgo, Alemania) de acuerdo con las instrucciones del manual del usuario. Después de que toda la muestra entrara en el lecho de gel, se realizó la elución con 3,5 ml de tampón de almacenamiento.

Las soluciones de HPPD eluidas de la columna de desalinización se congelaron a -80 °C en alícuotas de 1 ml.

Determinación de la concentración de proteína de HPPD usando el ensayo de proteína de Bradford

La concentración de proteína se determinó usando el ensayo de Bradford estándar (Bradford, (1976), Anal Biochem 72: 248-254).

Determinación de la pureza de soluciones de HPPD usando SDS-PAGE

La integridad de la proteína eluida se controló por medio de electroforesis en gel de proteína SDS-PAGE usando NuPAGE® Novex 4-12 % Bis-Tris Gels (Invitrogen, Karlsruhe, Alemania), se cargaron aproximadamente 10 µg de proteína. Se añadieron 10 µl de tampón de muestra de Laemmli a 1-10 µl de solución de proteína y la mezcla se incubó a 90 °C durante 10 min. Después de una corta etapa de centrifugación, toda la mezcla se cargó en una ranura de un gel SDS previamente fijado en una cámara de gel XCell SureLock™ Novex Mini-Cell rellena con tampón de corrida NuPAGE® MOPS SDS (diluido de la solución 20 x con ddH₂O). Luego se aplicó un voltaje de 150 a la cámara de gel durante 1 h. Para la tinción de las bandas de proteína, el gel se sumergió en solución de tinción azul brillante de Coomassie R-250. Para desteñir el gel de poliacrilamida, se sumergió en solución para desteñir azul brillante de Coomassie R-250 hasta que las bandas de proteína aparecieron azules en un gel blanco.

La actividad de HPPD se controló por medio de un ensayo espectrofotométrico estándar (procedimiento descrito extensivamente en el documento WO 2009/144079).

En este contexto, el valor de pI_{50} implica el valor de log de la concentración de inhibidor necesaria para inhibir el 50 % de la actividad enzimática en concentración molar.

Los valores de pI_{50} para inhibidores de HPPD se determinaron a partir de los gráficos de dosis-respuesta de la actividad de HPPD frente a la concentración de inhibidor usando el ensayo extensivamente descrito en el documento WO 2009/144079 a 2 mM de concentración de HPP fijo y 3 minutos de tiempo de incubación fijo usando el paquete informático ID Business Solutions Ltd. XLfit.

Tabla 7: Determinación de los pI_{50} de las enzimas de HPPD (HPPD de tipo silvestre de *Arabidopsis thaliana* "SEC ID N.º: 32", los mutantes S267A (SEC ID N.º: 34) y N282A (SEC ID N.º: 36)" y la tolerancia a los diversos inhibidores de HPPD enumerados más abajo tembotriona, dicetonitrilo, mesotriona, biciclopirona, pirasulfotol, sulcotriona, pirazolato, tefuriltriona y benzofenap. El símbolo ">" significa que el valor era mucho mayor que el indicado, pero no se podía calcular de forma precisa dentro del intervalo de concentración del inhibidor ensayado ($2,5 \times 10^{-6}$, $5,0 \times 10^{-6}$, $1,0 \times 10^{-5}$, $2,5 \times 10^{-5}$, $6,3 \times 10^{-5}$ y $2,5 \times 10^{-4}$ M).

	Biciclopirona	Benzofenap	Dicetonitrilo	Mesotriona
WT (SEC ID N.º: 32)	5,2	> 5,6	> 5,6	> 5,6
S267A (SEC ID N.º: 34)	nd - la (P)	nd - la (P)	nd - la (P)	nd - la (P)
N282A (SEC ID N.º: 36)	nd - la (P)	nd - la (P)	nd - la (P)	nd - la (P)

	Pirasulfotol	Pirazolato	Sulcotriona	Tefuriltriona	Tembotriona
WT (SEC ID N.º: 32)	5,4	5,4	> 5,6	> 5,6	> 5,6
S267A (SEC ID N.º: 34)	nd - la (P)	nd - la (P)	nd - la (P)	nd - la (P)	nd - la (P)
N282A (SEC ID N.º: 36)	nd - la (P)	nd - la (P)	nd - la (P)	nd - la (P)	nd - la (P)

Tal como se demostró previamente, los 2 mutantes de proteínas no son capaces de convertir HPP en homogentisato, lo que confirma que las 2 proteínas mutantes son inactivas. Esto confirma la hipótesis de que la posición 267 y 282 (haciendo referencia a la posición SEC ID N.º: 2) son absolutamente esenciales para obtener una HPPD activa.

Ejemplo 2: Mutantes de único punto revelaban mayor tolerancia a herbicidas inhibidores de HPPD:**Caracterización cinética y evaluación de tolerancia a inhibidores de HPPD de la enzima de HPPD "SEC ID N.º: 32".**

La actividad de HPPD se controló por medio del ensayo espectrofotométrico estándar (procedimiento extensivamente descrito en el documento WO 2009/144079).

Determinación de las propiedades cinéticas in vitro de HPPD

Los valores de K_m , $V_{m\max}$ y k_{cat} para diferentes preparaciones enzimáticas de HPPD y K_i , $K_1 = K_{on}$ y $K_{-1} = K_{off}$ para diferentes inhibidores de HPPD se determinaron usando un ensayo de HPLC para las mediciones de la actividad de HPPD. Las mezclas de ensayo contenían en un volumen de 1 ml 150 mM de tampón de Tris-HCl a pH 7,8, 10 mM de ascorbato de sodio, 650 unidades de catalasa bovina (Sigma C30 (Sigma-Aldrich, Múnich, Alemania), 34 mg de proteína/ml, 23.000 unidades/mg) y cantidades apropiadas de HPP, enzima de HPPD purificada e inhibidores de

HPPD. Para la determinación del valor de K_m , $V_{m\acute{a}x}$ y k_{cat} , las concentraciones de HPP en la mezcla de ensayo se variaron entre 10 y 400 μM . Para la determinación del valor K_i , $K_1 = K_{on}$ y $K_{-1} = K_{off}$, se usaron 2 mM de HPP. Todos los ensayos se iniciaron por adición de la enzima de HPPD a la mezcla de ensayo y se detuvieron en una serie de tiempos de entre 0 y 240 s por adición de 200 μl de la mezcla de reacción a los tubos de ensayo de reacción que contienen 20 μl de ácido perclórico al 10 %. La proteína precipitada se sedimentó por una centrifugación de 5 minutos a 10.000 g. 100 μl del sobrenadante se cargaron en una columna 250 x 4 mm Knauer (Berlín, Alemania) Eurospher 100-5 C18 equilibrada con 10 % de metanol, 0,1 % de ácido trifluoroacético (tampón A). La columna se eluyó, también a 1,5 ml/min, usando un lavado de 4 minutos con tampón A, seguido de un lavado de 3 min con 95 % de metanol y con un lavado adicional de 2 minutos con tampón A. La elución de HGA (ácido homogentísico) y HPP (hidroxifenilpiruvato) se controló a 292 nm. HGA se eluye a aproximadamente 5 minutos y HPP se eluye más tarde. Un grupo estándar de concentraciones de HGA se usaron para proporcionar una curva estándar a fin de calibrar la absorbancia de 292 nm del pico de HGA frente a la concentración de HGA.

Para las determinaciones del valor K_m y $V_{m\acute{a}x}$ las velocidades iniciales de la reacción de HPPD a diferentes concentraciones de sustrato se determinaron a partir de gráficos de HGA formados frente al tiempo y se ajustaron a la ecuación de Michaelis-Menten para enzimas no reaccionadas usando el paquete informático XLfit de ID Business Solutions Ltd. (www.idbs.com). Para la determinación de los valores K_i , $K_1 = K_{on}$ y $K_{-1} = K_{off}$ se ajustaron los transcurso de tiempo de la reacción de HPPD a diferentes concentraciones de inhibidor a las ecuaciones para el Mecanismo A, inhibición competitiva, para inhibidores de ligación estrecha (Cha, S. (1975) Tight-binding inhibitors – I. Kinetic behaviour. Biochemical Pharmacology 24, 2177-2185) usando el paquete informático XLfit de ID Business Solutions Ltd.

Tabla 8

SEC ID N.º: 32 de proteína consiste en la SEC ID N.º: 2 de proteína con una inserción de un aminoácido Ala y seis aminoácidos His directamente detrás de la primera metionina.

Caracterización cinética de enzimas de HPPD (“SEC ID N.º: 32” de Arabidopsis thaliana y la respectiva tolerancia de los inhibidores de la HPPD tembotriona y dicetonitrilo). En la siguiente tabla dada a continuación, “ K_m ” (constante de Michaelis-Menten) significa el parámetro cinético que se usa para caracterizar una enzima y se define como la concentración de sustrato que permite la velocidad máxima media de la reacción. K_m también se define como la concentración de sustrato a la que la velocidad de reacción alcanza la mitad de su valor máximo ($V_{m\acute{a}x}/2$) donde $V_{m\acute{a}x}$ tiene el significado de ser la velocidad máxima de la reacción. $K_{on} = K_1$ es igual a la constante de velocidad de asociación de la unión de enzima-sustrato y $K_{off} = K_{-1}$ es igual a la constante de velocidad de la disociación del complejo de enzima-inhibidor. K_i define la constante de inhibición.

		HPP		Tembotriona			Dicetonitrilo		
SEC ID N.º: 32	ID	K_m (μM)	$V_{m\acute{a}x}$ (μM)	k_1 ($\text{M}^{-1} \text{s}^{-1}$)	k_{-1} (s^{-1})	K_i (μM)	k_1 ($\text{M}^{-1} \text{s}^{-1}$)	k_{-1} (s^{-1})	K_i (μM)
		6,3	1,2	2,3E+05	3,5E-03	0,015	6,1E+05	1,1E-02	0,018

Los parámetros cinéticos K_m y $V_{m\acute{a}x}$ de la planta HPPD “SEC ID N.º: 32” resumidos en la Tabla 8 están en el intervalo de las actividades específicas medidas de la HPPD de planta purificada como por ejemplo para la HPPD de zanahoria (García y col. 2000, Biochemistry, 39, 7501-7507). La HPPD de Arabidopsis thaliana también es sensible en el intervalo similar a la HPPD de zanahoria a dicetonitrilo. Es la primera vez que se midió el parámetro cinético de la HPPD de Arabidopsis thaliana en presencia de tembotriona y está en un intervalo comparable al medido en presencia del inhibidor de HPPD dicetonitrilo.

Determinación de la actividad de HPPD en presencia de varios inhibidores de HPPD

En este contexto, el valor de pI_{50} significa el valor de log de la concentración de inhibidor necesaria para inhibir el 50 % de la actividad enzimática en la concentración molar.

Los valores de pI_{50} para los inhibidores de HPPD se determinaron a partir de gráficos de dosis-respuesta de la actividad de HPPD frente a la concentración de inhibidor usando el ensayo extensivamente descrito en el documento WO 2009/144079 a 2 mM de concentración de HPP fija y 3 minutos de tiempo de incubación fijo usando el paquete informático XLfit de ID Business Solutions Ltd.

Tabla 9: Determinación de los pI_{50} de las enzimas de HPPD (Arabidopsis thaliana “SEC ID N.º: 32” y mutantes de único punto de HPPD de Arabidopsis “SEC ID N.º: 35”) y sus respectivas tolerancias a los diversos inhibidores de HPPD enumerados más adelante tembotriona, dicetonitrilo, mesotriona, biclopirona, pirasulfotol, sulcotriona, pirazolato, tefuriltriona y benzofenap. El símbolo “>” significa que el valor era mucho mayor que el indicado pero no se podía calcular de forma precisa dentro del intervalo de concentración del inhibidor ensayado ($2,5 \times 10^{-6}$, $5,0 \times 10^{-6}$,

1,0 x 10⁻⁵, 2,5 x 10⁻⁵, 6,3 x 10⁻⁵ y 2,5 x 10⁻⁴ M).

	Biciclopirona	Benzofenap	Dicetonitrilo	Mesotriona	Pirasulfotol
SEC ID N.º: 32	5,2	> 5,6	> 5,6	> 5,6	5,4
E252S (SEC ID N.º: 36)	4,3	5,2	4,9	5,3	4,3

	Pirazolato	Sulcotriona	Tefuriltriona	Tembotriona
SEC ID N.º: 32	5,4	> 5,6	> 5,6	> 5,6
E252S (SEC ID N.º: 35)	4,7	5,1	5,2	5,5

5 Las mediciones muestran que el pI50 en el caso de mutantes puntuales se reduce en comparación con la proteína de tipo silvestre.

10 En la Tabla 9, se puede ver claramente que el mutante de HPPD está afectado en su tolerancia a al menos un inhibidor de HPPD y algo a varios inhibidores de HPPD. El mutante ensayado es más activo (muestra una tolerancia incrementada referida a los diversos inhibidores de HPPD aplicados) en presencia de al menos uno de los inhibidores de HPPD seleccionados, en consecuencia el complejo de enzima-inhibidor mutante debería ser menos estable que el complejo de enzima-inhibidor de tipo silvestre. Se puede concluir entonces que la posición 252 (tomando referencia en SEC ID N.º: 2) es importante en la modulación de la tolerancia de la enzima de HPPD a inhibidores de HPPD.

Listado de secuencias

15 <110> Bayer CropScience AG

<120> Variantes de HPPD y procedimientos de uso

<130> BCS 10-1029EP01

20 <140> EP10190629

<141> 10 de noviembre de 2010

<160> 50

25 <170> PatentIn version 3.3

<210> 1

<211> 1335

<212> ADN

30 <213> Arabidopsis thaliana

<400> 1

```

atggggccacc aaaacgccgc cgtttcagag aatcaaaacc atgatgacgg cgctgogtgc 60
tcgocgggat tcaagctcgt cggattttcc aagttcgtaa gaaagaatcc aaagtctgat 120
aaattcaagg ttaagcgctt ccatcacatc gagttctggt ggggcgacgc aaccaacgtc 180
gctcgtcget tctcctgggg tctgggggatg agattctccg ccaaatccga tttttccacc 240
ggaacatgg ttcacgctc ttacctact acctccggtg acctccgatt ctttttcaact 300
gctccttact ctccgtctct ctccgcgga gagattaaac cgacaaccac agctttotatc 360
ccaagtttcg atcacggctc ttgtcgttcc ttottctctt cacatggtct cggtggttaga 420
    
```

ES 2 594 284 T3

```

gccgttgcca ttgaagtaga agacgcagag tcagctttct ccatcagtgt agctaattggc 480
gctattccctt cgtcgccctcc tatcgtcctc aatgaagcag ttacgatcgc tgagggttaa 540
ctatacggcg atgttggtct cccgatatgtt agttacaaag cagaagatac cgaaaaatcc 600
gaattcttgc caggggttcga gcgtgtagag gatgcgctgt cgttccatt ggattatggt 660
atccggcgcc ttgaccacgc cgtgggaaac gttcctgagc ttggtcggc ttttaacttat 720
gtagcgggggt tcaactggtt tcaccaattc gcagagttca cagcagacga cgttggaacc 780
gccgagagcg gtttaaattc agcggctctg gctagcaatg atgaaatggt tcttctaccg 840
attaacgagc cagtgcacgg aacaaagagg aagagtcaga ttcagacgta tttggaacat 900
aacgaaggcg cagggctaca acatctggct ctgatgagtg aagacatatt caggaccctg 960
agagagatga ggaagaggag cagtattgga ggattcgcact tcatgccttc tctccgcct 1020
acttactacc agaatctcaa gaaacgggtc ggcgacgtgc tcagcgatga tcagatcaag 1080
gagtgtgagg aattagggat tctttagtag acagatgatc aagggacggt gcttcaaate 1140
ttcacaacac cactaggtga caggccgacg atatttatag agataatcca gagagtagga 1200
tgcattgatga aagatgagga aggggaaggct taccagagtg gaggatgtgg tggttttggc 1260
aaaggcaatt tctctgagct cttcaagtcc attgaagaat acgaaaagac tcttgaagcc 1320
aacagttag tggga 1335

```

<210> 2

<211> 445

5 <212> PRT

<213> *Arabidopsis thaliana*

<400> 2

```

Met Gly His Gln Asn Ala Ala Val Ser Glu Asn Gln Asn His Asp Asp
1           5           10           15

Gly Ala Ala Ser Ser Pro Gly Phe Lys Leu Val Gly Phe Ser Lys Phe
          20           25           30

Val Arg Lys Asn Pro Lys Ser Asp Lys Phe Lys Val Lys Arg Phe His
          35           40           45

His Ile Glu Phe Trp Cys Gly Asp Ala Thr Asn Val Ala Arg Arg Phe
          50           55           60

Ser Trp Gly Leu Gly Met Arg Phe Ser Ala Lys Ser Asp Leu Ser Thr
65           70           75           80

Gly Asn Met Val His Ala Ser Tyr Leu Leu Thr Ser Gly Asp Leu Arg
          85           90           95

Phe Leu Phe Thr Ala Pro Tyr Ser Pro Ser Leu Ser Ala Gly Glu Ile
          100          105          110

Lys Pro Thr Thr Thr Ala Ser Ile Pro Ser Phe Asp His Gly Ser Cys
          115          120          125

Arg Ser Phe Phe Ser Ser His Gly Leu Gly Val Arg Ala Val Ala Ile
          130          135          140

Glu Val Glu Asp Ala Glu Ser Ala Phe Ser Ile Ser Val Ala Asn Gly
145          150          155          160

```

ES 2 594 284 T3

Ala Ile Pro Ser Ser Pro Pro Ile Val Leu Asn Glu Ala Val Thr Ile
165 170 175

Ala Glu Val Lys Leu Tyr Gly Asp Val Val Leu Arg Tyr Val Ser Tyr
180 185 190

Lys Ala Glu Asp Thr Glu Lys Ser Glu Phe Leu Pro Gly Phe Glu Arg
195 200 205

Val Glu Asp Ala Ser Ser Phe Pro Leu Asp Tyr Gly Ile Arg Arg Leu
210 215 220

Asp His Ala Val Gly Asn Val Pro Glu Leu Gly Pro Ala Leu Thr Tyr
225 230 235 240

Val Ala Gly Phe Thr Gly Phe His Gln Phe Ala Glu Phe Thr Ala Asp
245 250 255

Asp Val Gly Thr Ala Glu Ser Gly Leu Asn Ser Ala Val Leu Ala Ser
260 265 270

Asn Asp Glu Met Val Leu Leu Pro Ile Asn Glu Pro Val His Gly Thr
275 280 285

Lys Arg Lys Ser Gln Ile Gln Thr Tyr Leu Glu His Asn Glu Gly Ala
290 295 300

Gly Leu Gln His Leu Ala Leu Met Ser Glu Asp Ile Phe Arg Thr Leu
305 310 315 320

Arg Glu Met Arg Lys Arg Ser Ser Ile Gly Gly Phe Asp Phe Met Pro
325 330 335

Ser Pro Pro Pro Thr Tyr Tyr Gln Asn Leu Lys Lys Arg Val Gly Asp
340 345 350

Val Leu Ser Asp Asp Gln Ile Lys Glu Cys Glu Glu Leu Gly Ile Leu
355 360 365

Val Asp Arg Asp Asp Gln Gly Thr Leu Leu Gln Ile Phe Thr Lys Pro
370 375 380

Leu Gly Asp Arg Pro Thr Ile Phe Ile Glu Ile Ile Gln Arg Val Gly
385 390 395 400

Cys Met Met Lys Asp Glu Glu Gly Lys Ala Tyr Gln Ser Gly Gly Cys
405 410 415

Gly Gly Phe Gly Lys Gly Asn Phe Ser Glu Leu Phe Lys Ser Ile Glu
420 425 430

ES 2 594 284 T3

Glu Tyr Glu Lys Thr Leu Glu Ala Lys Gln Leu Val Gly
 435 440 445

<210> 3
 <211> 1341
 5 <212> ADN
 <213> *Oryza sativa*

<400> 3

```

atgcctccca ctcccacccc caccgccacc accggcgccg tctcggccgc tgcggcgccg      60
ggggagaaacg cgggggttccg cctcgtcggg caccgccgct tcgtccgcgc caaccgcggg      120
agcgaccggt tccaggcgct cgcgttccac cacgtcgcgc tctgggtgcgc cgacgcccgg      180
tccgcgcggg gccgggttcgc cttcgccttg ggcgcgcgc tcgccgccag gtccgacctc      240
tccacgggga actccgcgca cgctccctc ctctccgct ccgcctcctg cgcgttctctc      300
ttcaccgccc cctacggcgg cgaccacggc gtcggcgccg accgggccac caccgcctcc      360
atcccttctt tctcccacgg cgcgcgcggg aggttcgcgc cggaccacgg cctcgcgggtg      420
cacgcggtgg cgtcgcgcgt cgcgcgcgc gccgcgcct tccgcgccag cgtcgcggcc      480
ggtgcgcgcc cggcgttcca gccgcgcgc ctcggcggtg gcttcggcct cgcggagggtg      540
gagctctacg gcgacgtcgt gctccgcttc gtcagccacc cggacggcgc cgacgcgccc      600
ttcctcccgg gtttcgaggg cgtcagcaac ccgggcgcgc tggactacgg cctccgcggg      660
ttcgaccacg tcgtcggcaa cgtgccggag ctcgcctcgg tagccgcgta catctccggg      720
ttcaccgggt tccacgagtt cgcgcgagtc accgccgagg acgtgggcac cgccgagagc      780
ggcctcaact cgggtggtgt cgcacaacaac ggggagaccg tgctgctgcc gctcaacgag      840
ccggtgcacg gcaccaagcg gcggagccag atacagacgt acctggacca ccacggcggc      900
ccgggggtgc agcacatcgc gctggccagc gacgacgtgc tggggacgct gagggagatg      960
cgggcgcgct ccgccatggg cggcttcgag ttcttggcgc cgcgcgcgc caactactac     1020
gacggcgtgc ggcggcgcgc cggggacgtg ctctcggagg agcagatcaa cgagtgccag     1080
gagctcgggg tgctcgtgga cagggatgac caggggtgtg tgctccagat cttaccaag     1140
ccagtaggag acaggccaac ctttttcttg gagatgatac aaaggattgg gtgcatggag     1200
aaggatgaga gtgggcagga gtaccagaag ggcggctgcg gcgggtttgg gaagggcaac     1260
ttctcggagc tgttcaagtc cattgaggag tatgagaaat cccttgaagc caagcaagcc     1320
cctacagttc aaggatccta g                                     1341
    
```

