

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 806 732**

51 Int. Cl.:

C12N 1/21 (2006.01)
C12N 9/06 (2006.01)
C12N 9/10 (2006.01)
C12N 9/12 (2006.01)
C12N 9/16 (2006.01)
C12N 9/88 (2006.01)
C12N 15/52 (2006.01)
C12N 15/55 (2006.01)
C12P 17/16 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **16.12.2016** **PCT/EP2016/081598**

87 Fecha y número de publicación internacional: **22.06.2017** **WO17103221**

96 Fecha de presentación y número de la solicitud europea: **16.12.2016** **E 16822958 (1)**

97 Fecha y número de publicación de la concesión europea: **22.04.2020** **EP 3390618**

54 Título: **Fábrica de células bacterianas modificadas genéticamente para la producción de tiamina**

30 Prioridad:

18.12.2015 EP 15201200

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

18.02.2021

73 Titular/es:

BIOSYNTIA APS (100.0%)
Fruebjergvej 3
2100 Copenhagen Ø, DK

72 Inventor/es:

GRONENBERG, LUISA;
SALOMONSEN, BO;
FERLA, MATTEO y
GENEE, HANS JASPER

74 Agente/Representante:

LINAGE GONZÁLEZ, Rafael

Observaciones:

Véase nota informativa (Remarks, Remarques o Bemerkungen) en el folleto original publicado por la Oficina Europea de Patentes

ES 2 806 732 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

Fábrica de células bacterianas modificadas genéticamente para la producción de tiamina

5 **Campo de la invención**

La invención se refiere a una bacteria modificada genéticamente para la producción de tiamina no fosforilada; en la que la bacteria se caracteriza por un transgén que codifica para una tiamina monofosfato fosfatasa (TMP fosfatasa que tiene EC.3.1.3.-) así como transgenes que codifican para polipéptidos que catalizan etapas en la ruta de la tiamina. La bacteria modificada genéticamente se caracteriza por una síntesis y liberación potenciadas de tiamina no fosforilada en el entorno extracelular. La invención se refiere además a un método para producir tiamina que usa la bacteria modificada genéticamente de la invención; así como al uso de la bacteria modificada genéticamente para la producción de tiamina.

15 **Antecedentes de la invención**

La tiamina, también conocida como vitamina B1, es un miembro del complejo B de vitaminas soluble en agua y es un requisito nutricional para los mamíferos. En la naturaleza, la tiamina existe en múltiples estados de fosforilación: tiamina, tiamina monofosfato (TMP) y tiamina difosfato (TPP) (también existen cantidades traza de tiamina trifosfato halladas en células). Todos los organismos vivos usan la tiamina, pero la tiamina y su forma activa tiamina pirofosfato (TPP) sólo se sintetizan en bacterias, hongos y plantas. Los animales dependen de su dieta como fuente de tiamina y, por tanto, para los seres humanos, es un nutriente esencial. La TPP actúa *in vivo* como la coenzima de enzimas que realizan varios procesos metabólicos vitales tales como la ruta de la pentosa fosfato y el ciclo de TCA. La tiamina se encuentra en la Lista de Medicamentos Esenciales de la Organización Mundial de la Salud que enumera los medicamentos más importantes necesarios en un sistema de salud básico, y la deficiencia de tiamina es un problema de salud extendido. La forma comercialmente relevante de tiamina es no fosforilada, dado que es la forma más estable que puede asimilarse y fosforilarse por seres humanos y animales para producir el cofactor biológicamente activo TPP.

Las rutas de biosíntesis de tiamina caracterizadas en bacterias, algunos protozoos, plantas y hongos comparten algunas características comunes (figura 1). Los restos tiazol y pirimidina se biosintetizan por separado. El resto pirimidina, 4-amino-5-hidroxi-metil-2-metilpirimidina fosfato (HMP-P), se deriva de ribótido de 5-aminoimidazol (AIR), un producto intermedio en la ruta de biosíntesis *de novo* de purina. En bacterias gramnegativas, la conversión de AIR en HMP-P está catalizada por el producto génico *thiC*. Entonces HMP-P se fosforila para dar HMP-PP mediante *ThiD* cinasa antes de acoplarse con la unidad de tiazol. El resto tiazol, fosfato de 5-(2-hidroxi-etil)-4-metiltiazol (HET-P), se deriva de L-tirosina y 1-desoxi-D-xilulosa fosfato (DXP) y cisteína; en el que el átomo de azufre se deriva probablemente de la L-cisteína. Esta última reacción requiere la expresión de al menos cinco genes *thiF*, *thiS*, *thiG*, *thiH* y *thiI*.

Los restos pirimidina y tiazol se combinan entonces para formar TMP mediante la acción de tiamina fosfato sintasa (EC.2.5.1.3) codificada por *thiE*. Por tanto, TMP es el primer producto de todas las rutas de biosíntesis de tiamina conocidas. En *E. coli* y otras *Enterobacteriaceae*, TMP puede fosforilarse para dar el cofactor TPP mediante una tiamina fosfato cinasa (EC.2.7.4.16) codificada por *thiL* en presencia de ATP. Algunas bacterias y eucariotas dependen de la vía de rescate de HMP y HET que requiere *ThiD*, *ThiE* y *ThiM*. *E. coli* usa una enzima de rescate, tiamina cinasa, codificada por *thiK* para convertir la tiamina exógena, captada por la célula, en TMP.

En vista de su capacidad para sintetizar TMP y TPP, pueden usarse microorganismos como fábricas de células para la producción recombinante de esta vitamina. La producción recombinante de medicamentos esenciales tales como la vitamina tiamina requiere el uso de un huésped adecuado que pueda producir y exportar tiamina. La tiamina no fosforilada (THI) es la diana deseada para la producción biológica, no sólo por motivos comerciales, sino también porque la tiamina atraviesa fácilmente membranas celulares, a diferencia de sus formas fosforiladas, y, por tanto, se acumula fuera del huésped de producción durante la fermentación, donde es la forma más estable.

Las ventajas de *Escherichia coli* recombinante como fábrica de células para la producción de productos biológicos se reconocen ampliamente debido al hecho de que: (i) tiene una cinética de crecimiento rápida única; con un tiempo de duplicación de aproximadamente 20 minutos cuando se cultiva en medios con sales de glucosa y en condiciones medioambientales óptimas, (ii) logra fácilmente una alta densidad celular; en la que se estima que el límite de densidad teórica de un cultivo líquido de *E. coli* es de aproximadamente 200 g de peso de células secas/l o alrededor de 1×10^{13} bacterias viables/ml. De manera adicional, existen muchos protocolos y herramientas moleculares disponibles para la producción de alto nivel de productos biológicos en *E. coli*; las proteínas heterólogas pueden expresarse fácilmente en *E. coli* y existen muchas cepas especializadas disponibles para la producción de productos terminados especializados.

En la mayoría de las bacterias, TPP se produce mediante la conversión directa de TMP en TPP, en la que el crecimiento celular requiere un suministro intracelular o extracelular de TPP. Por tanto, el uso de *E. coli* como fábrica de células para la producción de tiamina requiere una cepa modificada genéticamente que sea viable y, al mismo

tiempo, pueda producir y liberar tiamina en el entorno extracelular.