10 <210> 4
 <211> 446
 <212> PRT
 <213> *Oryza sativa*

15 <400> 4

```

Met Pro Pro Thr Pro Thr Pro Thr Ala Thr Thr Gly Ala Val Ser Ala
1                               5                               10                               15

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

Arg Phe Val Arg Ala Asn Pro Arg Ser Asp Arg Phe Gln Ala Leu Ala
                35                               40                               45

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

Arg Phe Ala Phe Ala Leu Gly Ala Pro Leu Ala Ala Arg Ser Asp Leu
    
```

ES 2 594 284 T3

65 70 75 80

Ser Thr Gly Asn Ser Ala His Ala Ser Leu Leu Leu Arg Ser Ala Ser
85 90 95

Val Ala Phe Leu Phe Thr Ala Pro Tyr Gly Gly Asp His Gly Val Gly
100 105 110

Ala Asp Ala Ala Thr Thr Ala Ser Ile Pro Ser Phe Ser Pro Gly Ala
115 120 125

Ala Arg Arg Phe Ala Ala Asp His Gly Leu Ala Val His Ala Val Ala
130 135 140

Leu Arg Val Ala Asp Ala Ala Asp Ala Phe Arg Ala Ser Val Ala Ala
145 150 155 160

Gly Ala Arg Pro Ala Phe Gln Pro Ala Asp Leu Gly Gly Gly Phe Gly
165 170 175

Leu Ala Glu Val Glu Leu Tyr Gly Asp Val Val Leu Arg Phe Val Ser
180 185 190

His Pro Asp Gly Ala Asp Ala Pro Phe Leu Pro Gly Phe Glu Gly Val
195 200 205

Ser Asn Pro Gly Ala Val Asp Tyr Gly Leu Arg Arg Phe Asp His Val
210 215 220

Val Gly Asn Val Pro Glu Leu Ala Pro Val Ala Ala Tyr Ile Ser Gly
225 230 235 240

Phe Thr Gly Phe His Glu Phe Ala Glu Phe Thr Ala Glu Asp Val Gly
245 250 255

Thr Ala Glu Ser Gly Leu Asn Ser Val Val Leu Ala Asn Asn Ala Glu
260 265 270

Thr Val Leu Leu Pro Leu Asn Glu Pro Val His Gly Thr Lys Arg Arg
275 280 285

Ser Gln Ile Gln Thr Tyr Leu Asp His His Gly Gly Pro Gly Val Gln
290 295 300

His Ile Ala Leu Ala Ser Asp Asp Val Leu Gly Thr Leu Arg Glu Met
305 310 315 320

Arg Ala Arg Ser Ala Met Gly Gly Phe Glu Phe Leu Ala Pro Pro Pro
325 330 335

Pro Asn Tyr Tyr Asp Gly Val Arg Arg Arg Ala Gly Asp Val Leu Ser
340 345 350

ES 2 594 284 T3

Glu Glu Gln Ile Asn Glu Cys Gln Glu Leu Gly Val Leu Val Asp Arg
 355 360 365

Asp Asp Gln Gly Val Leu Leu Gln Ile Phe Thr Lys Pro Val Gly Asp
 370 375 380

Arg Pro Thr Phe Phe Leu Glu Met Ile Gln Arg Ile Gly Cys Met Glu
 385 390 395 400

Lys Asp Glu Ser Gly Gln Glu Tyr Gln Lys Gly Gly Cys Gly Gly Phe
 405 410 415

Gly Lys Gly Asn Phe Ser Glu Leu Phe Lys Ser Ile Glu Glu Tyr Glu
 420 425 430

Lys Ser Leu Glu Ala Lys Gln Ala Pro Thr Val Gln Gly Ser
 435 440 445

<210> 5
 <211> 1281
 5 <212> ADN
 <213> Zea mays

<400> 5

atgcccccca	ccccacagc	cgccgcagcc	ggcgccgccc	tggcggcggc	atcagcagcg	60
gagcaggcgg	cggttcgcct	cgtgggccac	cgcaacttcg	tccgcttcaa	cccgcgctcc	120
gaccgcttcc	acacgctcgc	gttccaccac	gtggagctct	ggtgcgccga	cggggcctcc	180
gcccggggccc	gcttctcctt	cggcctgggc	gcgcccctcg	ccgcacgctc	cgacctctcc	240
acggggcaact	ccgcgcacgc	gtccctgctg	ctccgctccg	gctccctctc	cttctctctc	300
acgggcgcct	acgcgcacgg	cgccgacgct	gccaccgccc	cgetgccttc	cttctccgcc	360
gccgcgcgcg	ggcgcttcgc	agccgaccac	ggcctcgcgg	tgcgcgccgt	cgcgctccgc	420
gtcgccgacg	ccgaggacgc	cttcgcgcgc	ttccgcctcg	ccgaggtcga	gctctacggc	480
gacgtcgtgc	tccggtacgt	gagctaccgc	gacggcgccc	cgggcgagcc	cttctgcccg	540
gggttcgagg	gcgtggccag	ccccggggcg	gccgactacg	ggctgagcag	gttcgaccac	600
atcgtcggca	acgtgcccga	gctggcgccc	gccgcgcctc	acttcgccgg	cttcacgggg	660
ttccacgagt	tcgccgagtt	cacgacggag	gacgtgggca	ccgcggagag	cggcctcaac	720
tccatggtgc	tcgccaaaca	ctcggagaac	gtgctgctcc	cgcctcaacga	gccggtgcac	780
ggcaccaaagc	gccgcagcca	gatacaaacg	ttcctggacc	accacggcgg	ccccggcgtg	840
cagcacatgg	cgctggccag	cgacgacgtg	ctcaggacgc	tgaggagatg	gcaggcgcgc	900
tcggccatgg	gcggcttcga	gttcatggcg	cctcccacat	ccgactacta	tgacggcgtg	960
aggcggcgcg	ccggggacgt	gctcacggaa	gcacagatta	aggagtgccg	ggagctaggg	1020
gtgctggtgg	acagggatga	ccaggcgcgtg	ctgctccaaa	tcttcaccaa	gccagtgggg	1080
gacaggccaa	cgcgtgttctt	ggaaatcctc	caaaggatcg	ggtgcatgga	gaaggatgag	1140
aaggggcaag	aataccaaaa	gggtggctgc	ggcgggttcg	gcaagggaaa	cttctcgcag	1200
ctgttcaagt	ccatcgagga	ttatgagaag	tccttgaag	ccaagcaagc	tgctgcagca	1260
gctgcagctc	agggatccta	g				1281

10 <210> 6
 <211> 426
 <212> PRT
 <213> Zea mays

15 <400> 6

ES 2 594 284 T3

Met Pro Pro Thr Pro Thr Ala Ala Ala Ala Gly Ala Ala Val Ala Ala
1 5 10 15

Ala Ser Ala Ala Glu Gln Ala Ala Phe Arg Leu Val Gly His Arg Asn
20 25 30

Phe Val Arg Phe Asn Pro Arg Ser Asp Arg Phe His Thr Leu Ala Phe
35 40 45

His His Val Glu Leu Trp Cys Ala Asp Ala Ala Ser Ala Ala Gly Arg
50 55 60

Phe Ser Phe Gly Leu Gly Ala Pro Leu Ala Ala Arg Ser Asp Leu Ser
65 70 75 80

Thr Gly Asn Ser Ala His Ala Ser Leu Leu Leu Arg Ser Gly Ser Leu
85 90 95

Ser Phe Leu Phe Thr Ala Pro Tyr Ala His Gly Ala Asp Ala Ala Thr
100 105 110

Ala Ala Leu Pro Ser Phe Ser Ala Ala Ala Ala Arg Arg Phe Ala Ala
115 120 125

Asp His Gly Leu Ala Val Arg Ala Val Ala Leu Arg Val Ala Asp Ala
130 135 140

Glu Asp Ala Phe Arg Gly Phe Arg Leu Ala Glu Val Glu Leu Tyr Gly
145 150 155 160

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

Pro Phe Leu Pro Gly Phe Glu Gly Val Ala Ser Pro Gly Ala Ala Asp
180 185 190

Tyr Gly Leu Ser Arg Phe Asp His Ile Val Gly Asn Val Pro Glu Leu
195 200 205

Ala Pro Ala Ala Ala Tyr Phe Ala Gly Phe Thr Gly Phe His Glu Phe
210 215 220

Ala Glu Phe Thr Thr Glu Asp Val Gly Thr Ala Glu Ser Gly Leu Asn
225 230 235 240

Ser Met Val Leu Ala Asn Asn Ser Glu Asn Val Leu Leu Pro Leu Asn
245 250 255

Glu Pro Val His Gly Thr Lys Arg Arg Ser Gln Ile Gln Thr Phe Leu
260 265 270

ES 2 594 284 T3

Asp His His Gly Gly Pro Gly Val Gln His Met Ala Leu Ala Ser Asp
 275 280 285

Asp Val Leu Arg Thr Leu Arg Glu Met Gln Ala Arg Ser Ala Met Gly
 290 295 300

Gly Phe Glu Phe Met Ala Pro Pro Thr Ser Asp Tyr Tyr Asp Gly Val
 305 310 315 320

Arg Arg Arg Ala Gly Asp Val Leu Thr Glu Ala Gln Ile Lys Glu Cys
 325 330 335

Gln Glu Leu Gly Val Leu Val Asp Arg Asp Asp Gln Gly Val Leu Leu
 340 345 350

Gln Ile Phe Thr Lys Pro Val Gly Asp Arg Pro Thr Leu Phe Leu Glu
 355 360 365

Ile Ile Gln Arg Ile Gly Cys Met Glu Lys Asp Glu Lys Gly Gln Glu
 370 375 380

Tyr Gln Lys Gly Gly Cys Gly Gly Phe Gly Lys Gly Asn Phe Ser Gln
 385 390 395 400

Leu Phe Lys Ser Ile Glu Asp Tyr Glu Lys Ser Leu Glu Ala Lys Gln
 405 410 415

Ala Ala Ala Ala Ala Ala Ala Gln Gly Ser
 420 425

<210> 7
 <211> 1323
 5 <212> ADN
 <213> Avena sativa

<400> 7
 atgccgccca cccccgccac cgccaccggc gccgcgcggc ccgccgtgac tccagagcac 60
 gcggcccgga gctttccccg agtgggtccg gtcaaccgc gcagcgaccg cttccccgtg 120
 ctctccttcc accacgtcga gctctggtgc gccgacgccg cctcagcggc cggacgcttc 180
 tccttcgcgc tcggcgcgcc gctcgcgcc cggtccgacc tctccacggg gaactccgcg 240
 cacgcctccc tcctgctccg ctccgggcgcc ctccgcttcc tcttcacggc gccctacgcg 300
 ccgccgcgcc aggaggccgc cacggccgca gccaccgct ccatcccctc cttctccgcc 360
 gacgccgcgc ggacgttcgc cgcgcgccac ggctcgcgg tgcgctccgt cgggggtccgc 420
 gtcgctgacg ccgccgaggc cttccgcgtc agcgtagccg gcggcgctcg cccggccttc 480
 gccccagccg acctcggcca tggcttcggc ctccgccagg tcgagctcta cggcgacgctc 540
 gtgctacgct tcgtcagcta cccggacgag acagacctgc cattcctgcc aggggttcgag 600
 cgcgtgagca gccccggcgc cgtggactac ggctcaagc ggttcgacca cgtcgtgggc 660
 aacgtcccgg agatggcccc ggtcatagac tacatgaaag gcttcttggg gttccacgag 720
 ttccgcgagt tcaccgccga ggacgtgggc acgaccgaga gcgggctcaa ctcggtggtg 780
 ctgcacaaca actccgaggc cgtgctgctg ccgctcaacg agcccgtgca cggcacaag 840
 cgacggagcc agatacagac gtacctggag tatcacggcg ggcccggcgt gcagcacatc 900

ES 2 594 284 T3

```

gcgctcgcca gcaacgacgt gctcaggacg ctcagggaga tgcggggcgcg cacgcccattg 960
ggcggcttcg agttcatggc gccaccgcag gcgaaatact atgaaggcgt gcggcgcctc 1020
gcaggtgacg tgctctcggg agagcagatc aaggaatgcc aggagctggg ggtgctagtc 1080
gacagggatg atcaaggggt gttgctccaa atcttcacca agccagtagg ggacaggcca 1140
acgtttttcc tggagatgat ccaaagaatc ggggtgcatgg agaaggacga ggtcggggcaa 1200
gagtaccaga aggggtggctg cggcggggtt ggcaagggca atttctccga gctgttcaag 1260
tcattgagg actatgagaa atcccttgag gtcaagcaat ctgttgtagc tcagaaatcc 1320
tag 1323

```

<210> 8

<211> 440

5 <212> PRT

<213> Avena sativa

<400> 8

```

Met Pro Pro Thr Pro Ala Thr Ala Thr Gly Ala Ala Ala Ala Ala Val
1          5          10          15

Thr Pro Glu His Ala Ala Arg Ser Phe Pro Arg Val Val Arg Val Asn
          20          25          30

Pro Arg Ser Asp Arg Phe Pro Val Leu Ser Phe His His Val Glu Leu
          35          40          45

Trp Cys Ala Asp Ala Ala Ser Ala Ala Gly Arg Phe Ser Phe Ala Leu
          50          55          60

Gly Ala Pro Leu Ala Ala Arg Ser Asp Leu Ser Thr Gly Asn Ser Ala
65          70          75          80

His Ala Ser Leu Leu Leu Arg Ser Gly Ala Leu Ala Phe Leu Phe Thr
          85          90          95

Ala Pro Tyr Ala Pro Pro Pro Gln Glu Ala Ala Thr Ala Ala Thr
          100          105          110

Ala Ser Ile Pro Ser Phe Ser Ala Asp Ala Ala Arg Thr Phe Ala Ala
          115          120          125

Ala His Gly Leu Ala Val Arg Ser Val Gly Val Arg Val Ala Asp Ala
          130          135          140

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

Ala Pro Ala Asp Leu Gly His Gly Phe Gly Leu Ala Glu Val Glu Leu
          165          170          175

Tyr Gly Asp Val Val Leu Arg Phe Val Ser Tyr Pro Asp Glu Thr Asp
          180          185          190

Leu Pro Phe Leu Pro Gly Phe Glu Arg Val Ser Ser Pro Gly Ala Val

```

ES 2 594 284 T3

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

<210> 9
 <211> 1077

5 <212> ADN

<213> Pseudomonas fluorescens

ES 2 594 284 T3

<400> 9

```

atggcagatc tatacgaaaa cccaatgggc ctgatgggct ttgaattcat cgaattcgcg      60
tcgccgacgc cgggtaccct ggagccgata ttogagatca tgggcttcac caaagtcgcg     120
accacacggt ccaagaacgt gcacctgtac cgccagggcg agatcaacct gatcctcaac     180
aacgagccca acagcatcgc ctctactttt ggggcccgaac acggcccgtc ggtgtgcggc     240
atggcgttcc gcgtgaagga ctgcgaaaag gcctacaacc gcgccctgga actcggcgcc     300
cagccgatcc atattgacac cgggcccgatg gaattgaacc tgcccggcgat caagggcac     360
ggcggcgcgc cgttgtacct gatcgaccgt ttogggogaag gcagctcgat ctacgacac     420
gacttcgtgt acctcgaagg tgtggagcgc aatccggtcg gtgcaggtct caaagtcac     480
gaccacctga ccacaacgt ctatcgccgc cgcattgtct actgggcca cttctacgag     540
aaattgttca acttcctgta agcgcgttac ttogatatca agggcgagta caccggcctg     600
acttccaagg ccatgagtgc gccggacggc atgatccgca tcccgtgaa cgaagagtcg     660
tccaagggcg cggggcagat cgaagagttc ctgatgcagt tcaacggcga aggcattccag     720
cacgtggcgt tctcaccga cgacctggtc aagacctggg acgcgttgaa gaaaatcggc     780
atgcgcttca tgaccgcgcc gccagacact tattacgaaa tgctcgaagg ccgctgcct     840
gaccacggcg agccgggtga tcaactgcag gcacgcggta tctgctgga cggatcttc     900
gtggaaggcg acaaacgcct gctgctgcag atcttctcgg aaacctgat gggcccgggtg     960
ttcttcgaat tcatccagcg caagggcgac gatggggttg gcgagggcaa cttcaaggcg    1020
ctgttcgagt ccatcgaacg tgaccaggtg cgtcgtggtg tattgaccgc cgattaa      1077
    
```

<210> 10

5 <211> 358

<212> PRT

<213> *Pseudomonas fluorescens*

<400> 10

```

Met Ala Asp Leu Tyr Glu Asn Pro Met Gly Leu Met Gly Phe Glu Phe
1           5           10          15

Ile Glu Phe Ala Ser Pro Thr Pro Gly Thr Leu Glu Pro Ile Phe Glu
          20          25          30

Ile Met Gly Phe Thr Lys Val Ala Thr His Arg Ser Lys Asn Val His
          35          40          45

Leu Tyr Arg Gln Gly Glu Ile Asn Leu Ile Leu Asn Asn Glu Pro Asn
          50          55          60

Ser Ile Ala Ser Tyr Phe Ala Ala Glu His Gly Pro Ser Val Cys Gly
65          70          75          80

Met Ala Phe Arg Val Lys Asp Ser Gln Lys Ala Tyr Asn Arg Ala Leu
          85          90          95

Glu Leu Gly Ala Gln Pro Ile His Ile Asp Thr Gly Pro Met Glu Leu
          100         105         110

Asn Leu Pro Ala Ile Lys Gly Ile Gly Gly Ala Pro Leu Tyr Leu Ile
          115         120         125

Asp Arg Phe Gly Glu Gly Ser Ser Ile Tyr Asp Ile Asp Phe Val Tyr
          130         135         140
    
```

10

ES 2 594 284 T3

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

<210> 11

5 <400> 000

<210> 12

<400> 000

10

<210> 13

<211> 1053

<212> ADN

ES 2 594 284 T3

<213> Synechococcus sp.

<400> 13

```

atgaaccggt ccattcgaat tgtccaaggg atccaccacc tgcacttota cctttgggat      60
ctgccccgtt ggcgggaaca cttttgtcgg gtttggggct tcgggtggc aagcgacgcc      120
ggcaacaccc tggagctgga gcagggatcc ctgcgcttgc gcctgtctca gccggcacgg      180
gcgggggacg aggtggaccg ccatttgcag cggcatgggc cgggggtggt ggatgtggcc      240
ttggcgggtg gagagcagga gctaccggcc ttggcggagc tgttgccggg ccgagggccc      300
caactggcgt ggatcccggc agcagcggcg ctctgcctcc acacccccta cgggatccgg      360
cattctctga tccctggccc cttgggatgcc gccctgcgcg aagcgggcct gttttcccac      420
tgggatcacg tgggtgtttaa cgtggagcag ggatcccctgc aggcggcagc cgactggtat      480
gggcgggtgc tgggctggcg ggcgctgtac cgctacagca tgggcaccgc cacctccggc      540
ctggaaagcg tgggtggtggg ggatccggaa gcggggatcc aatgggcat caacgagccc      600
acctgtgccg cttcccagat tcaggagttt ttgcatgccc atggcggccc gggcattcag      660
cacgcggcgc tgcacagctc agacattggt gccagcctgc gccggttgcg gcagggggga      720
gtggactttt tgcaagtggc gccgcagtac tacaccagcc tggaaaggga gctggggttg      780
gcgctccgtt ctgcccttgg gcaggccatc tccctggcaag acctggtgga gcagcagatc      840
cttctggatg ctaccctgcc cgtttctgat ggccaggatc gcccccttct gctgcagacc      900
tttaccagc ccccttttgg tcggcccacc tttttctttg aagtcattca acggctaggc      960
ggggccacgg gctttggcga ggccaatttt caggccttgt togaggccct ggaacggcaa     1020
cagcgacagc gacaccaggc gctgaccctc tag                                     1053

```

5

<210> 14

<211> 350

<212> PRT

<213> Synechococcus sp.