Sumario

- 5 La invención proporciona una bacteria modificada genéticamente para la producción de tiamina no fosforilada; en la que dicha bacteria se caracteriza por tener siete transgenes que codifican para, respectivamente:
- a. un polipéptido que tiene actividad tiamina monofosfato fosfatasa (E.C.3.1.3.-), en la que la secuencia de aminoácidos de dicho polipéptido tiene al menos el 80% de identidad de secuencia con respecto a una secuencia seleccionada del grupo que consiste en SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72 y 74;
- 10 b. un polipéptido que tiene actividad 4-amino-5-hidroxitimetil-2-metilpirimidina fosfato sintasa (HMP-P) (E.C.4.1.99.17);
- 15 c. un polipéptido que tiene actividad tiamina fosfato sintasa (E.C.2.5.1.3);
- d. un polipéptido que tiene actividad adenilil transferasa de ThiS (E.C.2.7.7.73); un polipéptido que tiene actividad portadora de azufre de ThiS;
- 20 e. un polipéptido que tiene actividad tiazol sintasa (E.C.2.8.1.10);
- f. un polipéptido que tiene actividad 2-iminoacetato sintasa (E.C.4.1.99.19), o que tiene actividad glicina oxidasa (EC 1.4.3.19); y
- 25 g. un polipéptido que tiene actividad fosfohidroximetilpirimidina cinasa (E.C.2.7.4.7); en la que el género de la bacteria es *Escherichia*.
- Según una realización adicional, la bacteria modificada genéticamente de la invención se caracteriza además por un gen *thiL* endógeno modificado genéticamente que puede expresar actividad tiamina fosfato cinasa (EC.2.7.4.16) reducida en comparación con el gen *thiL* endógeno parental. Según una realización adicional, la bacteria modificada genéticamente de la invención se caracteriza además por la inactivación o delección de uno o más genes que codifican para una proteína seleccionada del grupo: proteína de unión periplásmica de transportadores ABC de tiamina; permeasa de transportadores ABC de tiamina; y ATPasa de transportadores ABC de tiamina.
- 30 La invención proporciona además un método para producir tiamina no fosforilada que comprende las etapas de:
- a. introducir una bacteria modificada genéticamente según cualquier realización de la invención en un medio de crecimiento para producir un cultivo;
- 40 b. cultivar el cultivo; y
- c. recuperar la tiamina producida por dicho cultivo, y opcionalmente purificar la tiamina recuperada.
- 45 La invención proporciona además el uso de un transgén que codifica para un polipéptido que tiene actividad tiamina monofosfato fosfatasa (E.C.3.1.3.-) para potenciar la exportación de tiamina no fosforilada en una bacteria *Escherichia*. Según una realización para el uso del transgén que codifica para un polipéptido que tiene actividad tiamina monofosfato fosfatasa (E.C.3.1.3.-), la secuencia de aminoácidos de dicho polipéptido tiene al menos el 80% de identidad de secuencia con respecto a una secuencia seleccionada de entre:
- 50 a. el grupo que consiste en SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28 y 30;
- b. el grupo que consiste en SEQ ID NO: 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64 y 66; y
- 55 c. el grupo que consiste en SEQ ID NO: 68, 70, 72 y 74.
- La invención proporciona además el uso de una bacteria modificada genéticamente según cualquier realización de la invención para la producción de tiamina no fosforilada.

Breve descripción de los dibujos

Figura 1. Esquema que muestra los productos intermedios de la ruta de la tiamina en microorganismos y las etapas enzimáticas respectivas que conducen a la síntesis de tiamina (THI); tiamina monofosfato (TMP) y tiamina difosfato (TPP).

Figura 2. Diagrama de barras que muestra los niveles extracelulares e intracelulares de THI, TMP y TPP detectados

en cultivos de cepas de *E. coli* modificadas genéticamente transformadas con un vector que expresa uno de los 5 genes de fosfatasa diferentes: pBS93 con el gen de *A. thaliana* (At5g32470); pBS94 con el gen *rgsA* de la fosfatasa de Newmann de *S. aureus*; pBS95 con el gen de fosfatasa de *Pseudomonas*; pBS96 con el gen *pho3* de *S. cerevisiae* (YBR092C); y pBS97 con el gen *nudJ* de *E. coli*, en comparación con una cepa que comprende pBS92, un vector de control. De manera adicional, cada cepa comprende el vector pBS140 que comprende los genes *thiMD* y *thiCEFSGH* que expresan constitutivamente enzimas en la ruta de biosíntesis de tiamina; que van a compararse con la cepa de control transformada con vector vacío (pGEN49) que carece de los genes que codifican para las enzimas de la ruta. Las concentraciones intracelulares se calculan usando todo el volumen de cultivo, es decir el nivel indicado es la concentración eficaz si el metabolito intracelular se hubiera liberado en el volumen de cultivo.

Figura 3. Árbol de unión adyacente que muestra la distribución de TMP fosfatasas en 3 clados basándose en la alineación de secuencias de aminoácidos. La barra de escala de distancia indica cambios/residuo de aminoácido. El clado 1 comprende el gen de *A. thaliana* (At5g32470); el clado 2 comprende un gen de *Anaerotruncus colihominis* (WP_006874980) (Hasnain *et al.*, 2015); el clado 3 comprende un gen de *Syntrophomonas wolfei* (WP_011640074.1) (Hasnain *et al.*, 2015).

Figura 4. Árbol filogenético de TMP fosfatasas pertenecientes a (A) el clado 1; (B) el clado 2 y (C) el clado 3 basándose en la alineación de secuencias de aminoácidos. Cada miembro de los tres clados comprende un dominio de haloácido (HAD) que se sabe que cataliza la desfosforilación de la tiamina monofosfato. El clado 1 consiste en homólogos de At5G32470 y abarca *Viridiplantae*, y presenta un dominio de la familia TenA/Teni-4 (PFam: PF03070) en el extremo N-terminal y un HAD en la mitad C-terminal (véase su alineación en la figura 5). El clado 2 representa genes bacterianos hallados con frecuencia como fusión con otros genes de tiamina e incluyen el gen de *Anaerotruncus colihominis* (WP_006874980). El clado 3 representa un clado de genes bacterianos incluyendo *Syntrophomonas wolfei* (WP_011640074.1), que se halla con frecuencia con otros genes de tiamina. Las hojas del grupo externo se indican en cursiva y los representativos de cada clado que se han caracterizado como TMP fosfatasas se indican en negrita.

Figura 5. Esta figura muestra una alineación de las secuencias de proteínas de 11 TMP fosfatasas pertenecientes al clado 1 (AT5G32470.1 de *Arabidopsis thaliana*; BnaC08g46820D_emb|CDY62623.1| de *Brassica napus*; isoforma X1_ref|XP_006484613.1| similar a C17G9.12c de *Citrus sinensis*; isoforma X1_ref|XP_003536133.1| de LOC100788250 de *Glycine max*; CGZ_23571 [gb|KDP23738.1| de *Jatropha curcas*; LOC104108252_ref|XP_009615535.1| de *Nicotiana tomentosiformis*; Os08g0566000_ref|NP_001062539.1| de *Oryza sativa*; isoforma X1_ref|XP_009379735.1| de LOC103968121 de *Pyrus x bretschneideri*; PRUPE_ppa003431mg_>ref|XP_007199656.1| de *Prunus persica*; POPTR_0019s03960g_ref|XP_002325785.2| de *Populus trichocarpa*; LOC103653246_ref|XP_008678418.1| de *Zea mays*). Las secuencias de proteínas se alinearon usando la función "alineación Geneious" (parámetros convencionales) del software Geneious 6.0.6 (Biomatters Ltd.). El dominio TenA está subrayado con una línea discontinua y el dominio HAD está subrayado con una línea punteada. Los residuos de los sitios activos de TenA (basándose en la estructura cristalina de una proteína representativa de TenA, TenA de *B. subtilis*, n.º de ID de PDB: 1YAF, están marcados con * y los residuos que se muestra que forman enlaces de H con HMP están marcados con +. Los residuos conservados de un dominio HAD están marcados con # y los motivos DXD y DXXXD conservados están subrayados con una línea continua.

Figura 6. Diagrama de barras que muestra los niveles extracelulares e intracelulares de THI, TMP y TPP detectados en cultivos de cepas de *E. coli* con las siguientes modificaciones genéticas: células delecionadas para el gen *thiK* (*thiK*⁻) en las barras centrales; células que portan un gen *thiL* mutado (parcialmente inactivado) (*thiL*^{*}) en las barras a la derecha; y células con genes *thiK* y *thiL* de tipo natural en las barras a la izquierda. Las cepas de *E. coli* en el panel inferior comprenden el vector pBS93 que codifica para la TMP fosfatasa de *A. thaliana* (At5g32470); y las cepas en el panel superior comprenden el vector de control pBS92. De manera adicional, todas las cepas comprenden el vector pBS140 que comprende los genes *thiMD* y *thiCEFSGH* que expresan constitutivamente enzimas en la ruta de biosíntesis de tiamina.

Abreviaturas y términos

Número gi: (identificador de genInfo) es un número entero único que identifica una secuencia particular, independientemente de la fuente de base de datos, que se asigna por el NCBI a todas las secuencias procesadas en Entrez, incluyendo secuencias de nucleótidos de DDBJ/EMBL/GenBank, secuencias de proteínas de SWISS-PROT, PIR y muchos otros.

Identidad de secuencia de aminoácidos: el término "identidad de secuencia", tal como se usa en el presente documento, indica una medida cuantitativa del grado de homología entre dos secuencias de aminoácidos de una longitud sustancialmente igual. Las dos secuencias que van a compararse deben estar alineadas para dar un ajuste lo mejor posible, por medio de la inserción de huecos o, alternativamente, truncamiento en los extremos de las secuencias de proteínas. La identidad de secuencia puede calcularse como ((Nref-Ndif)/Nref), en la que Ndif es el número total de residuos no idénticos en las dos secuencias cuando se alinean y en la que Nref es el número de residuos en una de las secuencias. Los cálculos de identidad de secuencia se automatizan preferiblemente usando el programa BLAST, por ejemplo, el programa BLASTP (Pearson W.R y D.J. Lipman (1988))

(www.ncbi.nlm.nih.gov/cgi-bin/BLAST). La alineación de múltiples secuencias se realiza con el método de alineación de secuencias ClustalW con parámetros por defecto tal como se describe por Thompson J., *et al.*, 1994, disponible en <http://www2.ebi.ac.uk/clustalw/>.

- 5 Preferiblemente, los números de sustituciones, inserciones, adiciones o deleciones de uno o más residuos de aminoácido en el polipéptido en comparación con su polipéptido comparador son limitados, es decir no más de 1, 2, 3, 4, 5, 6, 7, 8, 9 ó 10 sustituciones, no más de 1, 2, 3, 4, 5, 6, 7, 8, 9 ó 10 inserciones, no más de 1, 2, 3, 4, 5, 6, 7, 8, 9 ó 10 adiciones y no más de 1, 2, 3, 4, 5, 6, 7, 8, 9 ó 10 deleciones. Preferiblemente, las sustituciones son sustituciones de aminoácidos conservativas: limitadas a intercambios dentro de los miembros del grupo 1: glicina, alanina, valina, leucina, isoleucina; grupo 2: serina, cisteína, selenocisteína, treonina, metionina; grupo 3: prolina; grupo 4: fenilalanina, tirosina, triptófano; grupo 5: aspartato, glutamato, asparagina, glutamina.

15 La bacteria modificada genéticamente para la producción de tiamina se deriva de un miembro de un género de bacteria en el que la biosíntesis de TPP está mediada a través de la ruta de thiL y no la ruta de thiN (figura 1).

Gen mutante: un gen mutante en el genoma de una célula bacteriana puede presentar una función reducida del gen y, por tanto, cuando el gen mutante codifica para un polipéptido, la mutación puede dar como resultado una pérdida/reducción de la expresión del polipéptido codificado. Alternativamente, el gen mutante puede codificar para un polipéptido mutante, y cuando el polipéptido es una enzima, la mutación puede dar como resultado una pérdida de actividad enzimática detectable en la célula bacteriana.

Gen nativo: gen endógeno en un genoma de célula bacteriana, homólogo a un microorganismo huésped.

Descripción detallada

25 La adaptación de *E. coli* para su uso como fábrica de células para la producción de tiamina requiere la provisión y/o manipulación de enzimas que puedan controlar el estado de fosforilación intracelular de la tiamina. Pueden tomarse dos enfoques para desplazar el equilibrio intracelular lejos de TPP y TMP y hacia la tiamina.

30 El primer enfoque implica la modificación (por ejemplo, regulación por disminución) de una o más de las cinasas que fosforilan la tiamina (por ejemplo, cinasa de rescate, ThiK, en *E. coli*); las cinasas que pirofosforilan la tiamina (ThiN en *Bacillus subtilis*; hallada comúnmente en eucariotas, pero no hallada en *E. coli*); y las cinasas que fosforilan TMP (thiL en *E. coli*). Este enfoque se ha usado en *B. subtilis* (documento WO2004106557).

35 El segundo enfoque, que no se ha descrito previamente, implica la expresión de enzimas recombinantes en la célula para desfosforilar TMP y, opcionalmente, TPP en combinación con la regulación por incremento de la ruta de biosíntesis de la tiamina. Este enfoque puede combinarse adicionalmente con una o más modificaciones del primer enfoque. Las TMP fosfatasas y las TPP fosfatasas adecuadas para este enfoque se limitan a aquellas que actúan selectivamente sobre TMP y TPP. Sin embargo, las fosfatasas mejor estudiadas son fosfatasas con nucleótidos promiscuos (por ejemplo, fosfatasa codificada por *nudJ* de *E. coli*) que actúan tanto sobre nucleósido di y trifosfato como sobre la TMP y la TPP estructuralmente similares. Tales fosfatasas son inadecuadas, dado que su actividad inespecífica desestabilizará el metabolismo celular.

45 En algunos organismos (mayormente eucariotas), la TMP se desfosforila antes de pirofosforilarse para dar TPP (véase la figura 1); mediante lo cual THI es un producto intermedio clave en la biosíntesis de TDP y TPP. Sin embargo, la identidad de la fosfatasa responsable ha sido objeto de especulación durante mucho tiempo, y su gen codificante se describió como un gen ausente (Goyer *et al.*, 2013). El metabolismo de la tiamina, *in vivo*, no sólo requiere fosforilasas para la desfosforilación de TDP y TMP, sino también para la eliminación de sus formas corruptas, oxitiamina y oxotiamina. En una búsqueda de TMP fosfatasas que puedan eliminar selectivamente las formas tóxicas de la tiamina en plantas, Goyer *et al.*, 2013 sugirieron que una secuencia de proteína de *Arabidopsis* (At5g32470) era un candidato viable, ya que la anotación de su secuencia predijo un dominio de haloácido deshalogenoasa (HAD) fusionado con una proteína de la familia de tiaminasa 2 (TenA), que en otros organismos se sabía que las proteínas TenA sirven para esta función selectiva para la eliminación de oxitiamina y oxotiamina. Sin embargo, la única actividad fosfatasa detectada en extractos de *Arabidopsis* por Komeda Y., *et al.*, (1988) correspondía a una proteína de 10 kDa, significativamente más pequeña que la proteína de 69 kDa codificada por el gen de AT5G32470.1. La presente invención se refiere a la provisión de una bacteria modificada genéticamente tal como se define en las reivindicaciones que puede producir y liberar cantidades potenciadas de tiamina no fosforilada en el entorno extracelular en comparación con el progenitor del que se derivó.

60 I. Bacteria modificada genéticamente para la producción y exportación de tiamina no fosforilada

La presente invención proporciona una célula bacteriana modificada genéticamente del género *Escherichia* que puede producir y exportar niveles potenciados de tiamina no fosforilada. La célula bacteriana de la invención comprende un transgén que codifica para una tiamina monofosfato fosfatasa (TMP fosfatasa que tiene EC.3.1.3.-) así como transgenes que codifican para polipéptidos que catalizan etapas en la ruta de la tiamina. La actividad de los polipéptidos que catalizan etapas en la ruta de la tiamina potencia la síntesis de ambos productos intermedios en

la ruta de la tiamina y de productos de la ruta de la tiamina (por ejemplo, TMP y TPP) en la célula bacteriana.

Un polipéptido que tiene actividad tiamina monofosfato fosfatasa (E.C.3.1.3.-) es un miembro de una familia de enzimas TMP fosfatasa que tiene un dominio de haloácido que cataliza la desfosforilación de tiamina monofosfato.

5 Los miembros de esta familia están codificados por genes pertenecientes a tres clados (ejemplo 3). La secuencia de aminoácidos del polipéptido que tiene actividad TMP fosfatasa tiene al menos el 70, el 75, el 80, el 85, el 90, el 95, el 96, el 98, el 100% de identidad de secuencia de aminoácidos con respecto a una secuencia seleccionada de una cualquiera de: SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28 y 30 codificada por genes del clado 1; SEQ ID NO: 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64 y 66 codificada por genes del clado 2; y
10 SEQ ID NO: 68, 70, 72 y 74 codificada por genes del clado 3. El gen del clado 1 de *A. thaliana* (AT5G32470.1) y sus ortólogos enumerados en la figura 4A; los genes del clado 2 de *Anaerotruncus colihominis* y *Dorea longicatena* y sus ortólogos enumerados en la figura 4B; y el gen del clado 3 de *Syntrophomonas wolfei* y sus ortólogos enumerados en la figura 4C, todos codifican para polipéptidos que tienen actividad TMP fosfatasa (E.C.3.1.3.-).