10

<400> 14

```

Met Asn Pro Ser Ile Arg Ile Val Gln Gly Ile His His Leu His Phe
1                               5                               10                               15

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

Gly Phe Arg Val Ala Ser Asp Ala Gly Asn Thr Leu Glu Leu Glu Gln
                35                               40                               45

Gly Ser Leu Arg Leu Arg Leu Ser Gln Pro Ala Arg Ala Gly Asp Glu
                50                               55                               60

Val Asp Arg His Leu Gln Arg His Gly Pro Gly Val Val Asp Val Ala
65                               70                               75                               80

Leu Ala Val Gly Glu Gln Glu Leu Pro Ala Leu Ala Glu Leu Leu Arg
                85                               90                               95

Gly Arg Gly Ala Gln Leu Ala Trp Ile Pro Ala Ala Ala Ala Leu Cys
                100                               105                               110

Leu His Thr Pro Tyr Gly Ile Arg His Ser Leu Ile Pro Gly Pro Leu

```


ES 2 594 284 T3

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

<210> 15

<211> 1149

5 <212> ADN

<213> *Blepharisma japonicum*

<400> 15

atgacttatt acgacaagca agaaacgcgt ccagatcttg gcgaattcta tggtttccat 60
 cacgttcggt tttagctctc caactcagag caagccgctt cgttctacac atctcgcttt 120

ES 2 594 284 T3

```

gggttttctc cggttgccta tgaaggattg gaaacaggaa accaaaaatt ctgtaccaat 180
gtcgtccgaa gcaaccatgt agtcacgct tttacctcag ctctcactcc tgaagacaat 240
gaagtgaacc gtcacgttgg caagcatagt gatggagttc aagacattgc ctttagtgta 300
agtgacgcaa gagggatgta tgagaaagcg atagctaaag gctgtaaaag cttccgtgag 360
ccacaggttt tacaagatca atttgatct gttataatag cgtctctcca gacttatgga 420
gacactgttc acacattagt ccaaaatgtc gactatacag gacccttttt gcctggcttc 480
agagcaatca caaaagatga tccattaaac tctgcctttc ctcaggtaaa ttatgacatt 540
attgatcatg ttgtaggaaa tcagcctggg ggcgatatga ctctacagt agaatggat 600
gagaaatc tagaatttca tcgatattgg tctgctgatg agtctgtaat ccataccgat 660
tattcagcat taaggctctgt tgtggttgc gattgggatg aagtgatcaa aatgcctatt 720
aatgagcctg ctgatggact tagaaaaagt caaatccaag aatatgtcga atattatgg 780
ggagcaggcg tacaacatat tgccttaaaa gtcaatgata ttatttcagt aataagcacc 840
ttaagggcta gaggtgtgga attcttagaa gttcctccta aatattatga tagcttaaga 900
aaaagacttg cgcattctgc ggtacaaatt gaagaagact taaaaagaat tgaagacct 960
catattttgg ttgactttga cgaccgtggg tatttacttc agattttcac aaaaccagta 1020
gaagacagac ctactctgtt ttatgaaatt attcaaagac ataataacaa tggattcggga 1080
attggaatt ttaaagccct atttgaatca ttggaacaag agcaagaaag aagaggtaat 1140
ttgatctaa 1149

```

<210> 16

<211> 382

5 <212> PRT

<213> *Blepharisma japonicum*

<400> 16

```

Met Thr Tyr Tyr Asp Lys Gln Glu Thr Arg Pro Asp Leu Gly Glu Phe
1                5                10                15

Tyr Gly Phe His His Val Arg Phe Tyr Val Ser Asn Ser Glu Gln Ala
                20                25                30

Ala Ser Phe Tyr Thr Ser Arg Phe Gly Phe Ser Pro Val Ala Tyr Glu
                35                40                45

Gly Leu Glu Thr Gly Asn Gln Lys Phe Cys Thr Asn Val Val Arg Ser
50                55                60

Asn His Val Val Ile Ala Phe Thr Ser Ala Leu Thr Pro Glu Asp Asn
65                70                75                80

Glu Val Asn Arg His Val Gly Lys His Ser Asp Gly Val Gln Asp Ile
                85                90                95

Ala Phe Ser Val Ser Asp Ala Arg Gly Met Tyr Glu Lys Ala Ile Ala
                100                105                110

Lys Gly Cys Lys Ser Phe Arg Glu Pro Gln Val Leu Gln Asp Gln Phe
                115                120                125

Gly Ser Val Ile Ile Ala Ser Leu Gln Thr Tyr Gly Asp Thr Val His
130                135                140

```

ES 2 594 284 T3

Thr Leu Val Gln Asn Val Asp Tyr Thr Gly Pro Phe Leu Pro Gly Phe
 145 150 155 160

Arg Ala Ile Thr Lys Asp Asp Pro Leu Asn Ser Ala Phe Pro Gln Val
 165 170 175

Asn Tyr Asp Ile Ile Asp His Val Val Gly Asn Gln Pro Gly Gly Asp
 180 185 190

Met Thr Pro Thr Val Glu Trp Tyr Glu Lys Tyr Leu Glu Phe His Arg
 195 200 205

Tyr Trp Ser Ala Asp Glu Ser Val Ile His Thr Asp Tyr Ser Ala Leu
 210 215 220

Arg Ser Val Val Val Ala Asp Trp Asp Glu Val Ile Lys Met Pro Ile
 225 230 235 240

Asn Glu Pro Ala Asp Gly Leu Arg Lys Ser Gln Ile Gln Glu Tyr Val
 245 250 255

Glu Tyr Tyr Gly Gly Ala Gly Val Gln His Ile Ala Leu Lys Val Asn
 260 265 270

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

Leu Glu Val Pro Pro Lys Tyr Tyr Asp Ser Leu Arg Lys Arg Leu Ala
 290 295 300

His Ser Ala Val Gln Ile Glu Glu Asp Leu Lys Arg Ile Glu Asp Leu
 305 310 315 320

His Ile Leu Val Asp Phe Asp Asp Arg Gly Tyr Leu Leu Gln Ile Phe
 325 330 335

Thr Lys Pro Val Glu Asp Arg Pro Thr Leu Phe Tyr Glu Ile Ile Gln
 340 345 350

Arg His Asn Asn Asn Gly Phe Gly Ile Gly Asn Phe Lys Ala Leu Phe
 355 360 365

Glu Ser Leu Glu Gln Glu Gln Glu Arg Arg Gly Asn Leu Ile
 370 375 380

<210> 17

<211> 1206

5 <212> ADN

<213> Rhodococcus sp.

<400> 17

atgacgatcg agcagactct caccgacaag gaacgcctgg caggtctcga cctcggccag

60

ES 2 594 284 T3

```

ctcgagcagt tggtcgggct cgtcaggtac gacggcacc cgcaccggtt cccggtcagc 120
ggctgggatg ccgtcgtctg ggtggtcggc aacgccacc agaccgcca ctacttccag 180
tcoggttgc ggatgacct cgtcgcctac tccggacca ccaccgcaa ccgcgaccac 240
cacagcttcg tctcgaatc cggggccgtc cgttcgtca tcaaaggcgc cgtgaacccg 300
gacagcccc tgatcgacca ccaccgcacc cacggcgacg gcgtcgtcga catcgccctc 360
gocgtcccc acgtcgaaa gtgcatcgc cacgcccgc cccaggcgc caccgtcctc 420
gacgaacccc acgacgtgac cgacgaccac ggcaccgtcc gcctcgcgc gatcgccacc 480
tacggcgaca cccgccacac cctcgtcgc cgcagccact acaccgccc ctacctgcc 540
ggctacaccg cccgcacctc cggccacacc aaacgggacg gggcaccaa gcgcctgttc 600
caggccctcg accacgtcgt cggcaacgtc gaactcggca agatggacca ctgggtcgc 660
ttctacaacc gggtcattgg ctttacgaac atggccgagt tcgtcggcga ggacatcgc 720
accgactact ccgctgatg gagcaaggtc gtctccaacg gcaaccaccg ggtcaagttc 780
ccctcaacg aaccgacct cggcaagaaa cgtcgcaga tcgacgaata cctcgacttc 840
taccgggccc cggcgccca gcacctggcc ctggccacca atgacatcct caccgctgc 900
gaccagctga ccccgaggg cgtcaggttc ctggccacc cgcactccta ctacgaggac 960
ccgaaactgc gggcccggat cggcaacgtc cgcgccccca tcgcccgaact gcagaaaacgc 1020
ggcatcctcg tcgaccgca cgaagacgpc tacctcgtgc agatcttac caaacccctc 1080
gtcgaccggc ccaccgtgtt cttcgaactc atcgaacgpc accgctccct cggcttcggc 1140
atcggcaact tcaaagcct cttcaggcc atcgaacgpc aacaagcgc ccgcgaaaac 1200
ttctga 1206

```

<210> 18
 <211> 401
 5 <212> PRT
 <213> Rhodococcus sp.

<400> 18

```

Met Thr Ile Glu Gln Thr Leu Thr Asp Lys Glu Arg Leu Ala Gly Leu
1           5           10           15

Asp Leu Gly Gln Leu Glu Gln Leu Val Gly Leu Val Glu Tyr Asp Gly
20           25           30

Thr Arg Asp Pro Phe Pro Val Ser Gly Trp Asp Ala Val Val Trp Val
35           40           45

Val Gly Asn Ala Thr Gln Thr Ala His Tyr Phe Gln Ser Ala Phe Gly
50           55           60

Met Thr Leu Val Ala Tyr Ser Gly Pro Thr Thr Gly Asn Arg Asp His
65           70           75           80

His Ser Phe Val Leu Glu Ser Gly Ala Val Arg Phe Val Ile Lys Gly
85           90           95

Ala Val Asn Pro Asp Ser Pro Leu Ile Asp His His Arg Thr His Gly
100          105          110

Asp Gly Val Val Asp Ile Ala Leu Ala Val Pro Asp Val Asp Lys Cys
115          120          125

Ile Ala His Ala Arg Ala Gln Gly Ala Thr Val Leu Asp Glu Pro His

```

ES 2 594 284 T3

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

ES 2 594 284 T3

<210> 19
 <211> 1107
 <212> ADN
 <213> *Picrophilus torridus*

5

<400> 19
 atgtatggca aaaatttaat ctcagaacta agggaaaagg agatctttaa acgattacat 60
 cacgtggaat tttacgttag cagtgccaaa acatgggtcat atttcatgaa caggggtcctt 120
 ggatttaaaa cagtggcata tgccgggtcca gaaaccggga taagggacaa gatatcctat 180
 gttatgtccc agggcactgc aaggatatct tttacatcat caatgaatga tgatagctat 240
 atatcgaatc atgttaaaaa acacggggat ggcgtaaagg atatagcact tgaggtcgat 300
 gatctggacg aggcaaaaag cctgatagaa aagtatggaa caaaggtttc aaaaataaat 360
 gaaaataaagg atggaaatgg aaagataaga actgcagaga taaaaacgta cggtgaaacc 420
 gttcatacat taatagaaac cggggattac aatggcgtat tcatgcccggtt tatgaggaa 480
 tctgaaataa attcaaaaaa cactgggata aaaaagatcg atcatatagt tggaaatgtc 540
 tatgagggcg agatggatag ctgggttaat ttttacatag aaaaacttgg ctttgagcat 600
 ttaataacct ttgatgataa agatataaga actgattaca gcgcattaag atcaaagggtt 660
 gtaaaaataca atgacgatat cgtattttcca ataaatgagc ctgcaaaggg cttaagaaaa 720
 tcacagatag aggaatatct tgactattac aggtctgagg gcgttcagca catagcactg 780
 ttaactgatg atataataaa aactgtatcc atgatggagg aaaacggcat agaattttta 840
 aaaacaccag gatcactacta tgaatcccta tcatcaagga taggctcaat agacgaggat 900
 ttaaataaaa tagagaaaca taacatactt gtggatcgtg atgagaacgg atacctatta 960
 cagatcttca caaagcctgt tactgacagg ccaacgttct tctttgaggt catacagaga 1020
 aaggggtgcaa ggtcattcgg caacggtaac ttaaggcac tttttgaggg gatagaaagg 1080
 gagcaggcaa agagaggaaa cctatga 1107

<210> 20
 <211> 368
 <212> PRT
 <213> *Picrophilus torridus*

10

<400> 20
 Met Tyr Gly Lys Asn Leu Ile Ser Glu Leu Arg Glu Lys Glu Ile Phe
 1 5 10 15
 Lys Arg Leu His His Val Glu Phe Tyr Val Ser Ser Ala Lys Thr Trp
 20 25 30
 Ser Tyr Phe Met Asn Arg Gly Leu Gly Phe Lys Thr Val Ala Tyr Ala
 35 40 45
 Gly Pro Glu Thr Gly Ile Arg Asp Lys Ile Ser Tyr Val Met Ser Gln
 50 55 60
 Gly Thr Ala Arg Ile Ser Phe Thr Ser Ser Met Asn Asp Asp Ser Tyr
 65 70 75 80
 Ile Ser Asn His Val Lys Lys His Gly Asp Gly Val Lys Asp Ile Ala
 85 90 95
 Leu Glu Val Asp Asp Leu Asp Glu Ala Lys Ser Leu Ile Glu Lys Tyr

15

ES 2 594 284 T3

			100							105						110
Gly	Thr	Lys	Val	Ser	Lys	Ile	Asn	Glu	Ile	Lys	Asp	Gly	Asn	Gly	Lys	
		115					120					125				
Ile	Arg	Thr	Ala	Glu	Ile	Lys	Thr	Tyr	Gly	Glu	Thr	Val	His	Thr	Leu	
	130					135					140					
Ile	Glu	Thr	Gly	Asp	Tyr	Asn	Gly	Val	Phe	Met	Pro	Gly	Tyr	Glu	Glu	
145					150					155					160	
Ser	Glu	Ile	Asn	Ser	Lys	Asn	Thr	Gly	Ile	Lys	Lys	Ile	Asp	His	Ile	
				165					170					175		
Val	Gly	Asn	Val	Tyr	Glu	Gly	Glu	Met	Asp	Ser	Trp	Val	Asn	Phe	Tyr	
			180					185					190			
Ile	Glu	Lys	Leu	Gly	Phe	Glu	His	Leu	Ile	Thr	Phe	Asp	Asp	Lys	Asp	
		195					200					205				
Ile	Arg	Thr	Asp	Tyr	Ser	Ala	Leu	Arg	Ser	Lys	Val	Val	Lys	Tyr	Asn	
	210					215					220					
Asp	Asp	Ile	Val	Phe	Pro	Ile	Asn	Glu	Pro	Ala	Lys	Gly	Leu	Arg	Lys	
225					230					235					240	
Ser	Gln	Ile	Glu	Glu	Tyr	Leu	Asp	Tyr	Tyr	Arg	Ser	Glu	Gly	Val	Gln	
				245					250					255		
His	Ile	Ala	Leu	Leu	Thr	Asp	Asp	Ile	Ile	Lys	Thr	Val	Ser	Met	Met	
			260					265						270		
Glu	Glu	Asn	Gly	Ile	Glu	Phe	Leu	Lys	Thr	Pro	Gly	Ser	Tyr	Tyr	Glu	
		275					280					285				
Ser	Leu	Ser	Ser	Arg	Ile	Gly	Ser	Ile	Asp	Glu	Asp	Leu	Asn	Glu	Ile	
	290					295					300					
Glu	Lys	His	Asn	Ile	Leu	Val	Asp	Arg	Asp	Glu	Asn	Gly	Tyr	Leu	Leu	
305					310					315					320	
Gln	Ile	Phe	Thr	Lys	Pro	Val	Thr	Asp	Arg	Pro	Thr	Phe	Phe	Phe	Glu	
				325					330						335	
Val	Ile	Gln	Arg	Lys	Gly	Ala	Arg	Ser	Phe	Gly	Asn	Gly	Asn	Phe	Lys	
			340					345					350			
Ala	Leu	Phe	Glu	Ala	Ile	Glu	Arg	Glu	Gln	Ala	Lys	Arg	Gly	Asn	Leu	
		355					360						365			

<210> 21

<211> 1164

ES 2 594 284 T3

<212> ADN

<213> Kordia algicida

<400> 21

```

atggcagcag aaataaaaaa cttaaaagat ttacaaaata cagaatacgg actcaaaaaa    60
ttatttgacg aagcagaaga ctttcttcca cttttaggaa cagactacgt agaattatac    120
gtcgggaaac ccaaacaatc ggcacatttc tacaaaacgg cttttggttt tcaatcagaa    180
gcttacgcag gattggaaac aggattaacc gacagagttt catacgtatt aaaacaagat    240
aaaattcgct tggctttaac aacaccatta ggaaaagggt gcgaaatcaa tgagcatatc    300
gatttacacg gcgatggcgt aaaagtagta gcactttggg tagaagatgc tacaaaagcc    360
tttgaagaaa cgaccaaaaag aggcgcaaaa ccgtacatgg aaccaacaaa agaagaagat    420
gaaaacggat atgtaattcg ctcaggaatc tatacgtacg gagaacgggt tcatgttttt    480
gtagaacgta aaaactataa cggagtcttt ttaccaggat atcaaagatg ggaatctcac    540
tacaatccgg agccagttgg cttaaaattc atcgatcaca tggtaggaaa tgtaggttgg    600
ggagaaatga aagaatgggt tgaattctac gcgaaagtaa tgggatttgc gcaaattatc    660
tcctttacag atgatgatat ttctaccgat tttactgcgt tgatgagtaa agtaatgagt    720
aatggaaatg gtagaatcaa atttccaatc aatgaaccgg cagaaggaaa aaagaaatcg    780
caaattgaag aatatctaga cttttacaat ggttcaggag tacaacatat tgcggttgct    840
acagacaata ttattgatac ggtttcgcaa atgcgcgaac gtggagtata attcttatac    900
gttccagata catattatga tgacttgta gaacgtgttg gcgacatcga tgaagatgta    960
gaagaactca aaaaacacgg aatcttaatt gatcgtgatg aagaaggata cttattgcag   1020
ttatttacca aaaccattgt agacagacca acaatgttct ttgaagtcac tcagcgtaaa   1080
ggcgcacaat catttggagt aggaaacttt aaagctttat ttgaagcgat agaaagagaa   1140
caagctgctc gcggaacatt gtaa                                     1164

```

5

<210> 22

<211> 387

<212> PRT

10 <213> Kordia algicida

<400> 22

```

Met Ala Ala Glu Ile Lys Asn Leu Lys Asp Leu Gln Asn Thr Glu Tyr
1           5           10           15

Gly Leu Lys Lys Leu Phe Asp Glu Ala Glu Asp Phe Leu Pro Leu Leu
          20           25           30

Gly Thr Asp Tyr Val Glu Leu Tyr Val Gly Asn Ala Lys Gln Ser Ala
          35           40           45

His Phe Tyr Lys Thr Ala Phe Gly Phe Gln Ser Glu Ala Tyr Ala Gly
          50           55           60

Leu Glu Thr Gly Leu Thr Asp Arg Val Ser Tyr Val Leu Lys Gln Asp
65           70           75           80

Lys Ile Arg Leu Val Leu Thr Thr Pro Leu Gly Lys Gly Gly Glu Ile
          85           90           95

Asn Glu His Ile Asp Leu His Gly Asp Gly Val Lys Val Val Ala Leu
          100          105          110

```


ES 2 594 284 T3

Trp Val Glu Asp Ala Thr Lys Ala Phe Glu Glu Thr Thr Lys Arg Gly
 115 120 125

Ala Lys Pro Tyr Met Glu Pro Thr Lys Glu Glu Asp Glu Asn Gly Tyr
 130 135 140

Val Ile Arg Ser Gly Ile Tyr Thr Tyr Gly Glu Thr Val His Val Phe
 145 150 155 160

Val Glu Arg Lys Asn Tyr Asn Gly Val Phe Leu Pro Gly Tyr Gln Arg
 165 170 175

Trp Glu Ser His Tyr Asn Pro Glu Pro Val Gly Leu Lys Phe Ile Asp
 180 185 190

His Met Val Gly Asn Val Gly Trp Gly Glu Met Lys Glu Trp Cys Glu
 195 200 205

Phe Tyr Ala Lys Val Met Gly Phe Ala Gln Ile Ile Ser Phe Thr Asp
 210 215 220

Asp Asp Ile Ser Thr Asp Phe Thr Ala Leu Met Ser Lys Val Met Ser
 225 230 235 240

Asn Gly Asn Gly Arg Ile Lys Phe Pro Ile Asn Glu Pro Ala Glu Gly
 245 250 255

Lys Lys Lys Ser Gln Ile Glu Glu Tyr Leu Asp Phe Tyr Asn Gly Ser
 260 265 270

Gly Val Gln His Ile Ala Val Ala Thr Asp Asn Ile Ile Asp Thr Val
 275 280 285

Ser Gln Met Arg Glu Arg Gly Val Glu Phe Leu Tyr Val Pro Asp Thr
 290 295 300

Tyr Tyr Asp Asp Leu Leu Glu Arg Val Gly Asp Ile Asp Glu Asp Val
 305 310 315 320

Glu Glu Leu Lys Lys His Gly Ile Leu Ile Asp Arg Asp Glu Glu Gly
 325 330 335

Tyr Leu Leu Gln Leu Phe Thr Lys Thr Ile Val Asp Arg Pro Thr Met
 340 345 350

Phe Phe Glu Val Ile Gln Arg Lys Gly Ala Gln Ser Phe Gly Val Gly
 355 360 365

Asn Phe Lys Ala Leu Phe Glu Ala Ile Glu Arg Glu Gln Ala Ala Arg
 370 375 380

Gly Thr Leu

385

ES 2 594 284 T3

<210> 23
 <211> 1323
 <212> ADN
 5 <213> Sorghum bicolor

<400> 23
 atgccccga cccccaccac agccgccgca accggcgccg ccgtggcggc ggcacacagc 60
 gagcaggcgg cgttcgcgct cgtggggccac cgcaacttcg tcgcgctgaa cccgcgctcc 120
 gaccgcttcc acacgctcgc gttccaccac gtggagctct ggtgcgcgca cgcggcctcc 180
 gccgcggggc gcttctcctt cgggctcggc gcgcgcctcg ccgcgcggtc cgcacctctc 240
 acggggaaca ccgcgcacgc gtccctgctg ctccgctcgg gcgcctcgcg gttcctcttc 300
 acggcgccct acgcgcacgg cgcgcacgcc gccacggcct cgtgcgcctc cttctccgcc 360
 gccgaggcgc ggcgcttcgc ggccgaccac ggccctcggg tcgcgcgcgt cgcgctccgc 420
 gtggccgacg cggaggacgc ctccgcgcgc agcgtcgcgg ccggcgcgcg cccggcgctc 480
 gagcccgctg agctcggcct cggcttccgc ctgcgcgaag tcgagctcta cggcgacgct 540
 gtgctccggg acgtgagcta ccgcgacgac gcggacgcgt ccttctcgcg ggggttcgctg 600
 ggcgtgacca gccccggcgc ggccgactac gggctgagga ggttcgacca catcgtcggc 660
 aacgtgccgg agctggcgcg ggccggccgc tacttcgctg gcttcacggg gttccacgag 720
 ttgcgcgagt tcacggcgga ggacgtgggc accacggaga gcgggctcaa ctogatggtg 780
 ctgcccaaca acgcggagaa cgtgctgctc ccaactcaac agccggtgca cggcaccaag 840
 cgcgcagacc agatacagac gtacttggac caccacggcg gcccggcgt gcagcacatg 900
 gcgctggcca gcgacgacgt gctcaggacg ctgagggaga tgcaagcgcg ctccggccatg 960
 ggccgcttcg agttcatggc gcctccggcg cccgaatact atgacggcgt gaggcggcgc 1020
 gccggggacg tgctcacgga ggcacagatt aaggagtgtc aggaactagg ggtgctggtg 1080
 gacagagatg accaggcgtg gctgctccag atcttcacca agccagtggg ggacaggcca 1140
 acgttgttct tggagatcat tcaaaggatc ggggtgcatg agaaggatga gaaggggcaa 1200
 gaataccaga aggggtggctg tggcgggttt ggcaagggaa acttctcca gctgttcaaa 1260
 tccattgagg attatgagaa gtcccttgaa gctaagcaag ctgcagcagc tcagggatcc 1320
 tag 1323