15 En una realización, la secuencia de aminoácidos del polipéptido que tiene actividad TMP fosfatasa tiene al menos el 60, el 65, el 70, el 75, el 80, el 85, el 90, el 95, el 96, el 98, el 100% de identidad de secuencia de aminoácidos con respecto a SEQ ID NO: 2; cuando el polipéptido se caracteriza por comprender un dominio similar a TenA, probablemente importante para el reconocimiento de tiamina, y un dominio HAD para la hidrólisis del enlace fosfato (ejemplo 2). Las proteínas con esta estructura característica pueden sobreexpresarse en una célula bacteriana de la
20 invención con el fin de desplazar el perfil de fosforilación de la tiamina hacia la tiamina.

En una realización, la secuencia de aminoácidos del polipéptido que tiene actividad TMP fosfatasa tiene al menos el 80, el 85, el 90, el 95, el 96, el 98, el 100% de identidad de secuencia de aminoácidos con respecto a una secuencia seleccionada de una cualquiera de: SEQ ID NO: 2, 32, 40 y 68.

25 Los polipéptidos que están codificados por los transgenes en la bacteria modificada genéticamente, y cuya actividad sirve para potenciar la síntesis de ambos productos intermedios y de productos de la ruta de la tiamina, son los siguientes:

30 b) un polipéptido que tiene actividad 4-amino-5-hidroximetil-2-metilpirimidina fosfato sintasa (HMP-P) (E.C.4.1.99.17); tal como un polipéptido con una secuencia de aminoácidos que tiene el 80, el 85, el 90, el 95 o el 100% de identidad de secuencia con respecto a SEQ ID NO: 76, 78, 80 y 82;

35 c) un polipéptido que tiene actividad tiamina fosfato sintasa (2.5.1.3), tal como un polipéptido con una secuencia de aminoácidos que tiene el 80, el 85, el 90, el 95 o el 100% de identidad de secuencia con respecto a SEQ ID NO: 84;

d) un polipéptido que tiene actividad adenilil transferasa de ThiS (2.7.7.73), tal como un polipéptido con una secuencia de aminoácidos que tiene el 80, el 85, el 90, el 95 o el 100% de identidad de secuencia con respecto a
40 SEQ ID NO: 86;

e) un polipéptido que tiene actividad portadora de azufre de ThiS (donadora inmediata de azufre en la formación de tiazol), tal como un polipéptido con una secuencia de aminoácidos que tiene el 80, el 85, el 90, el 95 o el 100% de identidad de secuencia con respecto a SEQ ID NO: 88.

45 f) un polipéptido que tiene actividad tiazol sintasa (2.8.1.10), tal como un polipéptido con una secuencia de aminoácidos que tiene el 80, el 85, el 90, el 95 o el 100% de identidad de secuencia con respecto a SEQ ID NO: 90;

g) un polipéptido que tiene actividad 2-iminoacetato sintasa (4.1.99.19; también denominado L-tirosina 4-metilfenol liasa), tal como un polipéptido con una secuencia de aminoácidos que tiene el 80, el 85, el 90, el 95 o el 100% de identidad de secuencia con respecto a SEQ ID NO: 92; o un polipéptido que tiene actividad glicina oxidasa (EC.1.4.3.19), tal como un polipéptido con una secuencia de aminoácidos que tiene el 80, el 85, el 90, el 95 o el 100% de identidad de secuencia con respecto a SEQ ID NO: 94, 96 y 98; y
50

h) un polipéptido que tiene actividad fosfohidroximetilpirimidina cinasa (2.7.4.7), tal como un polipéptido con una secuencia de aminoácidos que tiene el 80, el 85, el 90, el 95 o el 100% de identidad de secuencia con respecto a
55 SEQ ID NO: 100.

La bacteria modificada genéticamente puede caracterizarse adicionalmente por una modificación genética para reducir la expresión de uno o más de los tres genes *thiBPQ* endógenos nativos y, por tanto, reducir la exportación de
60 tiamina de la bacteria. Los tres genes *thiB*, *thiP* y *thiQ* codifican para una proteína de unión periplásmica de transportadores ABC de tiamina; una permeasa de transportadores ABC de tiamina; y una ATPasa de transportadores ABC de tiamina, respectivamente. Por ejemplo, la exportación de tiamina puede reducirse mediante la inactivación o delección (mediante noqueado génico) de uno o más de los tres genes *thiB* (o *thiP* o *thiQ*) en el genoma de la bacteria. Preferiblemente, la secuencia de aminoácidos del polipéptido codificado por el gen *thiB*
65 inactivado/delecionado tiene al menos el 80%, el 85, el 90, el 95 o el 100% de identidad de secuencia con respecto a SEQ ID NO: 102; el polipéptido codificado por el gen *thiP* inactivado/delecionado tiene al menos el 80%, el 85, el 90,

el 95 o el 100% de identidad de secuencia con respecto a SEQ ID NO: 104; y el polipéptido codificado por el gen *thiQ* inactivado/deleciónado tiene al menos el 80%, el 85, el 90, el 95 o el 100% de identidad de secuencia con respecto a SEQ ID NO: 106.

- 5 La bacteria modificada genéticamente puede caracterizarse adicionalmente por un transgén adicional que codifica para un polipéptido que tiene actividad hidroxietiltiazol cinasa (2.7.1.50), tal como un polipéptido con una secuencia de aminoácidos que tiene el 80, el 85, el 90, el 95 o el 100% de identidad de secuencia con respecto a SEQ ID NO: 108.
- 10 La expresión de un transgén que codifica para una tiamina monofosfato fosfatasa (TMP fosfatasa que tiene EC.3.1.3.-) en combinación con transgenes que codifican para polipéptidos que catalizan etapas en la ruta de la tiamina en la bacteria modificada genéticamente de la invención, no sólo potencia la producción de compuestos de tiamina no fosforilada, sino que adicionalmente dirige la ruta hacia la producción de tiamina no fosforilada (ejemplo 1). Tal como se muestra en el ejemplo 4, los niveles de tiamina extracelular aumentaron en 6 veces, mientras que la
- 15 producción total de compuestos de tiamina no fosforilada se potenció en un 20%, cuando se compara con células que carecen del transgén que codifica para TMP fosfatasa. Los niveles sorprendentemente altos de tiamina no fosforilada producida y liberada en el medio extracelular por la bacteria modificada genéticamente de la invención cumplen los requisitos clave para su uso como fábrica de células para la producción de tiamina.
- 20 En una realización adicional, la bacteria modificada genéticamente de la invención se caracteriza además por un gen *thiL* endógeno modificado genéticamente que expresa actividad tiamina fosfato cinasa (EC.2.7.4.16) reducida en comparación con el gen *thiL* endógeno parental. El gen *thiL* modificado genéticamente puede mutarse para expresar menores cantidades del polipéptido codificado; o puede mutarse para codificar para un polipéptido que tiene actividad tiamina fosfato cinasa (EC.2.7.4.16) reducida. Por ejemplo, cuando la secuencia de aminoácidos del
- 25 polipéptido codificado tiene al menos el 80% de identidad de secuencia con respecto a SEQ ID NO: 110 y con la condición de que la secuencia tenga una sustitución G133D de residuos de aminoácido (o una sustitución G→D en una posición equivalente basándose en la alineación de secuencia); entonces tendrá una actividad tiamina fosfato cinasa (EC.2.7.4.16) reducida. Tal como se muestra en el ejemplo 4, cuando las bacterias modificadas genéticamente que comprenden transgenes para la ruta de la tiamina sola se compararon con bacterias que
- 30 comprenden además un transgén que codifica para una TMP fosfatasa y un gen *thiL* mutante (actividad de THIL reducida), los niveles de producción extracelular de tiamina no fosforilada aumentaron en 7 veces, mientras que la producción total de compuestos de tiamina no fosforilada se potenció en un 24%.
- 35 La bacteria modificada genéticamente de la invención tiene un transgén que comprende un promotor ligado operativamente a una secuencia codificante que codifica para una tiamina monofosfato fosfatasa (TMP fosfatasa que tiene EC.3.1.3.-); así como transgenes que codifican para polipéptidos que catalizan etapas en la ruta de la tiamina. Un promotor puede ligarse operativamente a cada secuencia codificante del polipéptido de la ruta de la tiamina; o las secuencias codificantes respectivas pueden formar un operón que se liga operativamente a un único promotor. Los promotores adecuados incluyen tanto promotores constitutivos (por ejemplo, apFAB46 [SEQ ID NO:
- 40 147], apFAB70 [SEQ ID NO: 148], apFAB71 [SEQ ID NO: 149]), así como promotores inducibles (por ejemplo, promotor pBAD ara [SEQ ID NO: 150] cuando se expresa conjuntamente con el gen araC para la regulación (Guzman *et al.*, 1995) o promotor pLac con el sitio de operador lacO [SEQ ID NO: 151] (Norlander *et al.*, 1983)).
- 45 La bacteria modificada genéticamente según la invención, para la producción y exportación de tiamina no fosforilada, es un miembro de un género de bacteria en el que la biosíntesis de TPP está mediada a través de la ruta de thiL (figura 1) y no la ruta de thiN. El género de bacteria es *Escherichia*. Preferiblemente, la bacteria de la invención es una especie de *Escherichia*, por ejemplo, *Escherichia coli*.

50 II. Método para producir tiamina usando un microorganismo recombinante que expresa una tiamina monofosfato fosfatasa

- 55 La tiamina no fosforilada puede producirse y exportarse usando células de *Escherichia* de la invención mediante la introducción de las células en un medio de cultivo que comprende una fuente de carbono para la biosíntesis de una o más de tiamina no fosforilada, TMP y TPP; y finalmente la recuperación de la tiamina no fosforilada producida por el cultivo, tal como se ilustra en los ejemplos. Las células bacterianas de la invención producirán tiamina no fosforilada cuando se suministra con una fuente de carbono adecuada incluyendo glucosa, maltosa, galactosa, fructosa, sacarosa, arabinosa, xilosa, rafinosa, manosa y lactosa.

60 III. Método de detección de tiamina producida y exportada por un microorganismo recombinante que expresa una tiamina monofosfato fosfatasa

- 65 Los métodos para detectar y cuantificar la tiamina extracelular e intracelular producida por un microorganismo incluyen cromatografía de líquidos de alta resolución, en relación con un patrón de tiamina. Por ejemplo, los compuestos individuales de tiamina, tiamina, TMP y TPP, pueden medirse usando un procedimiento de ensayo de HPLC con tiocromo modificado descrito previamente (Chie *et al.*, 1999). En resumen, se añaden 100 µl de sobrenadante de cultivo o extractos intracelulares a 200 µl de acetato de potasio 4 M. Luego se oxida la muestra

mediante la adición de 100 µl de ferricianuro de potasio 3,8 mM recién preparado en NaOH 7 M. Se mezcla vigorosamente la mezcla y luego se extingue mediante la adición de 100 µl de H₂O₂ al 0,06% recién preparado en KH₂PO₄ saturado. Se neutralizan las muestras con HCl 6 M y se transfieren a viales de HPLC y se inyectan en una columna Supelcosil LC-18-T (15 cm X 4,6 mm, 3 µm) (n.º de ref. de Supelco 58970-U). La elución se realiza mediante un gradiente del 10%-35% de metanol (el 50%-25% de H₂O) en presencia del 40% de K₂HPO₄ 0,1 M (pH 6,6) e hidrogenosulfato de tetrabutilamonio 4 mM. Se mide la fluorescencia a 444 nm después de la excitación a 365 nm. El orden cronológico de elución de la columna es tiamina, TMP y TPP. Este procedimiento se utilizó para monitorizar la producción de tiamina tanto interna como externa durante la fermentación.

Alternativamente, puede realizarse la medición directa de tiamina y de los productos intermedios HMP y HET en el caldo de fermentación mediante cromatografía de muestras en una columna Phenomenex LUNA C18, usando un sistema de HPLC Agilent 1100 equipado con un inyector automático termostatzado y un detector por red de diodos (DAD). Las dimensiones de la columna son 150 x 4,6 mm, tamaño de partícula de 5 micrómetros. La temperatura de la columna se mantiene constante a 20°C. La fase móvil es una mezcla de 0,4 g de pentanosulfonato en agua, pH 2 (A) y metanol (B).

IV. Métodos para producir una bacteria modificada genéticamente para la producción y exportación de tiamina no fosforilada

Los vectores de integración y autorreplicación adecuados para la clonación e introducción de uno o más genes que codifican para uno o más polipéptidos que tienen una actividad enzimática asociada con la síntesis de tiamina en una bacteria *Escherichia* de la invención están disponibles comercialmente y se conocen por los expertos en la técnica (véase, por ejemplo, Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, segunda edición, Cold Spring Harbor Laboratory Press, 1989). Las células de una bacteria se modifican por ingeniería genética mediante la introducción en las células de ADN heterólogo. La expresión heteróloga de genes que codifican para uno o más polipéptidos que tienen una actividad enzimática asociada con la síntesis de tiamina en una bacteria se demuestra en el ejemplo 1 y 3.

Una molécula de ácido nucleico, que codifica para uno o más polipéptidos que tienen una actividad enzimática asociada con la síntesis de tiamina, puede introducirse en una célula huésped por medio de un vector de autorreplicación u opcionalmente integrarse en el genoma de la célula huésped usando métodos y técnicas que son convencionales en la técnica. Por ejemplo, las moléculas de ácido nucleico pueden introducirse mediante protocolos convencionales tales como transformación incluyendo transformación química y electroporación, transducción, bombardeo de partículas, etc. La expresión de la molécula de ácido nucleico que codifica para las enzimas de la divulgación también puede lograrse integrando la molécula de ácido nucleico en el genoma.

La modificación genética del gen *thiL* endógeno nativo en una bacteria se realiza para reducir la expresión de la actividad tiamina fosfato cinasa (EC.2.7.4.16) en comparación con el gen *thiL* endógeno parental. Por ejemplo, el gen *thiL* nativo puede mutarse para codificar para un polipéptido que tiene actividad tiamina fosfato cinasa (EC.2.7.4.16) reducida, tal como se describe en el ejemplo 4. La modificación genética de uno o más de los genes *thiBPQ* endógenos nativos en una bacteria se realiza para reducir la recaptación de la exportación de tiamina extracelular en la bacteria. La delección (noqueado) de uno cualquiera de los tres genes *thiB* (o *thiP* o *thiQ*) en el genoma de la bacteria (por ejemplo, cepa K12 de *E. coli*), usando métodos de recombinación convencionales (Datsenko KA, *et al.*; 2000), es suficiente para reducir la importación de tiamina en comparación con una bacteria parental. Por ejemplo, el constructo *thiB::Kan* (*thpA* renombrado como *thiB*) puede amplificarse por PCR a partir de la cepa JW0067 de la colección Keio (*thiP* a partir de JW0066; *thiQ* a partir de JW0065) (Baba *et al.*, 2006). El fragmento amplificado puede transformarse mediante electroporación en células huésped competentes (por ejemplo, *E. coli*) que portan los genes de recombinasa λ Red expresados a partir de un promotor inducible (pKD46) (Datsenko KA, *et al.*, 2000). Los integrantes exitosos se seleccionan en medio de kanamicina y se confirman mediante PCR de colonias. El casete de resistencia a kanamicina (*kanR*) puede eliminarse mediante la transformación de las células con el plásmido sensible a la temperatura *pcp20*, que expresa la recombinasa FLP. Las colonias resistentes a ampicilina pueden aislarse a 30°C y luego volver a cultivarse en línea de manera no selectiva a 42°C. La pérdida de *KanR* puede confirmarse mediante PCR de colonias.

Ejemplos

Ejemplo 1. Identificación de una fosfatasa específica de tiamina

Un precursor clave para la biosíntesis de TDP y TPP en muchos microorganismos y plantas es THI; que a su vez se deriva de TMP mediante la acción de una fosfatasa. Se identificaron cinco supuestas fosfatasas diferentes, que se predijo que tenían actividad fosfatasa hacia los sustratos TMP o TPP, basándose en predicción estructural/funcional. Las moléculas de ácido nucleico, cuyas secuencias de nucleótidos codificaban para cada fosfatasa, se clonaron individualmente en un vector de expresión dando los siguientes cinco vectores de expresión de fosfatasa (+ control de vector vacío), tal como se muestra en la tabla 1.