10 <210> 24
 <211> 440
 <212> PRT
 <213> Sorghum bicolor

15 <400> 24
 Met Pro Pro Thr Pro Thr Thr Ala Ala Ala Thr Gly Ala Ala Val Ala
 1 5 10 15
 Ala Ala Ser Ala Glu Gln Ala Ala Phe Arg Leu Val Gly His Arg Asn
 20 25 30
 Phe Val Arg Val Asn Pro Arg Ser Asp Arg Phe His Thr Leu Ala Phe
 35 40 45
 His His Val Glu Leu Trp Cys Ala Asp Ala Ala Ser Ala Ala Gly Arg
 50 55 60
 Phe Ser Phe Gly Leu Gly Ala Pro Leu Ala Ala Arg Ser Asp Leu Ser
 65 70 75 80

ES 2 594 284 T3

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

ES 2 594 284 T3

Cys Gln Glu Leu Gly Val Leu Val Asp Arg Asp Asp Gln Gly Val Leu
 355 360 365

Leu Gln Ile Phe Thr Lys Pro Val Gly Asp Arg Pro Thr Leu Phe Leu
 370 375 380

Glu Ile Ile Gln Arg Ile Gly Cys Met Glu Lys Asp Glu Lys Gly Gln
 385 390 395 400

Glu Tyr Gln Lys Gly Gly Cys Gly Gly Phe Gly Lys Gly Asn Phe Ser
 405 410 415

Gln Leu Phe Lys Ser Ile Glu Asp Tyr Glu Lys Ser Leu Glu Ala Lys
 420 425 430

Gln Ala Ala Ala Ala Gln Gly Ser
 435 440

- <210> 25
- <211> 1311
- <212> ADN

5 <213> Triticum aestivum

<400> 25

```

atgccgccca cccccaccac ccccgcagcc accggcgccg gcgctgccgc cgcggtgacg      60
cgggagcag cgcggcccg cgaatggtc cgcttcaacc cgcgcagcga ccgcttccac      120
acgctctct tccaccacgt cgagttctgg tgcgcggacg ccgcctccgc cgcggccgc      180
ttcgccttcg cgctcggcgc gccgctcgcc gccaggtccg acctctccac ggggaactcc      240
gtgcacgcct ccagctgct ccgctcgggc aacctcgct tctctttcac cgcgccctac      300
gcaaacggct gcgacgccgc caccgcctcc ctgccctcct tctccgccga cgcgcgcgc      360
cggttctccg cggaccacgg gctcgcagtg cgctccatag cactgcgcgt cgcagacgcc      420
gcagaggcct tccgcgccag cgtcgcagga ggcgcgcgcc cggccttcag ccccgaggac      480
ctcggcccg gcttcggctt cgcggaggtc gagctctac gcgacgtcgt gctccgcttc      540
gtcagtcacc cggatgacac ggacgtgcc ttcttgccgg ggttcgaggg cgtgagcaac      600
ccggatgcc tggactacgg cctgacgcgg ttgcaccacg tcgtcggcaa cgtcccgag      660
cttgcccccg ccgccgcata cgtcgcgggg ttcgcggggg tccacgagtt cgcgcgagtt      720
acgacggagg acgtggggac ggccgagagc gggctcaact cgatggtgct cgcacaacaac      780
tcggagggcg tgctgctgcc gctcaacgag ccggtgcacg gcaccaagcg ccggagccag      840
atacagacgt tcttgaaca ccacggcggc tcgggcgtgc agcacatcgc ggtggccagc      900
agcgacgtgc tcaggacgt cagggagatg cgtgcgcgct ccgccatggg cggcttcgac      960
ttcctgccac cccgctgccg aaagtaactac gaaggcgtgc ggcgcacgc cggggatgtg     1020
ctctcggagg cgcagatcaa ggaatgccag gagctggggg tgctcgtcga cagggacgac     1080
caaggggtgt tgctacaaat cttcacaaag ccagtggggg acaggccaac gctgttctctg     1140
gagatgatcc aaaggatcgg gtgcattggag aaggacgaga gaggggaaga gtaccagaag     1200
ggtggctcgc gcgggttcgg caaaggcaac ttctccgagc tgttcaagtc cattgaagat     1260
tacgagaagt cccttgaagc caagcaatct gctgcagttc agggatcata g             1311
    
```

- 10 <210> 26
- <211> 436
- <212> PRT
- <213> Triticum aestivum

15 <400> 26

ES 2 594 284 T3

Met Pro Pro Thr Pro Thr Thr Pro Ala Ala Thr Gly Ala Gly Ala Ala
1 5 10 15

Ala Ala Val Thr Pro Glu His Ala Arg Pro Arg Arg Met Val Arg Phe
20 25 30

Asn Pro Arg Ser Asp Arg Phe His Thr Leu Ser Phe His His Val Glu
35 40 45

Phe Trp Cys Ala Asp Ala Ala Ser Ala Ala Gly Arg Phe Ala Phe Ala
50 55 60

Leu Gly Ala Pro Leu Ala Ala Arg Ser Asp Leu Ser Thr Gly Asn Ser
65 70 75 80

Val His Ala Ser Gln Leu Leu Arg Ser Gly Asn Leu Ala Phe Leu Phe
85 90 95

Thr Ala Pro Tyr Ala Asn Gly Cys Asp Ala Ala Thr Ala Ser Leu Pro
100 105 110

Ser Phe Ser Ala Asp Ala Ala Arg Arg Phe Ser Ala Asp His Gly Leu
115 120 125

Ala Val Arg Ser Ile Ala Leu Arg Val Ala Asp Ala Ala Glu Ala Phe
130 135 140

Arg Ala Ser Val Asp Gly Gly Ala Arg Pro Ala Phe Ser Pro Val Asp
145 150 155 160

Leu Gly Arg Gly Phe Gly Phe Ala Glu Val Glu Leu Tyr Gly Asp Val
165 170 175

Val Leu Arg Phe Val Ser His Pro Asp Asp Thr Asp Val Pro Phe Leu
180 185 190

Pro Gly Phe Glu Gly Val Ser Asn Pro Asp Ala Val Asp Tyr Gly Leu
195 200 205

Thr Arg Phe Asp His Val Val Gly Asn Val Pro Glu Leu Ala Pro Ala
210 215 220

Ala Ala Tyr Val Ala Gly Phe Ala Gly Phe His Glu Phe Ala Glu Phe
225 230 235 240

Thr Thr Glu Asp Val Gly Thr Ala Glu Ser Gly Leu Asn Ser Met Val
245 250 255

Leu Ala Asn Asn Ser Glu Gly Val Leu Leu Pro Leu Asn Glu Pro Val
260 265 270

ES 2 594 284 T3

His Gly Thr Lys Arg Arg Ser Gln Ile Gln Thr Phe Leu Glu His His
 275 280 285

Gly Gly Ser Gly Val Gln His Ile Ala Val Ala Ser Ser Asp Val Leu
 290 295 300

Arg Thr Leu Arg Glu Met Arg Ala Arg Ser Ala Met Gly Gly Phe Asp
 305 310 315 320

Phe Leu Pro Pro Arg Cys Arg Lys Tyr Tyr Glu Gly Val Arg Arg Ile
 325 330 335

Ala Gly Asp Val Leu Ser Glu Ala Gln Ile Lys Glu Cys Gln Glu Leu
 340 345 350

Gly Val Leu Val Asp Arg Asp Asp Gln Gly Val Leu Leu Gln Ile Phe
 355 360 365

Thr Lys Pro Val Gly Asp Arg Pro Thr Leu Phe Leu Glu Met Ile Gln
 370 375 380

Arg Ile Gly Cys Met Glu Lys Asp Glu Arg Gly Glu Glu Tyr Gln Lys
 385 390 395 400

Gly Gly Cys Gly Gly Phe Gly Lys Gly Asn Phe Ser Glu Leu Phe Lys
 405 410 415

Ser Ile Glu Asp Tyr Glu Lys Ser Leu Glu Ala Lys Gln Ser Ala Ala
 420 425 430

Val Gln Gly Ser
 435

<210> 27
 <211> 1305
 <212> ADN

5 <213> Hordeum vulgare

<400> 27

atgcgcgcca cccccaccac ccccgcggt accggcgccg ccgcccgggt gacgcccggag 60
 cacgcgcgac cgcaccgaat ggtccgcttc aaccgcgcga gcgaccgctt ccacacgctc 120
 tccttccacc acgtcgagtt ctggtgcgog gacgcccctt ccgcccggg ccgcttcgcg 180
 ttgcgctcgc gcgcgcccgt cgcgcaccag tccgacctct ccacggggaa ctccgcgcac 240
 gcctcccagc tgetccgctc gggctcccct gccttctctt tcaccgcgcc ctacgccaac 300
 ggotgcgacg ccgccaccgc ctccctgcc tctttctccg ccgacgccc gcgcgggttc 360
 tcgcgcgacc acgggatcgc ggtgcgctcc gtacgcgtgc gcgtcgcaga cgcgcgcgag 420
 gccttcgcgc ccagtcgctc acggggcgcg cgcgggctt tcgccccctt ggacctcggc 480
 cgcggettcg cgttcgcgga ggtcgagctc tacggcgacg tcgtgctccg cttcgtcagc 540
 caccgggacg gcacggacgt gcccttcttg ccggggttcg agggcgtaac caaccgggac 600
 gccgtggact acggcctgac gcggttcgac cacgtcgtcg gcaacgtccc ggagcttgcc 660
 ccgcgcgacg cctacatcgc cgggttcacg gggttccacg agttcgcgca gttcacggcg 720
 gaggacgtgg gcacgaccga gagcgggctc aactcgggtg tgetcgccaa caactcggag 780

ES 2 594 284 T3

```

ggcgtgctgc tgcgctcaa cgagccggtg cacggcacca agcgcgggag ccagatacag      840
acgttcctgg aacaccacgg cggccccggc gtgcagcaca tcgcggtggc cagcagtgac      900
gtgctcagga cgctcaggaa gatgcgtgcg cgctccgcca tgggcggcctt cgacttcctg      960
ccacccccgc tgcggaagta ctacgaaggc gtgcgacgcc ttgccgggga tgcctctctg    1020
gaggcgcaga tcaaggaatg ccaggagctg ggtgtgctcg tcgataggga cgaccaaggg      1080
gtgttgctcc aaatcttcac caagccagta ggggacaggc cgaccttgtt cctggagatg      1140
atccagagga tcgggtgcat ggagaaggac gagagagggg aagagtacca gaagggtggc      1200
tgcggcgggt tcggcaaagg caacttctcc gagctgttca agtccattga agattacgag      1260
aagtccttg  aagccaagca atctgctgca gttcagggat catag                          1305

```

<210> 28

<211> 434

5 <212> PRT

<213> *Hordeum vulgare*

<400> 28

```

Met Pro Pro Thr Pro Thr Thr Pro Ala Ala Thr Gly Ala Ala Ala Ala
1          5          10          15

```

```

Val Thr Pro Glu His Ala Arg Pro His Arg Met Val Arg Phe Asn Pro
          20          25          30

```

```

Arg Ser Asp Arg Phe His Thr Leu Ser Phe His His Val Glu Phe Trp
          35          40          45

```

```

Cys Ala Asp Ala Ala Ser Ala Ala Gly Arg Phe Ala Phe Ala Leu Gly
          50          55          60

```

```

Ala Pro Leu Ala Ala Arg Ser Asp Leu Ser Thr Gly Asn Ser Ala His
65          70          75          80

```

```

Ala Ser Gln Leu Leu Arg Ser Gly Ser Leu Ala Phe Leu Phe Thr Ala
          85          90          95

```

```

Pro Tyr Ala Asn Gly Cys Asp Ala Ala Thr Ala Ser Leu Pro Ser Phe
          100          105          110

```

```

Ser Ala Asp Ala Ala Arg Arg Phe Ser Ala Asp His Gly Ile Ala Val
          115          120          125

```

```

Arg Ser Val Ala Leu Arg Val Ala Asp Ala Ala Glu Ala Phe Arg Ala
          130          135          140

```

```

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

```

```

Arg Gly Phe Ala Phe Ala Glu Val Glu Leu Tyr Gly Asp Val Val Leu
          165          170          175

```

```

Arg Phe Val Ser His Pro Asp Gly Thr Asp Val Pro Phe Leu Pro Gly
          180          185          190

```

ES 2 594 284 T3

Phe Glu Gly Val Thr Asn Pro Asp Ala Val Asp Tyr Gly Leu Thr Arg
 195 200 205

Phe Asp His Val Val Gly Asn Val Pro Glu Leu Ala Pro Ala Ala Ala
 210 215 220

Tyr Ile Ala Gly Phe Thr Gly Phe His Glu Phe Ala Glu Phe Thr Ala
 225 230 235 240

Glu Asp Val Gly Thr Thr Glu Ser Gly Leu Asn Ser Val Val Leu Ala
 245 250 255

Asn Asn Ser Glu Gly Val Leu Leu Pro Leu Asn Glu Pro Val His Gly
 260 265 270

Thr Lys Arg Arg Ser Gln Ile Gln Thr Phe Leu Glu His His Gly Gly
 275 280 285

Pro Gly Val Gln His Ile Ala Val Ala Ser Ser Asp Val Leu Arg Thr
 290 295 300

Leu Arg Lys Met Arg Ala Arg Ser Ala Met Gly Gly Phe Asp Phe Leu
 305 310 315 320

Pro Pro Pro Leu Pro Lys Tyr Tyr Glu Gly Val Arg Arg Leu Ala Gly
 325 330 335

Asp Val Leu Ser Glu Ala Gln Ile Lys Glu Cys Gln Glu Leu Gly Val
 340 345 350

Leu Val Asp Arg Asp Asp Gln Gly Val Leu Leu Gln Ile Phe Thr Lys
 355 360 365

Pro Val Gly Asp Arg Pro Thr Leu Phe Leu Glu Met Ile Gln Arg Ile
 370 375 380

Gly Cys Met Glu Lys Asp Glu Arg Gly Glu Glu Tyr Gln Lys Gly Gly
 385 390 395 400

Cys Gly Gly Phe Gly Lys Gly Asn Phe Ser Glu Leu Phe Lys Ser Ile
 405 410 415

Glu Asp Tyr Glu Lys Ser Leu Glu Ala Lys Gln Ser Ala Ala Val Gln
 420 425 430

Gly Ser

<210> 29

<211> 1209

5 <212> ADN

<213> Rhodococcus sp.

ES 2 594 284 T3

<400> 29

```

atgactaccg ccgacattcg cctgacgccc cgcgaggtgg ccgcacatct ggagaccgac      60
gagctccggc agttggtcgg gctcgtcgaa cacgacgacg cgtcggatcc gtttcccgtg      120
gtcgcgatgg atgcctggtt gttcgtgtgc ggcaacgcga cgcagagcac gcagtacttc      180
gtctccacgt ggggcatgac cctcgtcgcc tacgccgggc cggagaccgg tcagcgctcg      240
cacaagtctt tegtctctga gtcggggctg gcacgggttc tgctgcacgg cgcctctgat      300
ccgaagagcc cgtcgcggga ccatcaccgg ggcacggcgg acggcgtggt ggacctggcg      360
atggaagttc tcgacgtcga ccgctgcac ggcgatgcac gctcgcaggg ggccaccatt      420
ctcgaggagc cgcgcgacgt cacggatcag ttccggcaccg tcgcgctcgc ggcgatcgcc      480
acgtacggca gcacccgga caccatcgtc gaccgaagcc gatacgacgg cccctacctc      540
cccggattcg tcgcgcgctc cagcggtttc gcggcgcgac cgggtaaacc cccgcgattg      600
ttccaggcgc tcgaccacgc cgtcggcaac gtcgagatgg gccggatgga tcaactgggtc      660
cggttctaca accgcgtcat gggcttcacg aacatggccg aattcgtcgg cgacgacatc      720
gccacggagt actcggcgtt gatgtcgaag gtcgtggcga acggcaatca ccgggtgaag      780
ttcccgtctc acgaaccgcg ggtgggaaag aagaagtctc agatcgacga atatctcgag      840
ttctacggtg agccgggctg ccagcatctg gccctcgcga ccggagacat cctcgcgacg      900
gtggacgcgt tgcgggcccga ggggtgtcga ttctgaaca cacccgacgc gtactacgag      960
gaccacagc tgcgcgcccg gatcggcagg gtgcgggtgc cggtgaggga actgcagaag     1020
cgcggaatcc tcgtcgaccg cgacgaggac ggatacctcc tgcagatctt caccaaaccg     1080
ctcggcgacc ggcgcgaccg gttcttcgag gtgatcgaac ggcacggttc gctcgggttc     1140
ggggcgggta acttcacggc cctgttcgaa tccatcgagc gtgagcaggc ggcgcgccc     1200
aatctgtga

```

<210> 30

5 <211> 402

<212> PRT

<213> Rhodococcus sp.

<400> 30

```

Met Thr Thr Ala Asp Ile Arg Leu Thr Pro Arg Glu Val Ala Ala His
1          5          10          15

Leu Glu Thr Asp Glu Leu Arg Gln Leu Val Gly Leu Val Glu His Asp
          20          25          30

Asp Ala Ser Asp Pro Phe Pro Val Val Ala Met Asp Ala Val Val Phe
          35          40          45

Val Cys Gly Asn Ala Thr Gln Ser Thr Gln Tyr Phe Val Ser Thr Trp
          50          55          60

Gly Met Thr Leu Val Ala Tyr Ala Gly Pro Glu Thr Gly Gln Arg Ser
65          70          75          80

His Lys Ser Phe Val Leu Glu Ser Gly Ser Ala Arg Phe Val Leu His
          85          90          95

Gly Ala Val Asp Pro Lys Ser Pro Leu Ala Asp His His Arg Ala His
          100          105          110

Gly Asp Gly Val Val Asp Leu Ala Met Glu Val Leu Asp Val Asp Arg

```

10

ES 2 594 284 T3

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

ES 2 594 284 T3

Asn Leu

<210> 31
 <211> 1353
 5 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> Secuencia de ácido nucleico que codifica HPPD de Arabidopsis y que contiene en el extremo 5' un ácido nucleico que codifica una Ala y 6 aminoácidos His
 10

 <220>
 <221> misc_feature
 <222> (4)..(6)
 15 <223> secuencia que codifica una Ala

 <220>
 <221> misc_feature
 <222> (7)..(24)
 20 <223> secuencia que codifica una cola de His que contiene 6 His

 <400> 31
 atggctcacc accatcacca tcacccaaac gccgcggtt cagagaatca aaaccatgat 60
 gacggcgctg cgtcgtcgc cggattcaag ctcgtcggat tttccaagtt cgtaaagaaag 120
 aatccaaagt ctgataaatt caaggttaag cgcttccatc acatcgagtt ctggtgcggc 180
 gacgcaacca acgtcgtcgc tcgcttctcc tggggtctgg ggatgagatt ctccgccaaa 240
 tcogattctt ccaccggaaa catggttccac gcctcttacc tactcacctc cggtgacctc 300
 cgattccttt tcaactgctc ttactctccg tctctctccg ccggagagat taaaccgaca 360
 accacagctt ctatoccaaag tttcgatcac ggctcttgtc gttccttctt ctcgtcacat 420
 ggtctcggtg ttagagccgt tgcgattgaa gtagaagacg cagagtcagc tttctccatc 480
 agtgtagcta atggcgctat tccttcgtcg cctcctatcg tcctcaatga agcagttacg 540
 atcgctgagg ttaaaactata cggcgatggt gttctccgat atgttagtta caaagcagaa 600
 gataccgaaa aatccgaatt cttgccaggg ttcgagcgtg tagaggatgc gtcgtcgttc 660
 ccattggatt atggtatccg gccgcttgac cacgccgtgg gaaacgttcc tgagcttggg 720
 ccggctttaa cttatgtagc ggggttccact ggttttcacc aattcgcaga gttcacagca 780
 gacgacgttg gaaccgccga gacgggttta aattcagcgg tcctggctag caatgatgaa 840
 atggttcttc taccgattaa cgagccagtg cacggaacaa agaggaagag tcagattcag 900
 acgtatttgg aacataacga aggcgcaggg ctacaacatc tggctctgat gagtgaagac 960
 atattcagga ccctgagaga gatgaggaag aggagcagta ttggaggatt cgacttcatg 1020
 ccttctctc cgcctactta ctaccagaat ctcaagaaac gggtcggcga cgtgctcagc 1080
 gatgatcaga tcaaggagtg tgaggaatta gggattcttg tagacagaga tgatcaaggg 1140
 acgttgcttc aaatcttcac aaaaccacta ggtgacaggc cgacgatatt tatagagata 1200
 atccagagag taggatgcat gatgaaagat gaggaaggga aggcctacca gagtggagga 1260
 tgtggtggtt ttggcaaagg caatttctct gagctcttca agtccattga agaatacga 1320
 aagactcttg aagccaaaca gttagtggga tga 1353

25 <210> 32
 <211> 450
 <212> PRT
 <213> Secuencia artificial
 <220>
 30 <223> codificado por la SEC ID N.º: 31