Tabla 1

Vector	Gen insertado	Proteína/SEQ ID NO. codificada
pBS92 (control)	-	-
pBS93	Gen sintético que codifica para el codón de fosfatasa de AT5G32470.1 de <i>Arabidopsis thaliana</i> optimizado para la expresión en <i>E. coli</i>	2
pBS94	Gen <i>rpsA</i> de fosfatasa de Newmann de <i>Staphylococcus aureus</i>	112
pBS95	Gen de fosfatasa de <i>Pseudomonas</i> identificado en una selección metagenómica	114
pBS96	Gen <i>pho3</i> de <i>Saccharomyces cerevisiae</i> (YBR092C)	116
pBS97	Gen <i>nudJ</i> de <i>E. coli</i>	118
pGEN49	Vector vacío usado para la construcción de pBS140	-
pGEN50	Vector vacío usado para la construcción de pBS140	-
pGEN51	Vector vacío usado como molde de PCR para la construcción de pBS92	-
pBS116	pGEN49 + operón de <i>thiC</i>	76
pBS117	pGEN50 + operón de <i>thiM</i>	108
pBS140	Vector que expresa los genes <i>thiCEFGHMD</i> de la ruta de la tiamina de <i>E. coli</i>	76, 84, 86, 88, 90, 92, 108, 100
pMA7-sacB	Vector que porta recombinasas usadas para la construcción de cepas para MAGE (Lennen <i>et al.</i> , 2015)	

1.1. Construcción de vectores

- 5 Se construyó el vector pBS92 usando el método de ensamblaje isotérmico de ADN de una sola etapa desarrollado por Gibson *et al.*, (2009). La estructura principal, que porta el casete SpecR y el origen de replicación SC101, se amplificó a partir de PZS4Int-tetR (http://www.expressys.com/main_tools.html) usando los cebadores oBS196 y oBS197, y se sintetizaron un casete de expresión génica que comprende un promotor apFAB70 (SEQ ID NO: 148) y un terminador apFAB381 (SEQ ID NO: 154) y luego se amplificaron usando los cebadores pBS194 y oBS195. Los
- 10 fragmentos de PCR purificados se ensamblaron mediante ensamblaje isotérmico en un solo recipiente usando la mezcla madre Gibson Assembly de NEB (n.º E2611) y el protocolo convencional según las instrucciones del fabricante (Gibson *et al.*, (2009). Los vectores ensamblados se transformaron en células DH10B electrocompetentes mediante electroporación y las células recuperadas se sembraron en placas de espectinomicina con caldo Luria (LB).
- 15 Se construyeron los vectores pBS93-97 mediante la amplificación de pBS92 con el par de cebadores oBS198 y oBS199; y el producto de PCR purificado se ensambló con el gen de fosfatasa respectivo usando la mezcla madre Gibson Assembly de NEB según las instrucciones del fabricante. Los genes de fosfatasa clonados en cada vector (tabla 1) se amplificaron de la siguiente manera:
- 20 pBS93: se usaron los cebadores oBS220, oBS221 para amplificar el gen sintético que codifica para la fosfatasa de AT5G32470.1 de *Arabidopsis thaliana* que se optimizó por codón para *E. coli*, y se sintetizó como dos bloques génicos;
- 25 pBS94: se usaron los cebadores oBS202, oBS203 para amplificar el gen *rsgA* a partir de ADN genómico de Newmann de *S. aureus*;
- pBS95: se usaron los cebadores oBS201, oBS200 para amplificar la secuencia codificante para un gen de fosfatasa ácida (EC.3.1.3.2) a partir de *Pseudomonas*.
- 30 pBS96: se usaron los cebadores oBS201, oBS202 para amplificar el gen *PHO3* de *S. cerevisiae*, que se optimizó por codón para *E. coli*, y se sintetizó como un bloque génico;
- pBS97: se usaron los cebadores oBS210, oBS211 para amplificar el gen *nudJ* (MG1655) a partir de ADN genómico de *E. coli*.
- 35

Cada uno de los vectores ensamblados se transformó en células DH10B electrocompetentes mediante electroporación y las células recuperadas se sembraron en placas con LB selectivas. Esta cepa DH10B huésped (F-mcrA Δ (mrr-hsdRMS-mcrBC) Φ 80dlacZ Δ M15 Δ lacX74 endA1 recA1 deoR Δ (ara,leu)7697 araD139 galU galK nupG rpsL λ -) comprende una delección cromosómica, que da como resultado la delección de tres genes (*thiBPQ* correspondientes a EcoGenes: EG11574, EG11573, EG11572) que codifican, respectivamente, para: proteína de unión periplásmica de transportadores ABC de tiamina; permeasa de transportadores ABC de tiamina; y ATPasa de transportadores ABC de tiamina.

Se construyó el vector pBS140 siguiendo un procedimiento de clonación jerárquica de dos etapas. En primer lugar, se ensamblaron el operón *thiC* y el operón *thiM* en los vectores de base pGEN49 (que comprende un promotor apFAB46 (SEQ ID NO: 147) y un terminador apFAB377 (SEQ ID NO: 153)) y pGEN50 (que comprende un promotor apFAB71 (SEQ ID NO: 149)) y un terminador apFAB378 (SEQ ID NO: 152)), respectivamente, y luego se combinaron los dos vectores resultantes para formar pBS140. Se amplificó el operón de *thiC* a partir de ADN genómico de MG1655 usando los cebadores oBS422 y oBS421. Los vectores pGEN49 y pGEN50 se amplificaron con los cebadores oGEN265 y oGEN266 y luego se ensamblaron con el operón de *thiC* y de *thiM*, respectivamente, usando ensamblaje isotérmico en un solo recipiente. Los vectores resultantes se denominaron pBS116 (pGEN49 + operón de *thiC*) y pBS117 (pGEN50 + operón de *thiM*). Los vectores ensamblados se transformaron en células DH10B electrocompetentes mediante electroporación y las células recuperadas se sembraron en placas de Kan con LB. La secuencia de los vectores resultantes se confirmó mediante secuenciación de Sanger.

Se purificó el vector pBS116 confirmado y se digirió durante > 8 h usando *Swa*I (n.º de NEB R0604). El ADN linealizado se purificó en gel. Se usó el vector pBS117 confirmado como molde para PCR con los cebadores oGEN182 y oGEN184. El fragmento de PCR resultante se purificó en gel. El vector amplificado y el operón de *thiM* se ensamblaron usando la mezcla madre Gibson Assembly de NEB, según las instrucciones del fabricante, y la mezcla se transformó en células DH10B electrocompetentes mediante electroporación y las células recuperadas se sembraron en placas de Kan con LB. La secuencia del plásmido pBS140 resultante se confirmó mediante secuenciación de Sanger.

Tabla 2. Lista de cebadores

Nombre de cebador	Secuencia	SEQ ID NO:
oBS194	GTCCTACTCAGGAGAGCGTTCACCGACAACCTCAGGAGAGC GTTCCACC	119
oBS195	CTTTCGTCTTCACCTCGAGGGAAATCAAATAGGCGTATCA CGAGGCC	120
oBS196	GATTTCCCTCGAGGTGAAGACGAAAG	121
oBS197	TGTCGGTGAACGCTCTCCTG	122
oBS198	ATTATTACTCGTGTGTTGTACAGAAAG	123
oBS199	CTAGTATTACCTCGCTATTAGTGACGTAATAGGAGGTAAGC	124
oBS200	CTATTACGTCACTAATAGCGAGGTAATACTAGATGTGCCAG CAGCATCCGC	125
oBS201	CGGAGGCCTTTCTGACAACACACGAGTAATAATCTAGTTTT CTAGAGGCAGCGC	126
oBS202	CTATTACGTCACTAATAGCGAGGTAATACTAGATGAAGACA GGTCGAATAGTG	127
oBS203	CGGAGGCCTTTCTGACAACACACGAGTAATAATTTAATATC TAACCTTTCTATTTG	128
oBS210	CCTATTACGTCACTAATAGCGAGGTAATACTAGATGTTTAA ACCGCACGTTACCG	129
oBS211	CTTTCTGACAACACACGAGTAATAATTTAGATGACACCCTTT GTAAAAGG	130
oBS220	GTACCTATAATGTGTGGATGTCCACCGCTTACCTCC	131
oBS221	CATCACCATCATCACCCTGAATTACTCGTGTGTTGTCA GAAAG	132
oGEN184	AAACCTCTTTATGTTGCAGTCG	133
oGEN182	AAATTCGCGAGTTCCACTAAGA	134

oGEN227	CCGCTTACCTCCTATTACGTCACTAATAGCTAAGGAGGTAA ATATGTCTGCAACAAACTGACCCGCC	135
oGEN264	CGGAGGCCTTTCTGACAACACACGAGTAATAATTCATAGTC TTTGCGAGGCG	136
oGEN265	ATTATTACTCGTGTGTTGTCAGAAAGGCCTCCG	137
oGEN266	ATTAGTGACGTAATAGGAGGTAAGCGGTGGG	138
oBS421	CGGAGGCCTTTCTGACAACACACGAGTAATAATTCACCACC AGGCGTGGAAG	139
oBS422	GCTTACCTCCTATTACGTCACTAATAGCTAAGGAGGTAAAT ATGCAAGTCGACCTGCTGG	140
oBS445	CGGAACAAAGCCGTGGATGTCCAACGTCATTGATAATGGC CCACGCGTGGTATCGCCGCAATGAGTTGCATATCGTAATA ATTGAGAAG	141
oBS446	CCAGAGCGCGTTAAGGCTCGTCCCATCGGAACAAAGCCGT GGATGTCCAACGTCATTGATAATGGCCCACGCGTGGTATC GCCGCCAATG	142
oBS456	CATGTGGCGAGTTCTCCCTG	143
oB5457	CAGGTAAACGGTACGCCAG	144
oB5521	GCAACTGTGCGATGGCGAAGC	145
oB5522	CCTGATCAACCGCCACCAC	146