 <220>

ES 2 594 284 T3

<221> MISC_FEATURE

<222> (2)..(2)

<223> Ala

5 <220>

<221> MISC_FEATURE

<222> (3)..(8)

<223> cola de His constituída por 6 His

10

<400> 32

```

Met Ala His His His His His His Gln Asn Ala Ala Val Ser Glu Asn
1           5           10           15

Gln Asn His Asp Asp Gly Ala Ala Ser Ser Pro Gly Phe Lys Leu Val
          20           25           30

Gly Phe Ser Lys Phe Val Arg Lys Asn Pro Lys Ser Asp Lys Phe Lys
          35           40           45

Val Lys Arg Phe His His Ile Glu Phe Trp Cys Gly Asp Ala Thr Asn
          50           55           60

Val Ala Arg Arg Phe Ser Trp Gly Leu Gly Met Arg Phe Ser Ala Lys
65           70           75           80

Ser Asp Leu Ser Thr Gly Asn Met Val His Ala Ser Tyr Leu Leu Thr
          85           90           95

Ser Gly Asp Leu Arg Phe Leu Phe Thr Ala Pro Tyr Ser Pro Ser Leu
          100          105          110

Ser Ala Gly Glu Ile Lys Pro Thr Thr Thr Ala Ser Ile Pro Ser Phe
          115          120          125

Asp His Gly Ser Cys Arg Ser Phe Phe Ser Ser His Gly Leu Gly Val
          130          135          140

Arg Ala Val Ala Ile Glu Val Glu Asp Ala Glu Ser Ala Phe Ser Ile
145          150          155          160

Ser Val Ala Asn Gly Ala Ile Pro Ser Ser Pro Pro Ile Val Leu Asn
          165          170          175

Glu Ala Val Thr Ile Ala Glu Val Lys Leu Tyr Gly Asp Val Val Leu
          180          185          190

Arg Tyr Val Ser Tyr Lys Ala Glu Asp Thr Glu Lys Ser Glu Phe Leu

```

ES 2 594 284 T3

	195					200							205				
Pro	Gly	Phe	Glu	Arg	Val	Glu	Asp	Ala	Ser	Ser	Phe	Pro	Leu	Asp	Tyr		
	210					215					220						
Gly	Ile	Arg	Arg	Leu	Asp	His	Ala	Val	Gly	Asn	Val	Pro	Glu	Leu	Gly		
225					230				235						240		
Pro	Ala	Leu	Thr	Tyr	Val	Ala	Gly	Phe	Thr	Gly	Phe	His	Gln	Phe	Ala		
				245					250					255			
Glu	Phe	Thr	Ala	Asp	Asp	Val	Gly	Thr	Ala	Glu	Ser	Gly	Leu	Asn	Ser		
			260					265					270				
Ala	Val	Leu	Ala	Ser	Asn	Asp	Glu	Met	Val	Leu	Leu	Pro	Ile	Asn	Glu		
	275						280					285					
Pro	Val	His	Gly	Thr	Lys	Arg	Lys	Ser	Gln	Ile	Gln	Thr	Tyr	Leu	Glu		
	290					295					300						
His	Asn	Glu	Gly	Ala	Gly	Leu	Gln	His	Leu	Ala	Leu	Met	Ser	Glu	Asp		
305					310					315					320		
Ile	Phe	Arg	Thr	Leu	Arg	Glu	Met	Arg	Lys	Arg	Ser	Ser	Ile	Gly	Gly		
				325					330					335			
Phe	Asp	Phe	Met	Pro	Ser	Pro	Pro	Pro	Thr	Tyr	Tyr	Gln	Asn	Leu	Lys		
			340						345				350				
Lys	Arg	Val	Gly	Asp	Val	Leu	Ser	Asp	Asp	Gln	Ile	Lys	Glu	Cys	Glu		
		355					360					365					
Glu	Leu	Gly	Ile	Leu	Val	Asp	Arg	Asp	Asp	Gln	Gly	Thr	Leu	Leu	Gln		
	370					375					380						
Ile	Phe	Thr	Lys	Pro	Leu	Gly	Asp	Arg	Pro	Thr	Ile	Phe	Ile	Glu	Ile		
385					390					395					400		
Ile	Gln	Arg	Val	Gly	Cys	Met	Met	Lys	Asp	Glu	Glu	Gly	Lys	Ala	Tyr		
				405					410					415			
Gln	Ser	Gly	Gly	Cys	Gly	Gly	Phe	Gly	Lys	Gly	Asn	Phe	Ser	Glu	Leu		
			420					425					430				
Phe	Lys	Ser	Ile	Glu	Glu	Tyr	Glu	Lys	Thr	Leu	Glu	Ala	Lys	Gln	Leu		
		435					440					445					
Val	Gly																
	450																

<210> 33

<211> 1353

ES 2 594 284 T3

<212> ADN
 <213> Secuencia artificial

<220>
 5 <223> secuencia de nucleótidos que codifica el polipéptido HPPD mutante

<220>
 <221> misc_feature
 <222> (4)..(6)
 10 <223> secuencia que codifica Ala

<220>
 <221> misc_feature
 <222> (7)..(24)
 15 <223> secuencia que codifica una cola de His que contiene 6 His

<400> 33
 atggctcacc accatcacca tcacccaaaac gccgcggtt cagagaatca aaaccatgat 60
 gacggcgctg cgtcgtcgcc gggattcaag ctcgtcggat ttccaagtt cgtaagaaaag 120
 aatccaaagt ctgataaatt caagggttaag cgcctccatc acatcgagtt ctgggtcggc 180
 gacgcaacca acgtcgtcgtc tcgcttctcc tggggctcgg ggatgagatt ctccgccaaa 240
 tccgatcttt ccaccggaaa catgggttac gcctcttacc tactcacctc cggtgacctc 300
 cgatttcctt tcaactgctcc ttactctcgc tctctctcgc ccggagagat taaaccgaca 360
 accacagctt ctatcccaag ttccgatcac ggctcttgtc gttccttctt ctcgtcacat 420
 ggtctcgggt ttagagccgt tcgcatgtaa gtagaagacg cagagtcagc tttctccatc 480
 agtgtagcta atggcgctat tccttcgtcg cctcctatcg tcctcaatga agcagttacg 540
 atcgtcgagg ttaaaactata cggcgatggt gttctccgat atgttagtta caaagcagaa 600
 gataccgaaa aatccgaatt cttgccaggg ttccgagcgtg tagaggatgc gtcgtcgttc 660
 ccattggatt atgggtatcc gccgcttgac cacgccgtgg gaaacgttcc tgagcttggg 720
 ccggctttaa cttatgtagc ggggttcact ggttttcacc aattcgcaga gttcacagca 780
 gacgacgttg gaaccgcga gagcgggtta aatgccggcg tcctggctag caatgatgaa 840
 atggttcttc taccgattaa cgagccagtg cacggaacaa agaggaagag tcagattcag 900
 acgtatcttg aacataacga aggcgcaggg ctacaacatc tggctctgat gagtgaagac 960
 atattcagga cctcgagaga gatgaggaag aggagcagta ttggaggatt cgacttcatg 1020
 ccttctctcc cgcctactta ctaccagaat ctcaagaaac gggtcggcga cgtgctcagc 1080
 gatgatcaga tcaaggagtg tgaggaatta gggattcttg tagacagaga tgatcaaggg 1140
 acgttgcttc aaatcttcac aaaaccacta ggtgacaggg cgacgatatt tatagagata 1200
 atccagagag taggatgcat gatgaaagat gaggaagggg aggcctacca gagtggagga 1260
 tgtgggtggt ttggcaaagg caatttctct gagctcttca agtccattga agaatacga 1320
 aagactcttg aagccaaaca gttagtggga tga 1353

20 <210> 34
 <211> 450
 <212> PRT
 <213> Secuencia artificial

25 <220>
 <223> codificado por la SEC ID N.º: 33

<220>
 <221> MISC_FEATURE
 30 <222> (2)..(2)
 <223> Ala

<220>

ES 2 594 284 T3

<221> MISC_FEATURE

<222> (3)..(8)

<223> cola de His constituída por 6 His

5 <400> 34

```

Met Ala His His His His His His Gln Asn Ala Ala Val Ser Glu Asn
1          5          10          15

Gln Asn His Asp Asp Gly Ala Ala Ser Ser Pro Gly Phe Lys Leu Val
          20          25          30

Gly Phe Ser Lys Phe Val Arg Lys Asn Pro Lys Ser Asp Lys Phe Lys
          35          40          45

Val Lys Arg Phe His His Ile Glu Phe Trp Cys Gly Asp Ala Thr Asn
          50          55          60

Val Ala Arg Arg Phe Ser Trp Gly Leu Gly Met Arg Phe Ser Ala Lys
65          70          75          80

Ser Asp Leu Ser Thr Gly Asn Met Val His Ala Ser Tyr Leu Leu Thr
          85          90          95

Ser Gly Asp Leu Arg Phe Leu Phe Thr Ala Pro Tyr Ser Pro Ser Leu
          100          105          110

Ser Ala Gly Glu Ile Lys Pro Thr Thr Thr Ala Ser Ile Pro Ser Phe
          115          120          125

Asp His Gly Ser Cys Arg Ser Phe Phe Ser Ser His Gly Leu Gly Val
          130          135          140

Arg Ala Val Ala Ile Glu Val Glu Asp Ala Glu Ser Ala Phe Ser Ile
145          150          155          160

Ser Val Ala Asn Gly Ala Ile Pro Ser Ser Pro Pro Ile Val Leu Asn
          165          170          175

Glu Ala Val Thr Ile Ala Glu Val Lys Leu Tyr Gly Asp Val Val Leu
          180          185          190

Arg Tyr Val Ser Tyr Lys Ala Glu Asp Thr Glu Lys Ser Glu Phe Leu
          195          200          205

Pro Gly Phe Glu Arg Val Glu Asp Ala Ser Ser Phe Pro Leu Asp Tyr
          210          215          220

Gly Ile Arg Arg Leu Asp His Ala Val Gly Asn Val Pro Glu Leu Gly
225          230          235          240

```

ES 2 594 284 T3

Pro Ala Leu Thr Tyr Val Ala Gly Phe Thr Gly Phe His Gln Phe Ala
 245 250 255

Glu Phe Thr Ala Asp Asp Val Gly Thr Ala Glu Ser Gly Leu Asn Ala
 260 265 270

Ala Val Leu Ala Ser Asn Asp Glu Met Val Leu Leu Pro Ile Asn Glu
 275 280 285

Pro Val His Gly Thr Lys Arg Lys Ser Gln Ile Gln Thr Tyr Leu Glu
 290 295 300

His Asn Glu Gly Ala Gly Leu Gln His Leu Ala Leu Met Ser Glu Asp
 305 310 315 320

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

Phe Asp Phe Met Pro Ser Pro Pro Pro Thr Tyr Tyr Gln Asn Leu Lys
 340 345 350

Lys Arg Val Gly Asp Val Leu Ser Asp Asp Gln Ile Lys Glu Cys Glu
 355 360 365

Glu Leu Gly Ile Leu Val Asp Arg Asp Asp Gln Gly Thr Leu Leu Gln
 370 375 380

Ile Phe Thr Lys Pro Leu Gly Asp Arg Pro Thr Ile Phe Ile Glu Ile
 385 390 395 400

Ile Gln Arg Val Gly Cys Met Met Lys Asp Glu Glu Gly Lys Ala Tyr
 405 410 415

Gln Ser Gly Gly Cys Gly Gly Phe Gly Lys Gly Asn Phe Ser Glu Leu
 420 425 430

Phe Lys Ser Ile Glu Glu Tyr Glu Lys Thr Leu Glu Ala Lys Gln Leu
 435 440 445

Val Gly
 450

- <210> 35
- <211> 1353
- 5 <212> ADN
- <213> Secuencia artificial
- <220>
- <223> secuencia de nucleótidos que codifica polipéptido HPPD mutante
- 10 <220>
- <221> misc_feature
- <222> (4)..(6)

ES 2 594 284 T3

<223> secuencia que codifica Ala

<220>

<221> misc_feature

5 <222> (7)..(24)

<223> secuencia que codifica una cola de His que contiene 6 His

<400> 35

atggctc	accatcacca	tcacccaaaac	gocgcggtt	cagagaaatca	aaaccatgat	60
gacggcgctg	cgctcgtcgcc	gggattcaag	ctcgtcggat	tttccaagtt	cgtaagaaaag	120
aatccaaagt	ctgataaatt	caagggttaag	cgcttccatc	acatcgagtt	ctggtgcggc	180
gacgcaacca	acgtcgctcg	tcgcttctcc	tggggctctg	ggatgagatt	ctccgccaaa	240
tccgatcttt	ccaccggaaa	catggttcac	gcctcttacc	tactcacctc	cggtgacctc	300
cgattccttt	tcaactgctcc	ttactctccg	tctctctccg	ccggagagat	taaaccgaca	360
accacagctt	ctatcccaag	tttcgatcac	ggctcttgtc	gttctctctt	ctcgtcacat	420
ggtctcgggtg	ttagagccgt	tgcgattgaa	gtagaagacg	cagagtcagc	tttctccatc	480
agtgtagcta	atggcgctat	tccttcgctg	cctcctatcg	tcctcaatga	agcagttacg	540
atcgctgagg	ttaaactata	cggcgatggt	gttctccgat	atgttagtta	caaagcagaa	600
gataccgaaa	aatccgaatt	cttgccaggg	ttcgagcgtg	tagaggatgc	gtcgtcgttc	660
ccattggatt	atggtatccg	gocgcttgac	cacgcgctgg	gaaacgttcc	tgagcttggg	720
ccggctttaa	cttatgtagc	ggggttcact	ggttttcacc	aattcgcaga	gttcacagca	780
gacgacgttg	gaaccgccga	gagcggttta	aattcagcgg	tcctggctag	caatgatgaa	840
atggttcttc	taccgattgc	cgagccagtg	cacggaacaa	agaggaagag	tcagattcag	900
acgtatcttg	aacataacga	aggcgcaggg	ctacaacatc	tggctctgat	gagtgaagac	960
atattcagga	ccctgagaga	gatgagggaag	aggagcagta	ttggaggatt	cgacttcatg	1020
ccttctcttc	cgctacttta	ctaccagaat	ctcaagaaac	gggtcggcga	cgtgctcagc	1080
gatgatcaga	tcaaggagtg	tgaggaatta	gggattcttg	tagacagaga	tgatcaaggg	1140
acgttgcttc	aaatcttcac	aaaaccacta	ggtgacaggc	cgacgatatt	tatagagata	1200
atccagagag	taggatgcat	gatgaaagat	gaggaagggg	aggcttacca	gagtggagga	1260
tgtggtggtt	ttggcaaagg	caatttctct	gagctcttca	agtccattga	agaatacгаа	1320
aagactcttg	aagccaaaca	gttagtggga	tga			1353

10

<210> 36

<211> 450

<212> PRT

<213> Secuencia artificial

15

<220>

<223> codificada por la SEC ID N.º: 35

<220>

20 <221> MISC_FEATURE

<222> (2)..(2)

<223> Ala

<220>

25 <221> MISC_FEATURE

<222> (3)..(8)

<223> cola de His constituida por 6 His

<400> 36

ES 2 594 284 T3

Met Ala His His His His His His Gln Asn Ala Ala Val Ser Glu Asn
1 5 10 15

Gln Asn His Asp Asp Gly Ala Ala Ser Ser Pro Gly Phe Lys Leu Val
20 25 30

Gly Phe Ser Lys Phe Val Arg Lys Asn Pro Lys Ser Asp Lys Phe Lys
35 40 45

Val Lys Arg Phe His His Ile Glu Phe Trp Cys Gly Asp Ala Thr Asn
50 55 60

Val Ala Arg Arg Phe Ser Trp Gly Leu Gly Met Arg Phe Ser Ala Lys
65 70 75 80

Ser Asp Leu Ser Thr Gly Asn Met Val His Ala Ser Tyr Leu Leu Thr
85 90 95

Ser Gly Asp Leu Arg Phe Leu Phe Thr Ala Pro Tyr Ser Pro Ser Leu
100 105 110

Ser Ala Gly Glu Ile Lys Pro Thr Thr Thr Ala Ser Ile Pro Ser Phe
115 120 125

Asp His Gly Ser Cys Arg Ser Phe Phe Ser Ser His Gly Leu Gly Val
130 135 140

Arg Ala Val Ala Ile Glu Val Glu Asp Ala Glu Ser Ala Phe Ser Ile
145 150 155 160

Ser Val Ala Asn Gly Ala Ile Pro Ser Ser Pro Pro Ile Val Leu Asn
165 170 175

Glu Ala Val Thr Ile Ala Glu Val Lys Leu Tyr Gly Asp Val Val Leu
180 185 190

Arg Tyr Val Ser Tyr Lys Ala Glu Asp Thr Glu Lys Ser Glu Phe Leu
195 200 205

Pro Gly Phe Glu Arg Val Glu Asp Ala Ser Ser Phe Pro Leu Asp Tyr
210 215 220

Gly Ile Arg Arg Leu Asp His Ala Val Gly Asn Val Pro Glu Leu Gly
225 230 235 240

Pro Ala Leu Thr Tyr Val Ala Gly Phe Thr Gly Phe His Gln Phe Ala
245 250 255

Glu Phe Thr Ala Asp Asp Val Gly Thr Ala Glu Ser Gly Leu Asn Ser
260 265 270

ES 2 594 284 T3

Ala Val Leu Ala Ser Asn Asp Glu Met Val Leu Leu Pro Ile Ala Glu
 275 280 285

Pro Val His Gly Thr Lys Arg Lys Ser Gln Ile Gln Thr Tyr Leu Glu
 290 295 300

His Asn Glu Gly Ala Gly Leu Gln His Leu Ala Leu Met Ser Glu Asp
 305 310 315 320

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

Phe Asp Phe Met Pro Ser Pro Pro Pro Thr Tyr Tyr Gln Asn Leu Lys
 340 345 350

Lys Arg Val Gly Asp Val Leu Ser Asp Asp Gln Ile Lys Glu Cys Glu
 355 360 365

Glu Leu Gly Ile Leu Val Asp Arg Asp Asp Gln Gly Thr Leu Leu Gln
 370 375 380

Ile Phe Thr Lys Pro Leu Gly Asp Arg Pro Thr Ile Phe Ile Glu Ile
 385 390 395 400

Ile Gln Arg Val Gly Cys Met Met Lys Asp Glu Glu Gly Lys Ala Tyr
 405 410 415

Gln Ser Gly Gly Cys Gly Gly Phe Gly Lys Gly Asn Phe Ser Glu Leu
 420 425 430

Phe Lys Ser Ile Glu Glu Tyr Glu Lys Thr Leu Glu Ala Lys Gln Leu
 435 440 445

Val Gly
 450

- <210> 37
- <211> 1353
- 5 <212> ADN
- <213> Secuencia artificial

- <220>
- 10 <223> secuencia de nucleótidos que codifica polipéptido HPPD mutante

- <220>
- <221> misc_feature
- <222> (4)..(6)
- <223> secuencia que codifica Ala

- 15 <220>
- <221> misc_feature
- <222> (7)..(24)
- <223> secuencia que codifica una cola de His que contiene 6 His

- 20

ES 2 594 284 T3

<400> 37

```

atggctcacc accatcacca tcacccaaaac gccgcgcttt cagagaatca aaacccatgat      60
gacggcgctg cgtcgctgcc gggattcaag ctcgctcggat tttccaagtt cgtaagaaaag    120
aatccaaagt ctgataaatt caaggttaag cgcttccatc acatcgagtt ctgggtgcggc    180
gacgcaacca acgtcgctcg tcgctttctcc tggggctctgg ggatgagatt ctccgccaaa    240
tccgatcttt ccaccggaaa catggttcac gcctcttacc tactcacctc cggtgacctc    300
cgattccttt tcaactgctc ttactctccg tctctctccg ccggagagat taaaccgaca    360
accacagctt ctatcccaag tttcgatcac ggctcttgtc gttccttctt ctcgtcacat    420
ggctctgggtg ttagagccgt tgcgattgaa gtagaagacg cagagtcagc tttctccatc    480
agtgtagcta atggcgctat tccttcgctc cctcctatcg tcctcaatga agcagttacg    540
atcgctgagg ttaaaactata cggcgatggt gttctccgat atgttagtta caaagcagaa    600
gataccgaaa aatccgaatt cttgccaggg ttcgagcgtg tagaggatgc gtcgctcgttc    660
ccattggatt atggtatccg gcggcttgac cagcccgctgg gaaacgttcc tgagcttggg    720
ccggctttaa cttatgtagc ggggttcaact ggttttcacc aattcgcatc tttcacagca    780
gacgacgttg gaaccgccga gagcggttta aattcagcgg tcctggctag caatgatgaa    840
atggttcttc taccgattaa cgagccagtg cacggaacaa agaggaagag tcagattcag    900
acgtatcttg aacataacga aggcgcaggg ctacaacatc tggctctgat gagtgaagac    960
atattcagga cctgagaga gatgaggaag aggagcagta ttggaggatt cgacttcatg   1020
ccttctcctc cgcctactta ctaccagaat ctcaagaaac gggtcggcga cgtgctcagc   1080
gatgatcaga tcaaggagtg tgaggaatta gggattcttg tagacagaga tgatcaaggg   1140
acgttgcttc aaatcttcac aaaaccacta ggtgacaggc cgacgatatt tatagagata   1200
atccagagag taggatgcat gatgaaagat gaggaaggga aggcttacca gagtggagga   1260
tgtggtggtt ttggcaaagg caatttctct gagctcttca agtccattga agaatacгаа   1320
aagactcttg aagccaaaca gttagtggga tga                                     1353

```

<210> 38

5 <211> 450

<212> PRT

<213> Secuencia artificial

<220>

10 <223> codificada por la SEC ID N.º: 37

<220>

<221> MISC_FEATURE

<222> (2)..(2)

15 <223> Ala

<220>

<221> MISC_FEATURE

<222> (3)..(8)

20 <223> cola de His constituida por 6 His

<400> 38

```

Met Ala His His His His His His Gln Asn Ala Ala Val Ser Glu Asn
  1                   5                   10                   15

Gln Asn His Asp Asp Gly Ala Ala Ser Ser Pro Gly Phe Lys Leu Val
                   20                   25                   30

```

ES 2 594 284 T3

Gly Phe Ser Lys Phe Val Arg Lys Asn Pro Lys Ser Asp Lys Phe Lys
35 40 45

Val Lys Arg Phe His His Ile Glu Phe Trp Cys Gly Asp Ala Thr Asn
50 55 60

Val Ala Arg Arg Phe Ser Trp Gly Leu Gly Met Arg Phe Ser Ala Lys
65 70 75 80

Ser Asp Leu Ser Thr Gly Asn Met Val His Ala Ser Tyr Leu Leu Thr
85 90 95

Ser Gly Asp Leu Arg Phe Leu Phe Thr Ala Pro Tyr Ser Pro Ser Leu
100 105 110

Ser Ala Gly Glu Ile Lys Pro Thr Thr Thr Ala Ser Ile Pro Ser Phe
115 120 125

Asp His Gly Ser Cys Arg Ser Phe Phe Ser Ser His Gly Leu Gly Val
130 135 140

Arg Ala Val Ala Ile Glu Val Glu Asp Ala Glu Ser Ala Phe Ser Ile
145 150 155 160

Ser Val Ala Asn Gly Ala Ile Pro Ser Ser Pro Pro Ile Val Leu Asn
165 170 175

Glu Ala Val Thr Ile Ala Glu Val Lys Leu Tyr Gly Asp Val Val Leu
180 185 190

Arg Tyr Val Ser Tyr Lys Ala Glu Asp Thr Glu Lys Ser Glu Phe Leu
195 200 205

Pro Gly Phe Glu Arg Val Glu Asp Ala Ser Ser Phe Pro Leu Asp Tyr
210 215 220

Gly Ile Arg Arg Leu Asp His Ala Val Gly Asn Val Pro Glu Leu Gly
225 230 235 240

Pro Ala Leu Thr Tyr Val Ala Gly Phe Thr Gly Phe His Gln Phe Ala
245 250 255

Ser Phe Thr Ala Asp Asp Val Gly Thr Ala Glu Ser Gly Leu Asn Ser
260 265 270

Ala Val Leu Ala Ser Asn Asp Glu Met Val Leu Leu Pro Ile Asn Glu
275 280 285

Pro Val His Gly Thr Lys Arg Lys Ser Gln Ile Gln Thr Tyr Leu Glu
290 295 300

His Asn Glu Gly Ala Gly Leu Gln His Leu Ala Leu Met Ser Glu Asp

ES 2 594 284 T3

305					310					315					320
Ile	Phe	Arg	Thr	Leu	Arg	Glu	Met	Arg	Lys	Arg	Ser	Ser	Ile	Gly	Gly
				325					330					335	
Phe	Asp	Phe	Met	Pro	Ser	Pro	Pro	Pro	Thr	Tyr	Tyr	Gln	Asn	Leu	Lys
			340					345					350		
Lys	Arg	Val	Gly	Asp	Val	Leu	Ser	Asp	Asp	Gln	Ile	Lys	Glu	Cys	Glu
		355					360					365			
Glu	Leu	Gly	Ile	Leu	Val	Asp	Arg	Asp	Asp	Gln	Gly	Thr	Leu	Leu	Gln
	370					375					380				
Ile	Phe	Thr	Lys	Pro	Leu	Gly	Asp	Arg	Pro	Thr	Ile	Phe	Ile	Glu	Ile
385					390					395					400
Ile	Gln	Arg	Val	Gly	Cys	Met	Met	Lys	Asp	Glu	Glu	Gly	Lys	Ala	Tyr
				405					410					415	
Gln	Ser	Gly	Gly	Cys	Gly	Gly	Phe	Gly	Lys	Gly	Asn	Phe	Ser	Glu	Leu
			420					425					430		
Phe	Lys	Ser	Ile	Glu	Glu	Tyr	Glu	Lys	Thr	Leu	Glu	Ala	Lys	Gln	Leu
		435					440					445			
Val	Gly														
	450														

<210> 39
 <211> 1353
 5 <212> ADN
 <213> Secuencia artificial

<220>
 <223> secuencia de nucleótidos que codifica polipéptido HPPD mutante
 10

<220>
 <221> misc_feature
 <222> (4)..(6)
 <223> secuencia que codifica Ala
 15

<220>
 <221> misc_feature
 <222> (7)..(24)
 <223> secuencia que codifica una cola de His que contiene 6 His
 20

<400> 39
 atggctcacc accatcacca tcacccaaaac gccgcggttt cagagaatca aaaccatgat 60
 gacggcgctg cgctcgtcgcc gggattcaag ctcgctcgat ttccaagtt cgtaagaaag 120
 aatccaaagt ctgataaatt caaggttaag cgcttcacac acatcgagtt ctgggtgcggc 180
 gacgcaacca acgtcgctcg tcgcttctcc tggggctctgg ggatgagatt ctccgccaaa 240

ES 2 594 284 T3

```
tccgatcttt ccaccgaaa catggttac gccctttacc tactcacctc cggtgacctc 300
cgattccttt tcaactgctc ttactctcog tctctctcog cgggagagat taaaccgaca 360
accacagctt ctatcccaag tttegateac ggctcttgtc gttccttctt ctogtcacat 420
ggtctcggtg ttagagccgt tgcgattgaa gtagaagacg cagagtcagc tttctccatc 480
agtgtagcta atggcgctat tccttcgctg cctcctatcg tcctcaatga agcagttaog 540
atcgctgagg ttaaactata cggcgatggt gttctccgat atgttagtta caaagcagaa 600
gataccgaaa aatccgaatt cttgccaggg ttcgagcgtg tagaggatgc gtcgctgctc 660
ccattggatt atggtatccg gcggcttgac cacgccgtgg gaaacgttcc tgagcttggg 720
ccggctttaa cttatgtagc ggggttccact ggttttcacc aattcgcaga gttcacagca 780
gacgacgttg gaaccgccga gagcggttta aattcagcgg cgctggctag caatgatgaa 840
atggttcttc taccgattaa cgagccagtg cacggaacaa agaggaagag tcagattcag 900
acgtatthgg aacataacga aggcgcaggg ctacaacatc tggtctgat gagtgaagac 960
atattcagga ccctgagaga gatgaggaag aggagcagta ttggaggatt cgacttcatg 1020
ccttctctc cgcctactta ctaccagaat ctcaagaaac gggtcggcga cgtgctcagc 1080
gatgatcaga tcaaggagtg tgaggaatta gggattcttg tagacagaga tgatcaaggg 1140
acgttgcttc aaatcttcac aaaaccacta ggtgacaggc cgacgatatt tatagagata 1200
atccagagag taggatgcat gatgaaagat gaggaagggg aggcttacca gagtggagga 1260
tgtggtggtt ttggcaaagg caatttctct gagctcttca agtcattga agaatacгаа 1320
aagactcttg aagccaaaca gttagtggga tga 1353
```

<210> 40
 <211> 450
 5 <212> PRT
 <213> Secuencia artificial

<220>
 <223> codificada por la SEC ID N.º: 39

10 <220>
 <221> MISC_FEATURE
 <222> (2)..(2)
 <223> Ala

15 <220>
 <221> MISC_FEATURE
 <222> (3)..(8)
 <223> cola de His constituida por 6 His

20 <400> 40

```
Met Ala His His His His His His Gln Asn Ala Ala Val Ser Glu Asn
1                5                10                15
```

```
Gln Asn His Asp Asp Gly Ala Ala Ser Ser Pro Gly Phe Lys Leu Val
                20                25                30
```

```
Gly Phe Ser Lys Phe Val Arg Lys Asn Pro Lys Ser Asp Lys Phe Lys
                35                40                45
```

```
Val Lys Arg Phe His His Ile Glu Phe Trp Cys Gly Asp Ala Thr Asn
                50                55                60
```

```
Val Ala Arg Arg Phe Ser Trp Gly Leu Gly Met Arg Phe Ser Ala Lys
```

ES 2 594 284 T3

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

ES 2 594 284 T3

Lys Arg Val Gly Asp Val Leu Ser Asp Asp Gln Ile Lys Glu Cys Glu
 355 360 365

Glu Leu Gly Ile Leu Val Asp Arg Asp Asp Gln Gly Thr Leu Leu Gln
 370 375 380

Ile Phe Thr Lys Pro Leu Gly Asp Arg Pro Thr Ile Phe Ile Glu Ile
 385 390 395 400

Ile Gln Arg Val Gly Cys Met Met Lys Asp Glu Glu Gly Lys Ala Tyr
 405 410 415

Gln Ser Gly Gly Cys Gly Gly Phe Gly Lys Gly Asn Phe Ser Glu Leu
 420 425 430

Phe Lys Ser Ile Glu Glu Tyr Glu Lys Thr Leu Glu Ala Lys Gln Leu
 435 440 445

Val Gly
 450

- <210> 41
- <211> 1353
- 5 <212> ADN
- <213> Secuencia artificial
- <220>
- <223> secuencia de nucleótidos que codifica polipéptido HPPD mutante
- 10 <220>
- <221> misc_feature
- <222> (4)..(6)
- <223> secuencia que codifica Ala
- 15 <220>
- <221> misc_feature
- <222> (7)..(24)
- <223> secuencia que codifica una cola de His que contiene 6 His

20 <400> 41

atggctcatc	accatcacca	tcaccaaacc	gocgcggttt	cagagaatca	aaaccatgat	60
gacggcgcgtg	cgtcgctgcc	gggattcaag	ctcgtcggat	ttccaagtt	cgtaagaaag	120
aatccaaagt	ctgataaatt	caaggttaag	cgcttccatc	acatcgagtt	ctgggtgcggc	180
gacgcaacca	acgtcgtctg	tcgcttctcc	tggggctctgg	ggatgagatt	ctccgcaaaa	240
tccgatcttt	ccaccggaaa	catggttcac	gcctcttacc	tactcacctc	cggtgacctc	300
cgattccttt	tcaactgctc	ttactctccg	tctctctccg	ccggagagat	taaaccgaca	360
accacagctt	ctatcccaag	tttcgatcac	ggctcttgtc	gttccttctt	ctcgtcacat	420
ggtctcgggtg	ttagagccgt	tgcgattgaa	gtagaagacg	cagagtcagc	tttctccatc	480
agtgtagcta	atggcgctat	tccttcgctg	cctcctatcg	tcctcaatga	agcagttacg	540
atcgctgagg	ttaaactata	cggcgatggt	gttctccgat	atgtagtta	caaagcagaa	600
gataccgaaa	aatccgaatt	cttgccaggg	ttcgagcgtg	tagaggatgc	gtcgtcgttc	660
ccattggatt	atggtatccg	gocgcttgac	cacgcgctgg	gaaacgttcc	tgagcttggg	720

ES 2 594 284 T3

```

ccggctttaa cttatgtagc ggggttcact ggttttcacc aattcgcaga gttcacagca 780
gacgacgttg gaaccgccga gagcggttta aattcagcgg tcttggctag caatgatgaa 840
atggttcttc tagcgattaa cgagccagtg cacggaacaa agaggaagag tcagattcag 900
acgtatttgg aacataacga aggcgcaggg ctacaacatc tggctctgat gagtgaagac 960
atattcagga ccoctgagaga gatgaggaag aggagcagta ttggaggatt cgacttcattg 1020
ccttctcttc cgctactta ctaccagaat ctcaagaaac gggtcggcga cgtgctcagc 1080
gatgatcaga tcaaggagtg tgaggaatta gggattcttg tagacagaga tgatcaaggg 1140
acgttgcttc aaatcttcac aaaaccacta ggtgacaggc cgacgatatt tatagagata 1200
atccagagag taggatgcat gatgaaagat gaggaagggg aggcctacca gagtggagga 1260
tgtgggtggtt ttggcaaagg caatttctct gagctcttca agtccattga agaatacгаа 1320
aagactcttg aagccaaaca gttagtggga tga 1353

```

- <210> 42
- <211> 450
- 5 <212> PRT
- <213> Secuencia artificial
- <220>
- <223> codificada por la SEC ID N.º: 41
- 10 <220>
- <221> MISC_FEATURE
- <222> (2)..(2)
- <223> Ala
- 15 <220>
- <221> MISC_FEATURE
- <222> (3)..(8)
- <223> cola de His constituida por 6 His
- 20 <400> 42

```

Met Ala His His His His His His Gln Asn Ala Ala Val Ser Glu Asn
1                5                10                15

Gln Asn His Asp Asp Gly Ala Ala Ser Ser Pro Gly Phe Lys Leu Val
                20                25                30

Gly Phe Ser Lys Phe Val Arg Lys Asn Pro Lys Ser Asp Lys Phe Lys
                35                40                45

Val Lys Arg Phe His His Ile Glu Phe Trp Cys Gly Asp Ala Thr Asn
50                55                60

Val Ala Arg Arg Phe Ser Trp Gly Leu Gly Met Arg Phe Ser Ala Lys
65                70                75                80

Ser Asp Leu Ser Thr Gly Asn Met Val His Ala Ser Tyr Leu Leu Thr
                85                90                95

Ser Gly Asp Leu Arg Phe Leu Phe Thr Ala Pro Tyr Ser Pro Ser Leu
100                105                110

Ser Ala Gly Glu Ile Lys Pro Thr Thr Thr Ala Ser Ile Pro Ser Phe

```

ES 2 594 284 T3

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

ES 2 594 284 T3

Ile Gln Arg Val Gly Cys Met Met Lys Asp Glu Glu Gly Lys Ala Tyr
 405 410 415

Gln Ser Gly Gly Cys Gly Gly Phe Gly Lys Gly Asn Phe Ser Glu Leu
 420 425 430

Phe Lys Ser Ile Glu Glu Tyr Glu Lys Thr Leu Glu Ala Lys Gln Leu
 435 440 445

Val Gly
 450

5 <210> 43
 <211> 1353
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> secuencia de nucleótidos que codifica polipéptido HPPD mutante

15 <220>
 <221> misc_feature
 <222> (4)..(6)
 <223> secuencia que codifica Ala

20 <220>
 <221> misc_feature
 <222> (7)..(24)
 <223> secuencia que codifica una cola de His que contiene 6 His

<400> 43

atggtcctc	accatcacca	tcacccaaaac	gcccgcgttt	cagagaatca	aaaccatgat	60
gacggcgctg	cgctcgtcgc	gggattcaag	ctcgtcggat	ttccaagtt	cgtaagaaag	120
aatccaaagt	ctgataaatt	caaggttaag	cgcttccatc	acatcgagtt	ctggtgcggc	180
gacgcaacca	acgtcgtctg	tcgcttctcc	tggggctctg	ggatgagatt	ctccgccaaa	240
tccgatcttt	ccaccgaaa	catggttac	gcctcttacc	tactcacctc	cggtgacctc	300
cgattccttt	tactgctcc	ttactctccg	tctctctccg	cggagagat	taaaccgaca	360
accacagctt	ctatcccaag	tttcgatcac	ggctcttgtc	gttccttctt	ctcgtcacat	420
ggtctcggtg	ttagagccgt	tgcgattgaa	gtagaagacg	cagagtcagc	tttctccatc	480
agtgtagcta	atggcgctat	tccttcgtcg	cctcctatcg	tcctcaatga	agcagttacg	540
atcgtctgagg	ttaaaactata	cggcgatggt	gttctccgat	atgttagtta	caaagcagaa	600
gataccgaaa	aatccgaatt	cttgccaggg	ttcagagctg	tagaggatgc	gtcgtcgttc	660
ccattggatt	atggtatccg	gcccgttgac	cagcccgtag	gaaacgttcc	tgagcttggg	720
ccggctttaa	cttatgtagc	ggggttcact	ggttttcacc	aattcgcaga	gttcacagca	780
gacgacgttg	gaaccgcca	gagcggttta	aattcagcgg	tcctggctag	caatgatgaa	840
atggttcttc	taccgattaa	cgagccagtg	cacggaacaa	agaggaagag	tcagattcag	900
acgtatttgg	aacataacga	aggcgcaggg	ctacaacatc	tggctctgat	gagtgaagac	960
atattcagga	ccctgagaga	gatgaggaag	aggagcagta	ttggaggatt	cgacttcatg	1020
ccttctcctc	cgctactta	ctaccagaat	ctcaagaaac	gggtcggcga	cgtgctcagc	1080
gatgatcaga	tcaaggagtg	tgaggaatta	gggattatgg	tagacagaga	tgatcaaggg	1140
acgttgcttc	aaatcttcac	aaaaccacta	ggtgacaggc	cgacgatatt	tatagagata	1200

ES 2 594 284 T3

atccagagag taggatgcat gatgaaagat gaggaaggga aggcttacca gagtggagga 1260
 tgtggtggtt ttggcaaagc caatttctct gagctcttca agtccattga agaatacga 1320
 aagactcttg aagccaaaca gttagtggga tga 1353

- <210> 44
- <211> 450
- 5 <212> PRT
- <213> Secuencia artificial

- <220>
- <223> codificada por la SEC ID N.º: 43
- 10 <220>
- <221> MISC_FEATURE
- <222> (2)..(2)
- <223> Ala
- 15 <220>
- <221> MISC_FEATURE
- <222> (3)..(8)
- <223> cola de His constituida por 6 His
- 20 <400> 44

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

ES 2 594 284 T3

Ser Val Ala Asn Gly Ala Ile Pro Ser Ser Pro Pro Ile Val Leu Asn
165 170 175

Glu Ala Val Thr Ile Ala Glu Val Lys Leu Tyr Gly Asp Val Val Leu
180 185 190

Arg Tyr Val Ser Tyr Lys Ala Glu Asp Thr Glu Lys Ser Glu Phe Leu
195 200 205

Pro Gly Phe Glu Arg Val Glu Asp Ala Ser Ser Phe Pro Leu Asp Tyr
210 215 220

Gly Ile Arg Arg Leu Asp His Ala Val Gly Asn Val Pro Glu Leu Gly
225 230 235 240

Pro Ala Leu Thr Tyr Val Ala Gly Phe Thr Gly Phe His Gln Phe Ala
245 250 255

Glu Phe Thr Ala Asp Asp Val Gly Thr Ala Glu Ser Gly Leu Asn Ser
260 265 270

Ala Val Leu Ala Ser Asn Asp Glu Met Val Leu Leu Pro Ile Asn Glu
275 280 285

Pro Val His Gly Thr Lys Arg Lys Ser Gln Ile Gln Thr Tyr Leu Glu
290 295 300

His Asn Glu Gly Ala Gly Leu Gln His Leu Ala Leu Met Ser Glu Asp
305 310 315 320

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

Phe Asp Phe Met Pro Ser Pro Pro Pro Thr Tyr Tyr Gln Asn Leu Lys
340 345 350

Lys Arg Val Gly Asp Val Leu Ser Asp Asp Gln Ile Lys Glu Cys Glu
355 360 365

Glu Leu Gly Ile Met Val Asp Arg Asp Asp Gln Gly Thr Leu Leu Gln
370 375 380

Ile Phe Thr Lys Pro Leu Gly Asp Arg Pro Thr Ile Phe Ile Glu Ile
385 390 395 400

Ile Gln Arg Val Gly Cys Met Met Lys Asp Glu Glu Gly Lys Ala Tyr
405 410 415

Gln Ser Gly Gly Cys Gly Gly Phe Gly Lys Gly Asn Phe Ser Glu Leu
420 425 430

ES 2 594 284 T3

Phe Lys Ser Ile Glu Glu Tyr Glu Lys Thr Leu Glu Ala Lys Gln Leu
 435 440 445

Val Gly
 450

<210> 45
 <211> 1353
 5 <212> ADN
 <213> Secuencia artificial

<220>
 <223> secuencia de nucleótidos que codifica polipéptido HPPD mutante

10 <220>
 <221> misc_feature
 <222> (4)..(6)
 <223> secuencia que codifica Ala

15 <220>
 <221> misc_feature
 <222> (7)..(24)
 <223> secuencia que codifica una cola de His que contiene 6 His

20 <400> 45
 atgggtcattc accatcacca tcacccaaaac gccgcggttt cagagaatca aaacctatgat 60
 gacggcgctg cgtcgtcgcc gggattcaag ctcgtcggat tttccaagtt cgtaagaaaag 120
 aatccaaagt ctgataaatt caaggttaag cgcttccatc acatcgagtt ctgggtgcggc 180
 gacgcaacca acgtcgtctg tcgcttctcc tggggctctgg ggatgagatt ctccgccaaa 240
 tccgatcttt ccaccggaaa catggttcac gcctcttacc tactcacctc cggtgacctc 300
 cgattccttt tcaactgctc ttactctccg tctctctccg ccggagagat taaaccgaca 360
 accacagctt ctatcccaag tttcgatcac ggctcttgtc gttccttctt ctcgtcacat 420
 ggtctcggtg ttagagccgt tgcgattgaa gtagaagacg cagagtcagc tttctccatc 480
 agtgtagcta atggcgctat tccttcgtcg cctcctatcg tcctcaatga agcagttacg 540
 atcgctgagg ttaaactata cggcgatggt gttctccgat atgttagtta caaagcagaa 600
 gataaccgaaa aatccgaatt cttgccaggg ttcgagcgtg tagaggatgc gtcgtcgttc 660
 ccattggatt atggatccg gccgcttgac cacgcctggg gaaacgttcc tgagcttggg 720
 ccggctttaa cttatgtagc ggggttccact ggttttcacc aattcgcaga gttcacagca 780
 gacgacgttg gaaccgccga gagcggttta aattcagcgg tcctggctag caatgatgaa 840
 atggttcttc taccgattaa cgagccagtg cacggaaca agaggaagag tcagattcag 900
 acgtatttgg aacataacga aggcgcaggg ctacaacatc tggctctgat gagtgaagac 960
 atattcagga ccctgagaga gatgaggaag aggagcagta ttggaggatt cgacttcatg 1020
 ccttctctc cgctactta ctaccagaat ctcaagaaac gggtcggcga cgtgctcagc 1080
 gatgatcaga tcaaggagtg tgaggaatta gggattcttg tagacagaga tgatcaaggg 1140
 acgttgcttc aaatcttcac aaaaccacta ggtgacaggg cgacgatatt tatagagata 1200
 atccagagag taggatgcat gatgaaagat gaggaagga aggcctacca gagtggagga 1260
 tgtgtgggtt ttggcattgg caatttctct gagctcttca agtccattga agaatacгаа 1320
 aagactcttg aagccaaaca gttagtggga tga 1353