1.2. Expresión de supuestos genes de TMP fosfatasa en células de *E. coli* modificadas genéticamente

- Se expresaron cada uno de los cinco genes de TMP fosfatasa candidatos: gen de *A. thaliana* (At5g32470); gen *rgsA* de fosfatasa de Newmann de *S. aureus*; gen de fosfatasa de *Pseudomonas*; gen *pho3* de *S. cerevisiae* (YBR092C); y gen *nudJ* de *E. coli* en una cepa huésped de *E. coli* que comprendía el vector pBS140. El vector pBS140 comprende los genes *thiMD* y *thiCEFSGH* de la ruta de la tiamina; cuya expresión en las células huésped conduce a la sobreexpresión de la ruta de biosíntesis de la tiamina.
- En primer lugar, las células de las cepas huésped de *E. coli* transformadas que expresan cada uno de los genes de fosfatasa candidatos (y las cepas huésped de *E. coli* de control) se cultivaron previamente en 400 µl de medio MOPS modificado (que comprende antibióticos para el mantenimiento de los vectores) en placas de cultivo de 96 pocillos profundos a 37°C, agitando a 300 rpm durante la noche; y luego se usaron los cultivos producidos para inocular el mismo volumen de medio hasta una D.O. a 600 nm de 0,0015. Se hicieron crecer estos cultivos durante 24 horas en las mismas condiciones; y luego se usaron para el análisis de tiamina. La D.O. a 600 nm se midió usando un lector de placas, y luego se convirtió a la D.O. a 600 nm correspondiente en una cubeta.

1.3. Extracción y detección de TPP, TMP y tiamina producidas por células de *E. coli* modificadas genéticamente

- Se recuperó la TPP, la TMP y la tiamina extracelular e intracelular en cada cultivo y se extrajeron de la siguiente manera: se sembraron 0,4 ml de cada cultivo a 4°C mediante centrifugación en la placa de cultivo a 4000 x g durante 5 minutos. Todas las etapas restantes se realizaron en hielo. Se retiraron suavemente 40 µl de sobrenadante para el análisis de TPP, TMP y tiamina extracelulares. Después de decantar el sobrenadante restante, se invirtió la placa de cultivo para retirar el medio residual y luego se agitó con vórtex. Se añadieron 100 µl de metanol helado de calidad para HPLC a cada pocillo de la placa de cultivo; y las células se agitaron con vórtex de nuevo. Después de la incubación en hielo durante un mínimo de 20 minutos, se sedimentó el residuo celular mediante centrifugación a 4000 x g durante 5 minutos. El sobrenadante se usó como extracto intracelular para el análisis posterior.

- Con el fin de detectar TPP, TMP y tiamina usando un detector de fluorescencia, los compuestos de tiamina producidos por cada cultivo se derivatizaron en tiocromos, que son fuertemente fluorescentes. Todas las etapas se realizan a temperatura ambiente. Se añadieron volúmenes de 40 µl de los extractos extracelulares e intracelulares a 80 µl de acetato de potasio 4 M y se mezcló mediante pipeteo. Se añadieron 40 µl de ferricianuro de potasio 3,8 mM recién preparado en NaOH 7 M y se mezcló. Se extinguió la reacción mediante la adición de 40 µl H₂O₂ al 0,06% recién preparado en KH₂PO₄ saturado. Los extractos se neutralizaron mediante la adición de 47 µl de HCl 6 M y luego se analizaron mediante HPLC tal como se describe a continuación. Todos los compuestos derivatizados se cuantificaron usando curvas patrón de fluorescencia de patrones de TPP, TMP y tiamina recién preparados que se derivatizaron con tiacromos en paralelo con los extractos analizados.

Se usó HPLC para cuantificar el contenido intracelular y extracelular de tiamina no fosforilada; TMP y TPP en cada

cultivo mediante un método adaptado del método descrito por Schyns G., *et al.*, (2005). Columna: Hypersil Gold, tamaño de partícula de 3 μm , dimensiones: 150 x 2,1 mm (n.º de producto de Thermo: 25003-152130) con precolumna del mismo material. Tampón A: K_2HPO_4 10 mM, pH: 7, hidrogenosulfato de t-butilamonio 4 mM. Tampón B: metanol. Tampón C: agua destilada. Gradiente: concentración constante de tampón A al 40%. 0-2 min: concentración constante de B al 10%; 2-6 min: gradiente de B desde el 10 hasta el 30%; 6-8 min: concentración constante de B al 30%; 8-8,1 min: gradiente de B desde el 30 hasta el 10%; 8,1-12 min: concentración constante de B al 10%. Velocidad de flujo: 0,8 ml/min. Volumen de inyección: 2 μl para muestras intracelulares, 5 μl para muestras extracelulares. Orden de elución: tiamina, TMP, TPP. Detección usando un detector de fluorescencia con $\lambda_{\text{excitación}}$: 365 nm y $\lambda_{\text{emisión}}$: 444 nm.

1.4. La proteína AT5G32470.1 de *A. thaliana* expresada en células de *E. coli* modificadas genéticamente muestra actividad TMP fosfatasa

La composición de los compuestos de tiamina producidos por células de *E. coli* modificadas genéticamente, modificadas por ingeniería genética para sobreexpresar la ruta de la tiamina, y que expresan uno de los cinco genes de fosfatasa diferentes, se muestra en la figura 2. Las únicas células que muestran un desplazamiento significativo en el perfil de fosforilación de la tiamina hacia tiamina no fosforilada, con respecto al de células que comprenden el vector de control pBS92, fueron aquellas que expresan un gen que codifica para la proteína AT5G32470.1 de *A. thaliana*. La TMP no fue detectable en estas células; la cantidad de TPP se redujo ligeramente; sobre cuya base puede deducirse que la proteína AT5G32470.1 de *A. thaliana* tiene actividad TMP fosfatasa, y también puede tener actividad de TPP. La actividad TMP fosfatasa en células que expresan TMP fosfatasa de *A. thaliana* dieron como resultado un notable aumento tanto en la producción de compuestos de tiamina extracelulares totales como en tiamina no fosforilada extracelular, que representó el 80% de los compuestos de tiamina totales detectados en el cultivo. Los compuestos de tiamina apenas fueron detectables en cultivos de células de *E. coli* transformadas con los vectores vacíos (pGEN49 y pBS92) que no expresaron ninguna TMP fosfatasa ni enzimas de la ruta de la tiamina. Una comparación de los niveles de compuestos de tiamina no fosforilada producidos por células que expresan enzimas de la ruta de la tiamina, solas o junto con TMP fosfatasa, revela su efecto sinérgico sobre la producción de compuestos de tiamina.