<210> 46
 25 <211> 450
 <212> PRT
 <213> Secuencia artificial

ES 2 594 284 T3

<220>

<223> codificada por la SEC ID N.º: 45

<220>

5 <221> MISC_FEATURE

<222> (2)..(2)

<223> Ala

<220>

10 <221> MISC_FEATURE

<222> (3)..(8)

<223> cola de His constituida por 6 His

<400> 46

Met Ala His His His His His His Gln Asn Ala Ala Val Ser Glu Asn
1 5 10 15

Gln Asn His Asp Asp Gly Ala Ala Ser Ser Pro Gly Phe Lys Leu Val
20 25 30

Gly Phe Ser Lys Phe Val Arg Lys Asn Pro Lys Ser Asp Lys Phe Lys
35 40 45

Val Lys Arg Phe His His Ile Glu Phe Trp Cys Gly Asp Ala Thr Asn
50 55 60

Val Ala Arg Arg Phe Ser Trp Gly Leu Gly Met Arg Phe Ser Ala Lys
65 70 75 80

Ser Asp Leu Ser Thr Gly Asn Met Val His Ala Ser Tyr Leu Leu Thr
85 90 95

Ser Gly Asp Leu Arg Phe Leu Phe Thr Ala Pro Tyr Ser Pro Ser Leu
100 105 110

Ser Ala Gly Glu Ile Lys Pro Thr Thr Thr Ala Ser Ile Pro Ser Phe
115 120 125

Asp His Gly Ser Cys Arg Ser Phe Phe Ser Ser His Gly Leu Gly Val
130 135 140

Arg Ala Val Ala Ile Glu Val Glu Asp Ala Glu Ser Ala Phe Ser Ile
145 150 155 160

Ser Val Ala Asn Gly Ala Ile Pro Ser Ser Pro Pro Ile Val Leu Asn
165 170 175

Glu Ala Val Thr Ile Ala Glu Val Lys Leu Tyr Gly Asp Val Val Leu
180 185 190

15

ES 2 594 284 T3

Arg Tyr Val Ser Tyr Lys Ala Glu Asp Thr Glu Lys Ser Glu Phe Leu
 195 200 205

Pro Gly Phe Glu Arg Val Glu Asp Ala Ser Ser Phe Pro Leu Asp Tyr
 210 215 220

Gly Ile Arg Arg Leu Asp His Ala Val Gly Asn Val Pro Glu Leu Gly
 225 230 235 240

Pro Ala Leu Thr Tyr Val Ala Gly Phe Thr Gly Phe His Gln Phe Ala
 245 250 255

Glu Phe Thr Ala Asp Asp Val Gly Thr Ala Glu Ser Gly Leu Asn Ser
 260 265 270

Ala Val Leu Ala Ser Asn Asp Glu Met Val Leu Leu Pro Ile Asn Glu
 275 280 285

Pro Val His Gly Thr Lys Arg Lys Ser Gln Ile Gln Thr Tyr Leu Glu
 290 295 300

His Asn Glu Gly Ala Gly Leu Gln His Leu Ala Leu Met Ser Glu Asp
 305 310 315 320

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

Phe Asp Phe Met Pro Ser Pro Pro Pro Thr Tyr Tyr Gln Asn Leu Lys
 340 345 350

Lys Arg Val Gly Asp Val Leu Ser Asp Asp Gln Ile Lys Glu Cys Glu
 355 360 365

Glu Leu Gly Ile Leu Val Asp Arg Asp Asp Gln Gly Thr Leu Leu Gln
 370 375 380

Ile Phe Thr Lys Pro Leu Gly Asp Arg Pro Thr Ile Phe Ile Glu Ile
 385 390 395 400

Ile Gln Arg Val Gly Cys Met Met Lys Asp Glu Glu Gly Lys Ala Tyr
 405 410 415

Gln Ser Gly Gly Cys Gly Gly Phe Gly Ile Gly Asn Phe Ser Glu Leu
 420 425 430

Phe Lys Ser Ile Glu Glu Tyr Glu Lys Thr Leu Glu Ala Lys Gln Leu
 435 440 445

Val Gly
 450

ES 2 594 284 T3

<211> 1353
 <212> ADN
 <213> Secuencia artificial

5 <220>
 <223> secuencia de nucleótidos que codifica polipéptido HPPD mutante

<220>
 <221> misc_feature

10 <222> (4)..(6)
 <223> secuencia que codifica Ala

<220>
 <221> misc_feature

15 <222> (7)..(24)
 <223> secuencia que codifica una cola de His que contiene 6 His

<400> 47

atgggtcctc	accatcacca	tcaccaaacc	gccgcccgttt	cagagaaatca	aaaccatgat	60
gacggcgctg	cgctcgtcgc	gggattcaag	ctcgtcggat	ttccaagtt	cgtaagaaag	120
aatccaaagt	ctgataaatt	caaggttaag	cgcttccatc	acatcgagtt	ctgggtcggc	180
gacgcaacca	acgtcgtcgc	tcgcttctcc	tggggctctgg	ggatgagatt	ctccgccaaa	240
tccgatcttt	ccaccggaaa	catggttcac	gcctcttacc	tactcacctc	cggtgacctc	300
cgattccttt	tactcgtccc	ttactctccg	tctctctccg	ccggagagat	taaaccgaca	360
accacagctt	ctatcccaag	tttcgatcac	ggctcttgtc	gttccttctt	ctcgtcacat	420
ggtctcggtg	ttagagccgt	tgcgattgaa	gtagaagacg	cagagtcagc	tttctccatc	480
agtgtagcta	atggcgctat	tccttcgtcg	cctcctatcg	tcctcaatga	agcagttacg	540
atcgtctgagg	ttaaactata	cggcgatggt	gttctccgat	atgttagtta	caaagcagaa	600
gataccgaaa	aatccgaatt	cttgccaggg	ttcgagcgtg	tagaggatgc	gtcgtcgttc	660
ccattggatt	atgggatccg	gcccgttgac	cacgccgtgg	gaaacgttcc	tgagcttggg	720
ccggctttaa	cttatgtagc	ggggttcact	ggttttcacc	aattcgcaga	gttcacagca	780
gacgacgttg	gaaccgccga	gagcggttta	aattcagcgg	tcctggctag	caatgatgaa	840
atggttcttc	taccgattaa	cgagccagtg	cacggaacaa	agaggaagag	tcagattcag	900
acgtatttgg	aacataacga	aggcgcaggg	ctacaacatc	tggctctgat	gagtgaagac	960
atattcagga	ccctgagaga	gatgaggaag	aggagcagta	ttggaggatt	cgacttcatg	1020
ccttctctc	cgctactta	ctaccagaat	ctcaagaaac	gggtcggcga	cgtgctcagc	1080
gatgatcaga	tcaaggagtg	tgaggaatta	gggattcttg	tagacagaga	tgatcaaggg	1140
acgttgcttc	aaatcttcac	aaaaccacta	ggtgacaggg	cgacgatatt	tatagagata	1200
atccagagag	taggatgcat	gatgaaagat	gaggaagga	aggcttacca	gagtggagga	1260
tgtggtggtt	ttggccaggg	caatttctct	gagctcttca	agtccattga	agaatacgaa	1320
aagactcttg	aagccaaaca	gttagtggga	tga			1353

20 <210> 48
 <211> 450
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> codificada por la SEC ID N.º: 47

<220>
 <221> MISC_FEATURE
 <222> (2)..(2)
 <223> Ala

30

ES 2 594 284 T3

<220>

<221> MISC_FEATURE

<222> (3)..(8)

<223> cola de His constituída por 6 His

5

<400> 48

```

Met Ala His His His His His His Gln Asn Ala Ala Val Ser Glu Asn
1           5           10           15

Gln Asn His Asp Asp Gly Ala Ala Ser Ser Pro Gly Phe Lys Leu Val
          20           25           30

Gly Phe Ser Lys Phe Val Arg Lys Asn Pro Lys Ser Asp Lys Phe Lys
          35           40           45

Val Lys Arg Phe His His Ile Glu Phe Trp Cys Gly Asp Ala Thr Asn
          50           55           60

Val Ala Arg Arg Phe Ser Trp Gly Leu Gly Met Arg Phe Ser Ala Lys
65           70           75           80

Ser Asp Leu Ser Thr Gly Asn Met Val His Ala Ser Tyr Leu Leu Thr
          85           90           95

Ser Gly Asp Leu Arg Phe Leu Phe Thr Ala Pro Tyr Ser Pro Ser Leu
          100          105          110

Ser Ala Gly Glu Ile Lys Pro Thr Thr Thr Ala Ser Ile Pro Ser Phe
          115          120          125

Asp His Gly Ser Cys Arg Ser Phe Phe Ser Ser His Gly Leu Gly Val
130           135           140

Arg Ala Val Ala Ile Glu Val Glu Asp Ala Glu Ser Ala Phe Ser Ile
145           150           155           160

Ser Val Ala Asn Gly Ala Ile Pro Ser Ser Pro Pro Ile Val Leu Asn
          165          170          175

Glu Ala Val Thr Ile Ala Glu Val Lys Leu Tyr Gly Asp Val Val Leu
          180          185          190

Arg Tyr Val Ser Tyr Lys Ala Glu Asp Thr Glu Lys Ser Glu Phe Leu
          195          200          205

Pro Gly Phe Glu Arg Val Glu Asp Ala Ser Ser Phe Pro Leu Asp Tyr
210           215           220

Gly Ile Arg Arg Leu Asp His Ala Val Gly Asn Val Pro Glu Leu Gly

```

ES 2 594 284 T3

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

<210> 49
 <211> 1353
 <212> ADN
 <213> Secuencia artificial

5

<220>
 <223> secuencia de nucleótidos que codifica polipéptido HPPD mutante

10

<220>

ES 2 594 284 T3

<221> misc_feature
 <222> (4)..(6)
 <223> secuencia que codifica Ala

5 <220>
 <221> misc_feature
 <222> (7)..(24)
 <223> secuencia que codifica una cola de His que contiene 6 His

10 <400> 49
 atgggtcacc accatcacc taccacaaac gccgcggtt cagagaaatca aaaccatgat 60
 gacggcgctg cgtcgtcgc gggattcaag ctcgctcggat ttccaagtt cgtaaagaaag 120
 aatccaaagt ctgataaatt caaggttaag cgtttccatc acatcgagtt ctgggtcggc 180
 gacgcaacca acgtcgtcgtc tgcgttctcc tggggtctgg ggatgagatt ctccgccaaa 240
 tccgatcttt ccaccggaaa catggttcc gctcttacc tactcacctc cggtgacctc 300
 cgatttcctt tcaactgctc ttactctccg tctctctccg ccggagagat taaaccgaca 360
 accacagctt ctatcccaag tttagatcac ggctcttgct gttccttctt ctcgtcacat 420
 ggtctcggtg ttagagccgt tgcgattgaa gtagaagacg cagagtcagc tttctccatc 480
 agtgtagcta atggcgctat tccttcgtcg cctcctatcg tcctcaatga agcagttacg 540
 atcgtctgagg ttaaaactata ccggcgatggt gttctccgat atgttagtta caaagcagaa 600
 gataccgaaa aatccgaatt cttgccaggg ttcgagcgtg tagaggatgc gtcgtcgttc 660
 ccattggatt atgggtatcc gggccttgac cacgcctgg gaaacgttcc tgagcttggg 720
 ccggctttaa cttatgtagc ggggttccact ggttttcacc aattcgcaga gttcacagca 780
 gacgacgttg gaaccgcga gagcggttta aattcagcgg tcctggctag caatgatgaa 840
 atggttcttc taccgattaa ccgagccagtg cacggaacaa agaggaagag tcagattcag 900
 acgtatcttg aacataacga aggcgcaggg ctacaacatc tggctctgat gagtgaagac 960
 atattcagga ccttgagaga gatgaggaag aggagcagta ttggaggatt cgacttcatg 1020
 ccttctctc cgcctactta ctaccagaat ctcaagaaac gggtcggcga cgtgctcagc 1080
 gatgatcaga tcaaggagtg tgaggaatta gggattcttg tagacagaga tgatcaaggg 1140
 acgttgcttc aaatcttcc aaaaccacta ggtgacaggc cgacgatatt tatagagata 1200
 atccagagag taggatgcat gatgaaagat gaggaaggga aggcttacca gagtggagga 1260
 tgtggtggtt ttggcgtggg caatttctct gagctcttca agtccattga agaatacga 1320
 aagactcttg aagccaaaca gttagtgagg tga 1353

<210> 50
 <211> 450
 15 <212> PRT
 <213> Secuencia artificial

<220>
 <223> codificada por la SEC ID N.º: 49

20 <220>
 <221> MISC_FEATURE
 <222> (2)..(2)
 <223> Ala

25 <220>
 <221> MISC_FEATURE
 <222> (3)..(8)
 <223> cola de His constituida por 6 His

30 <400> 50

ES 2 594 284 T3

Met Ala His His His His His His Gln Asn Ala Ala Val Ser Glu Asn
1 5 10 15

Gln Asn His Asp Asp Gly Ala Ala Ser Ser Pro Gly Phe Lys Leu Val
20 25 30

Gly Phe Ser Lys Phe Val Arg Lys Asn Pro Lys Ser Asp Lys Phe Lys
35 40 45

Val Lys Arg Phe His His Ile Glu Phe Trp Cys Gly Asp Ala Thr Asn
50 55 60

Val Ala Arg Arg Phe Ser Trp Gly Leu Gly Met Arg Phe Ser Ala Lys
65 70 75 80

Ser Asp Leu Ser Thr Gly Asn Met Val His Ala Ser Tyr Leu Leu Thr
85 90 95

Ser Gly Asp Leu Arg Phe Leu Phe Thr Ala Pro Tyr Ser Pro Ser Leu
100 105 110

Ser Ala Gly Glu Ile Lys Pro Thr Thr Thr Ala Ser Ile Pro Ser Phe
115 120 125

Asp His Gly Ser Cys Arg Ser Phe Phe Ser Ser His Gly Leu Gly Val
130 135 140

Arg Ala Val Ala Ile Glu Val Glu Asp Ala Glu Ser Ala Phe Ser Ile
145 150 155 160

Ser Val Ala Asn Gly Ala Ile Pro Ser Ser Pro Pro Ile Val Leu Asn
165 170 175

Glu Ala Val Thr Ile Ala Glu Val Lys Leu Tyr Gly Asp Val Val Leu
180 185 190

Arg Tyr Val Ser Tyr Lys Ala Glu Asp Thr Glu Lys Ser Glu Phe Leu
195 200 205

Pro Gly Phe Glu Arg Val Glu Asp Ala Ser Ser Phe Pro Leu Asp Tyr
210 215 220

Gly Ile Arg Arg Leu Asp His Ala Val Gly Asn Val Pro Glu Leu Gly
225 230 235 240

Pro Ala Leu Thr Tyr Val Ala Gly Phe Thr Gly Phe His Gln Phe Ala
245 250 255

Glu Phe Thr Ala Asp Asp Val Gly Thr Ala Glu Ser Gly Leu Asn Ser
260 265 270

ES 2 594 284 T3

Ala Val Leu Ala Ser Asn Asp Glu Met Val Leu Leu Pro Ile Asn Glu
 275 280 285

Pro Val His Gly Thr Lys Arg Lys Ser Gln Ile Gln Thr Tyr Leu Glu
 290 295 300

His Asn Glu Gly Ala Gly Leu Gln His Leu Ala Leu Met Ser Glu Asp
 305 310 315 320

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

Phe Asp Phe Met Pro Ser Pro Pro Pro Thr Tyr Tyr Gln Asn Leu Lys
 340 345 350

Lys Arg Val Gly Asp Val Leu Ser Asp Asp Gln Ile Lys Glu Cys Glu
 355 360 365

Glu Leu Gly Ile Leu Val Asp Arg Asp Asp Gln Gly Thr Leu Leu Gln
 370 375 380

Ile Phe Thr Lys Pro Leu Gly Asp Arg Pro Thr Ile Phe Ile Glu Ile
 385 390 395 400

Ile Gln Arg Val Gly Cys Met Met Lys Asp Glu Glu Gly Lys Ala Tyr
 405 410 415

Gln Ser Gly Gly Cys Gly Gly Phe Gly Val Gly Asn Phe Ser Glu Leu
 420 425 430

Phe Lys Ser Ile Glu Glu Tyr Glu Lys Thr Leu Glu Ala Lys Gln Leu
 435 440 445

Val Gly
 450

1

REIVINDICACIONES

1. Un ácido nucleico aislado que comprende una secuencia de nucleótidos que codifica una proteína de HPPD mutada, en el que dicha proteína de HPPD mutada tiene actividad de HPPD y en el que en dicha proteína de HPPD mutada el aminoácido en una posición correspondiente a la posición 252 de la SEC ID N.º: 2 ha sido reemplazado de tal modo que la secuencia de aminoácidos resultante comprende un aminoácido seleccionado de His, Ile, Leu, Asn o Ser en la posición correspondiente a la posición 252 de la secuencia de aminoácidos de la SEC ID N.º: 2:
2. El ácido nucleico aislado de la reivindicación 1, en el que dicha proteína HPPD se deriva de *Streptomyces avermitilis* (Genebank SAV11864), *Daucus carota* (Genebank DCU 87257), *Arabidopsis thaliana* (Genebank AF047834), *Mycosphaerella graminicola* (Genebank AF038152), *Oryza sativa* / arroz [BAD26248], *Zea mays* / maíz [ACN36372], *Avena sativa* [ABZ23427], *Pseudomonas fluorescens* [ABF50055], *Synechococcus* sp. [YP_473959], *Blepharisma japonicum* [BAF91881], *Rhodococcus* RHA1 sp. ro0240 [YP_702005], *Rhodococcus* RHA1 sp. ro0341 [YP_703002], *Picrophilus torridus* [YP_024147], *Kordia algicida* [ZP_02161490], *Sorghum bicolor* [XP_002453359], *Triticum aestivum* / trigo [AAZ67144] u *Hordeum vulgare* / cebada [O48604].
3. El ácido nucleico aislado de la reivindicación 1, en el que dicha proteína HPPD comprende la secuencia de aminoácidos de la SEC ID N.º: 4 y en el que la secuencia de aminoácidos resultante comprende un aminoácido seleccionado de His, Ile, Leu, Asn o Ser en la posición 249 de la SEC ID N.º: 4 (que corresponde a la posición 252 de la SEC ID N.º: 2).
4. El ácido nucleico aislado de la reivindicación 1, en el que dicha proteína HPPD comprende la secuencia de aminoácidos de la SEC ID N.º: 6 y en el que la secuencia de aminoácidos resultante comprende un aminoácido seleccionado de His, Ile, Leu, Asn o Ser en la posición 226 de la SEC ID N.º: 6 (que corresponde a la posición 252 de la SEC ID N.º: 2).
5. El ácido nucleico aislado de la reivindicación 1, en el que dicha proteína HPPD comprende la secuencia de aminoácidos de la SEC ID N.º: 8 y en el que la secuencia de aminoácidos resultante comprende un aminoácido seleccionado de His, Ile, Leu, Asn o Ser en la posición 243 de la SEC ID N.º: 8 (que corresponde a la posición 252 de la SEC ID N.º: 2).
6. El ácido nucleico aislado de la reivindicación 1, en el que dicha proteína HPPD comprende la secuencia de aminoácidos de la SEC ID N.º: 10 y en el que la secuencia de aminoácidos resultante comprende un aminoácido seleccionado de His, Ile, Leu, Asn o Ser en la posición 190 de la SEC ID N.º: 10 (que corresponde a la posición 252 de la SEC ID N.º: 2).
7. El ácido nucleico aislado de la reivindicación 1, en el que dicha proteína HPPD comprende la secuencia de aminoácidos de la SEC ID N.º: 14 y en el que la secuencia de aminoácidos resultante comprende un aminoácido seleccionado de His, Ile, Leu, Asn o Ser en la posición 171 de la SEC ID N.º: 14 (que corresponde a la posición 252 de la SEC ID N.º: 2).
8. El ácido nucleico aislado de la reivindicación 1, en el que dicha proteína HPPD comprende la secuencia de aminoácidos de la SEC ID N.º: 16 y en el que la secuencia de aminoácidos resultante comprende un aminoácido seleccionado de His, Ile, Leu, Asn o Ser en la posición 211 de la SEC ID N.º: 16 (que corresponde a la posición 252 de la SEC ID N.º: 2).
9. El ácido nucleico aislado de la reivindicación 1, en el que dicha proteína HPPD comprende la secuencia de aminoácidos de la SEC ID N.º: 18 y en el que la secuencia de aminoácidos resultante comprende un aminoácido seleccionado de His, Ile, Leu, Asn o Ser en la posición 233 de la SEC ID N.º: 18 (que corresponde a la posición 252 de la SEC ID N.º: 2).
10. El ácido nucleico aislado de la reivindicación 1, en el que dicha proteína HPPD comprende la secuencia de aminoácidos de la SEC ID N.º: 30 y en el que la secuencia de aminoácidos resultante comprende un aminoácido seleccionado de His, Ile, Leu, Asn o Ser en la posición 234 de la SEC ID N.º: 30 (que corresponde a la posición 252 de la SEC ID N.º: 2).
11. El ácido nucleico aislado de la reivindicación 1, en el que dicha proteína HPPD comprende la secuencia de aminoácidos de la SEC ID N.º: 20 y en el que la secuencia de aminoácidos resultante comprende un aminoácido seleccionado de His, Ile, Leu, Asn o Ser en la posición 203 de la SEC ID N.º: 20 (que corresponde a la posición 252 de la SEC ID N.º: 2).
12. El ácido nucleico aislado de la reivindicación 1, en el que dicha proteína HPPD comprende la secuencia de aminoácidos de la SEC ID N.º: 22 y en el que la secuencia de aminoácidos resultante comprende un aminoácido seleccionado de His, Ile, Leu, Asn Val en la posición 2421 de la SEC ID N.º: 22 (que corresponde a la posición 252 de la SEC ID N.º: 2).
13. El ácido nucleico aislado de la reivindicación 1, en el que dicha proteína HPPD comprende la secuencia de

aminoácidos de la SEC ID N.º: 24 y en el que la secuencia de aminoácidos resultante comprende un aminoácido seleccionado de His, Ile, Leu, Asn o Ser en la posición 243 de la SEC ID N.º: 24 (que corresponde a la posición 252 de la SEC ID N.º: 2).

- 5 14. El ácido nucleico aislado de la reivindicación 1, en el que dicha proteína HPPD comprende la secuencia de aminoácidos de la SEC ID N.º: 26 y en el que la secuencia de aminoácidos resultante comprende un aminoácido seleccionado de His, Ile, Leu o Asn en la posición 239 de la SEC ID N.º: 26 (que corresponde a la posición 252 de la SEC ID N.º: 2).
- 10 15. El ácido nucleico aislado de la reivindicación 1, en el que dicha proteína HPPD comprende la secuencia de aminoácidos de la SEC ID N.º: 2 y en el que la secuencia de aminoácidos resultante comprende un aminoácido seleccionado de His, Ile, Leu, Asn o Ser en la posición 252 de la SEC ID N.º: 2.
16. El ácido nucleico de una cualquiera de las reivindicaciones 1 a 15, en el que dicha proteína HPPD es capaz de incrementar la tolerancia de una planta a al menos un herbicida que actúa sobre la HPPD.
17. Una proteína codificada por el ácido nucleico aislado de una cualquiera de las reivindicaciones 1 a 15.
- 15 18. Un gen quimérico que comprende una secuencia codificada que comprende el ácido nucleico de una cualquiera de las reivindicaciones 1 a 15 unido operativamente a un promotor expresable en plantas y opcionalmente una terminación de transcripción y una región de poliadenilación.
19. Un procedimiento de obtención de una proteína HPPD mutada capaz de modular la tolerancia de una planta a al menos un herbicida que actúa sobre la HPPD, en donde dicha proteína HPPD mutada tiene actividad HPPD, comprendiendo el procedimiento
- 20 i) proporcionar una proteína HPPD;
- ii) reemplazar un aminoácido en dicha proteína HPPD, de tal forma que la secuencia de aminoácidos resultante comprende una seleccionada entre His, Ile, Leu, Asn o Ser en una posición en una proteína HPPD, correspondiendo dicha posición a la posición 252 de la secuencia de aminoácidos de la SEC ID N.º: 2;
- iii) determinar la inhibición de la proteína HPPD resultante por al menos un herbicida que actúa sobre la HPPD,
- 25 en el que una inhibición de la proteína resultante de menos o más que la observada con una proteína de HPPD de referencia es indicativa de que la proteína resultante es capaz de modular la tolerancia de una planta a dicho herbicida.
20. El procedimiento de la reivindicación 19, en el que dicha proteína HPPD mutada es capaz de incrementar la tolerancia de una planta a al menos un herbicida inhibidor de la HPPD.
- 30 21. El ácido nucleico de la reivindicación 16 o el procedimiento de la reivindicación 19, en el que el herbicida que actúa sobre la HPPD es tembrotriona.
22. Un procedimiento de producción de una planta transgénica que comprende introducir dentro de dicha planta el ácido nucleico de una cualquiera de las reivindicaciones 1 a 16 o un ácido nucleico que codifica la enzima HPPD identificada por el procedimiento de una cualquiera de las reivindicaciones 19 a 20, ambos unidos operativamente a un promotor expresable en plantas o el gen quimérico de la reivindicación 18.
- 35 23. El procedimiento de la reivindicación 22, en el que el ácido nucleico de la reivindicación 16 o un ácido nucleico identificado por el procedimiento de las reivindicaciones 19 o 20, ambos unidos operativamente a un promotor expresable en plantas, o el gen quimérico de la reivindicación 18 que comprende el ácido nucleico de una cualquiera de las reivindicaciones 1 a 16, se introducen en dicha planta.
- 40 24. El ácido nucleico de una cualquiera de las reivindicaciones 1 a 15 o el procedimiento de una cualquiera de las reivindicaciones 22 a 23, en el que dicha proteína HPPD comprende:
- a) una His en una posición en una proteína de HPPD, correspondiendo dicha posición a la posición 226 de la secuencia de aminoácidos de la SEC ID N.º: 2;
- 45 b) una Ser en una posición en una proteína de HPPD, correspondiendo dicha posición a la posición 267 de la secuencia de aminoácidos de la SEC ID N.º: 2;
- c) una Asn en una posición en una proteína de HPPD, correspondiendo dicha posición a la posición 282 de la secuencia de aminoácidos de la SEC ID N.º: 2;
- d) una His en una posición en una proteína de HPPD, correspondiendo dicha posición a la posición 308 de la secuencia de aminoácidos de la SEC ID N.º: 2;
- 50 e) una Tyr en una posición en una proteína de HPPD, correspondiendo dicha posición a la posición 342 de la

secuencia de aminoácidos de la SEC ID N.º: 2;

f) una Glu en una posición en una proteína de HPPD, correspondiendo dicha posición a la posición 394 de la secuencia de aminoácidos de la SEC ID N.º: 2;

5 g) una Gly en una posición en una proteína de HPPD, correspondiendo dicha posición a la posición 420 de la secuencia de aminoácidos de la SEC ID N.º: 2;

h) una Asn en una posición en una proteína de HPPD, correspondiendo dicha posición a la posición 423 de la secuencia de aminoácidos de la SEC ID N.º: 2.

25. Una célula vegetal que comprende el ácido nucleico aislado de una cualquiera de las reivindicaciones 1 a 15 o el gen quimérico de la reivindicación 18 en su información genética.

10 26. Una planta, una parte de la misma o un tejido vegetal que consisten esencialmente en las células vegetales de la reivindicación 25.

27. La planta de la reivindicación 26 que se selecciona de trigo, algodón, colza, arroz, maíz, soja y sorgo.

15 28. Un procedimiento de modulación de la tolerancia de una planta a al menos un herbicida que actúa sobre la HPPD o de obtención una planta tolerante a un herbicida inhibidor de la HPPD que comprende introducir el ácido nucleico aislado de una cualquiera de las reivindicaciones 1 a 15 unido operativamente a un promotor expresable en plantas o el gen quimérico de la reivindicación 18 dentro del genoma de una planta.

29. Uso de un gen quimérico de la reivindicación 18 o del ácido nucleico de una cualquiera de las reivindicaciones 1 a 15 unido operativamente a un promotor expresable en plantas para modular la tolerancia de una planta a al menos un herbicida inhibidor de la HPPD aplicado para controlar malas hierbas.

20 30. El procedimiento de las reivindicaciones 19, 22 o 28 o el uso de la reivindicación 30, en los que se incrementa la tolerancia de una planta a al menos un herbicida inhibidor de la HPPD.

Figura 1a: superposición de la estructura de rayos X de HPPD de *Arabidopsis thaliana* (gris oscuro) y *Pseudomonas fluorescens* (gris claro). La estructura se muestra como gráfico de cintas.

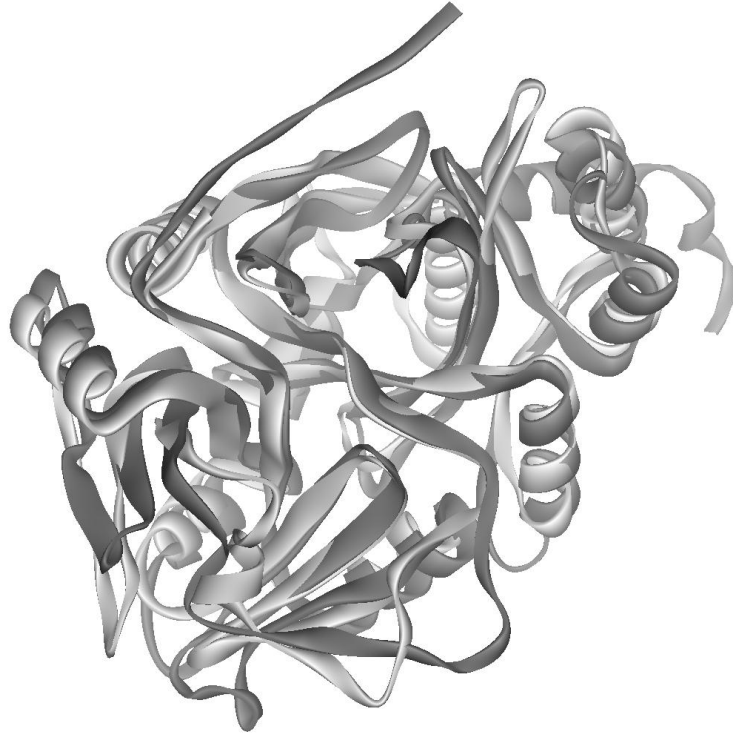


Figura 1b: superposición de la estructura de rayos X de HPPD de *Arabidopsis thaliana* (gris oscuro) y *Streptomyces avermitilis* (gris claro). La estructura se muestra como gráfico de cintas.

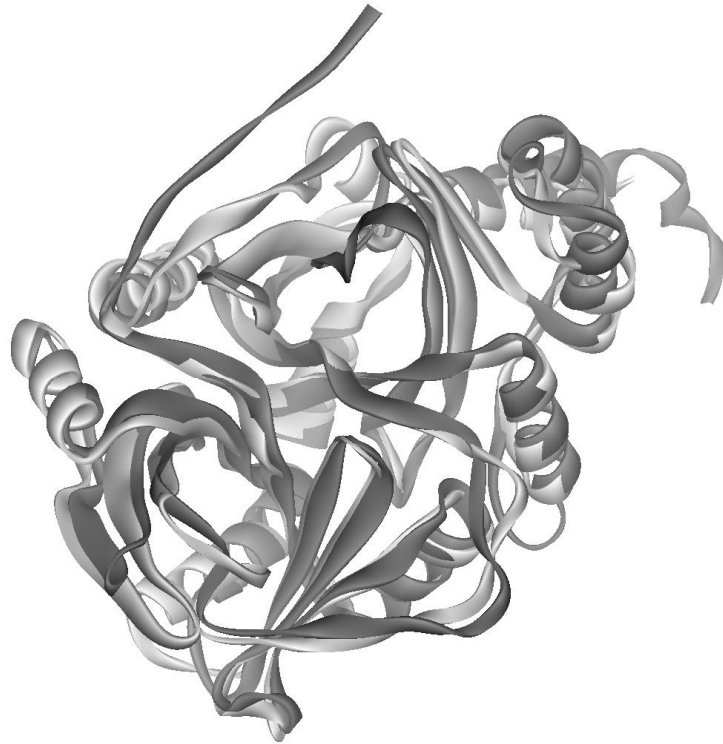


Figura 1c: superposición de la estructura de rayos X de HPPD de *Arabidopsis thaliana* (gris oscuro) y *Homo sapiens* (gris claro). La estructura se muestra como gráfico de cintas.



Figura 1d: superposición de la estructura de rayos X de HPPD de *Arabidopsis thaliana* (gris oscuro) y *Rattus norvegicus* (gris claro). La estructura se muestra como gráfico de cintas.

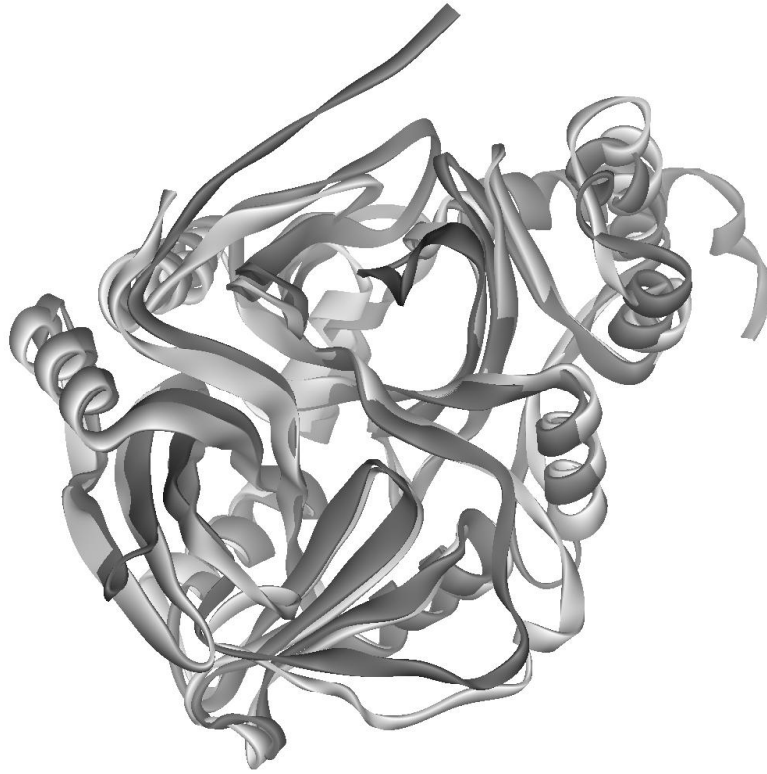


Figura 2a: aminoácidos que forman el sitio de unión de *Arabidopsis thaliana* (representado por líneas en negrita). El núcleo de la proteína de HPPD de *Arabidopsis thaliana* se muestra como traza de C_{α} y el hierro está marcado.



Figura 2b: aminoácidos que forman el sitio de unión de *Pseudomonas fluorescens* (representado por líneas en negrita). El núcleo de la proteína de HPPD de *Pseudomonas fluorescens* se muestra como traza de C_{alfa} y el hierro está marcado.

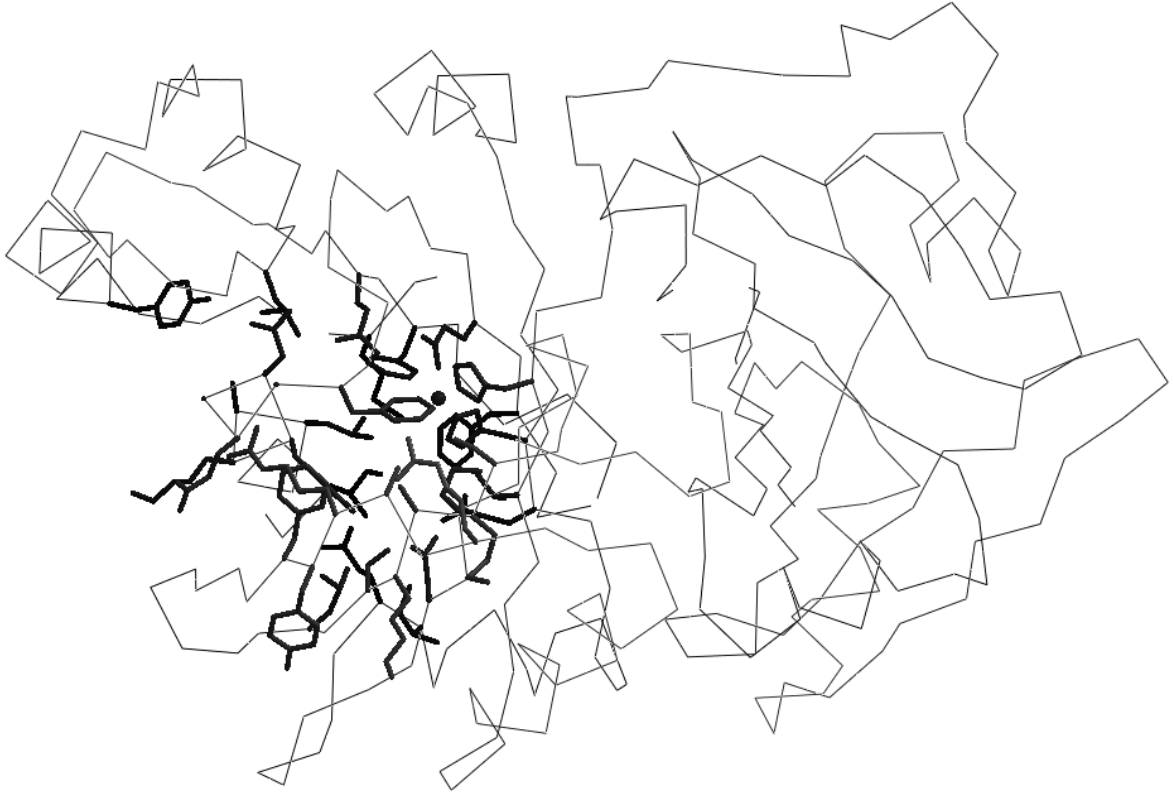


Figura 2c: aminoácidos que forman el sitio de unión de *Streptomyces avermitilis* (representado por líneas en negrita). El núcleo de la proteína de HPPD de *Streptomyces avermitilis* se muestra como traza de C_{α} y el hierro está marcado.

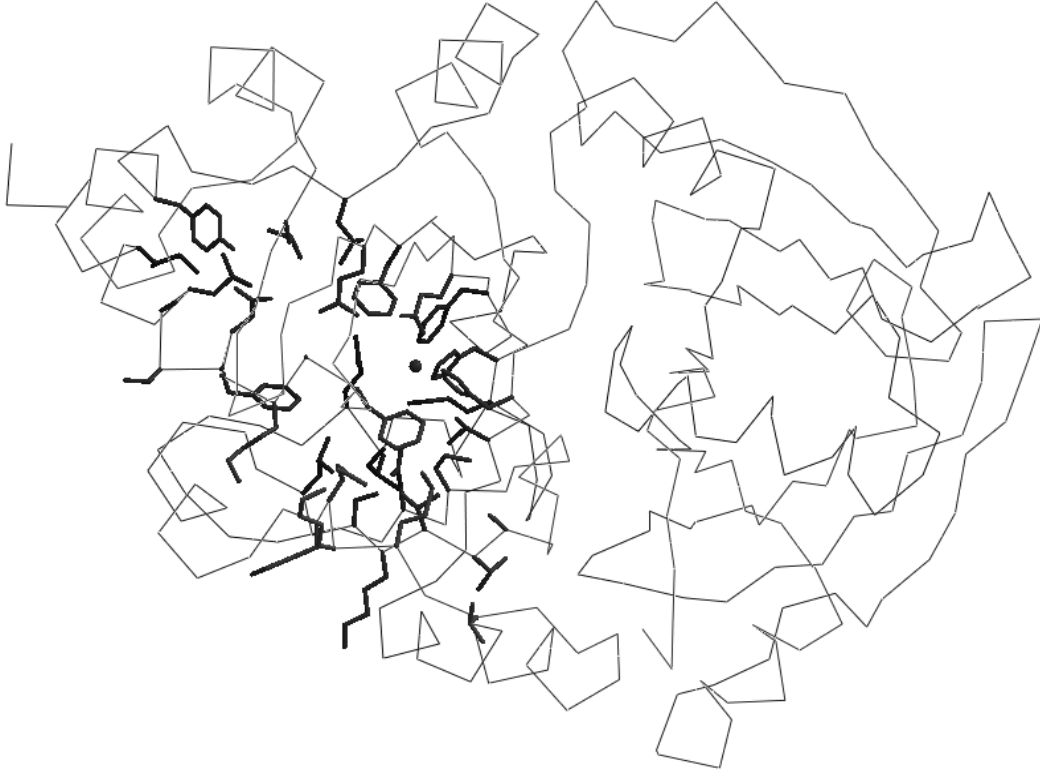


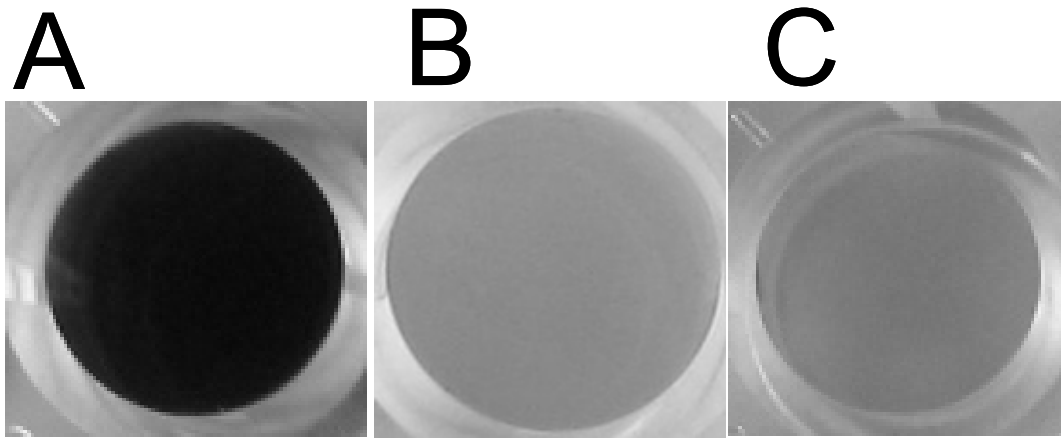
Figura 2d: aminoácidos que forman el sitio de unión de *Homo sapiens* (representado por líneas en negrita). El núcleo de la proteína de HPPD de *Homo sapiens* se muestra como traza de C_{alfa} y el hierro está marcado.



Figura 2e: aminoácidos que forman el sitio de unión de *Rattus norvegicus* (representado por líneas en negrita). El núcleo de la proteína de HPPD de *Rattus norvegicus* se muestra como traza de C_{α} y el hierro está marcado.



FIGURA 3: color de ensayo marrón



A: WT, B: N282A, C: S267A