Ejemplo 2. Alineación y anotación estructural de la proteína de *A. thaliana* (AT5G32470.1)

La alineación de secuencia y la anotación estructural de la proteína de *A. thaliana* (AT5G32470.1) revela dos dominios estructurales y funcionales. El primer dominio es un dominio similar a TenA ubicado entre los aminoácidos 85-292, que se basa en la alineación con la proteína TenA de *Bacillus subtilis*, se le asigna la responsabilidad del reconocimiento de sustratos de tiamina fosforilada. La estructura cristalina de la proteína TenA de *B. subtilis* se ha resuelto sola y en forma de complejo con el precursor de tiamina HMP (Toms *et al.*, 2005). TenA de *B. subtilis* es una aminopirimidina aminohidrolasa que cataliza la hidrólisis de 4-amino-5-aminometil-2-metilpirimidina para dar 4-amino-5-hidroximetil-2-metilpirimidina (HMP). Dado que HMP corresponde a la mitad de la molécula de tiamina, se predice que los residuos del sitio de unión a tiamina en la proteína de *A. thaliana* corresponden a los sitios de unión a HMP en TenA de *B. subtilis*. Una alineación de secuencia muestra que los residuos clave, que forman contactos por enlaces de H con HMP y que recubren el sitio activo, se conservan entre TenA de *B. subtilis* y AT5G32470.1 (que se identifican en la figura 5). TenA de *B. subtilis* puede hidrolizar la tiamina, aunque a una velocidad mucho más lenta que HMP, reforzando el argumento de que el dominio de TenA es responsable de la unión de tiamina.

Un segundo dominio en la proteína de *A. thaliana* pertenece a una familia conservada de dominios de hidrolasa similares a haloácido deshalogenasa (superfamilia similar a HAD) (<http://www.ncbi.nlm.nih.gov/Structure/cdd/cddsrv.cgi?uid=277525>). Esta superfamilia incluye carbono hidrolasas y fosfato hidrolasas. Se sabe que los miembros de la superfamilia de dominios similares a HAD catalizan una reacción de sustitución nucleófila en el carbono o fósforo en cuestión. Una característica distintiva de este dominio HAD (código de PFAM PF00702) es la presencia de varios residuos conservados que forman parte del pliegue de Rossmann (Burroughs, A.M., *et al.*, 2006). Un aspartato conservado en el extremo de la lámina 1, una serina o treonina conservada en el extremo de la lámina 2, una lisina en una estructura denominada caperuza C2 y un aspartato conservado en la lámina 4, que juntos funcionan en catálisis. Por consiguiente, el dominio similar a HAD en la proteína de *A. thaliana* (figura 5) se le asignó actividad fosfato hidrolasa. Estos dominios funcionales conservados permiten la identificación de polipéptidos que tienen actividad TMP fosfatasa.

Ejemplo 3. Identificación y anotación estructural de familias de proteínas TMP fosfatasas

Los miembros de la familia de proteínas TMP fosfatasas, que pueden potenciar los niveles de tiamina en una célula bacteriana, se identificaron basándose en la secuencia, homología estructural y funcional, empleando los siguientes criterios y herramientas: se sabía empíricamente que cuatro genes codificaban para enzimas con una actividad de hidrólisis de tiamina monofosfato, ya sea promiscua o fisiológica. Los tres genes eran el gen de AT5G32470.1, el primero identificado en el presente documento como fosfatasa, y tres genes notificados recientemente (Hasnain *et al.*, 2015). Las secuencias que eran homólogas a estos genes, además de pertenecer a cualquier grupo de proteínas de fusión TenA-HAD, o alternativamente fusionadas a un gen de la ruta de la tiamina o presentes en los operones que comprenden genes de la ruta de la tiamina, se tomaron como un fuerte indicador de conservación de su función

ya que codifican para una TMP fosfatasa. Las secuencias que codifican para un dominio de haloácido, que es un rasgo característico de proteínas que catalizan desfosforilaciones, se tomaron como un indicador adicional de la función de TMP fosfatasa. El conjunto de datos de semillas de PF012710 se usó como un conjunto de datos de base para la diversidad de proteínas de HAD. En el caso de proteínas que se componen de múltiples dominios, se retiraron los dominios adicionales basándose en una alineación preliminar con secuencias de dominio HAD conocidas. Los conjuntos de datos de los genes de interés se eligieron manualmente y las cabeceras se cambiaron con unas pequeñas instrucciones de programa de Python. Luego se alinearon los genes con el conjunto de datos de Pfam y, por consiguiente, se eligieron los grupos externos para cada grupo. Las secuencias se alinearon independientemente para cada grupo con Muscle y se cortaron con Gblocks con parámetros permisivos, y las alineaciones resultantes se usaron para la inferencia de árbol mediante la máxima probabilidad en un modelo WAG con una distribución CAT.

Las proteínas de dominio de haloácido de TMP fosfatasa caracterizadas que catalizan la desfosforilación de tiamina monofosfato se encuentran dentro de tres clados (figura 3), indicativos de su evolución independiente paralela.

El clado 1 comprende homólogos de At5G32470 y abarca *Viridiplantae* (figura 4A), y todos sus miembros presentan un dominio de la familia TenA/Ten1-4 (PFam: PF03070) en el extremo N-terminal (figura 5). Los miembros de este clado comprenden las proteínas: NP-198287.3 de *Arabidopsis thaliana* (AT5G32470.1) [SEQ ID NO: 2]; XP_009379735.1 de *Pyrus x bretschneideri* [SEQ ID NO: 4]; CDY62623.1 de *Brassica napus* [SEQ ID NO: 6]; XP_003536133.1 de *Glycine max* [SEQ ID NO: 8]; XP_009615535.1 de *Nicotiana tomentosiformis* [SEQ ID NO: 10]; XP_002325785.2 de *Populus trichocarpa* [SEQ ID NO: 12]; KDP23738.1 de *Jatropha curcas* [SEQ ID NO: 14]; XP_006484613.1 de *Citrus sinensis* [SEQ ID NO: 16]; XP_007199656.1 de *Prunus persica* [SEQ ID NO: 18]; *Phoenix dactylifera*_XP_008796407 [SEQ ID NO: 20]; XP_008678418.1 de *Zea mays* [SEQ ID NO: 22]; NP_001062539.1 de *Oryza sativa* [SEQ ID NO: 24]; *Picea sitchensis*_ABR16455 [SEQ ID NO: 26]; *Physcomitrella patens*_XP_001769831 [SEQ ID NO: 28]; y *Selaginella moellendorffii*_XP_002990363 [SEQ ID NO: 30].

El clado 2 comprende genes bacterianos, muchos de los cuales se fusionan con genes de la ruta de la tiamina (figura 4B). El clado 2 incluye un gen de *Anaerotruncus colihominis* (WP_006874980) caracterizado como una TMP fosfatasa (Hasnain et al., 2015). Los miembros de este clado comprenden las proteínas: *Anaerotruncus colihominis*_WP_006874980 [SEQ ID NO: 32]; *Eubacterium ventriosum*_WP_005362972 [SEQ ID NO: 34]; *Coprococcus eutactus*_(ATCC_27759)_EDP27707 [SEQ ID NO: 36]; *Ruminococcus bromii*_L2-63_CBL14666 [SEQ ID NO: 38]; *Dorea longicatena*_(DSM_13814)_EDM62146 [SEQ ID NO: 40]; bacteria *Lachnospiraceae*_EPC05128 [SEQ ID NO: 42]; *Fusicatenibacter*_CUQ30753 [SEQ ID NO: 44]; *Clostridium*_sp._ERI68966 [SEQ ID NO: 46]; *Eubacterium hallii*_EEG35494 [SEQ ID NO: 48]; *Eubacterium*_sp._CDB67556 [SEQ ID NO: 50]; *Lachnospira pectinoschiza*_CUQ76318 [SEQ ID NO: 52]; bacteria *Peptostreptococcaceae*_WP_009530263 [SEQ ID NO: 54]; bacteria *Peptostreptococcaceae*_WP_009527854 [SEQ ID NO: 56]; *Atopobium*_sp._WP_035427744 [SEQ ID NO: 58]; *Atopobium parvulum*_WP_035433109 [SEQ ID NO: 60]; *Atopobium rimae*_WP_003148415 [SEQ ID NO: 62]; *Olsenella uli*_WP_013251930 [SEQ ID NO: 64]; y *Atopobium minutum*_KRN55115 [SEQ ID NO: 66].

El clado 3 comprende otro grupo de genes bacterianos, muchos de los cuales se encuentran adyacentes a genes de la ruta de la tiamina (figura 4C); e incluye una TMP fosfatasa caracterizada codificada por un gen de *Syntrophomonas wolfei* (WP_011640074.1) (Hasnain et al., 2015). Los miembros de este clado comprenden las proteínas: *Syntrophomonas wolfei*_WP_011640074 [SEQ ID NO: 68]; *Desulfotobacterium hafniense*_WP_018212876 [SEQ ID NO: 70]; *Pelotomaculum thermopropionicum*_WP_012032097 [SEQ ID NO: 72]; y *Desulfotomaculum ruminis*_WP_013840216 [SEQ ID NO: 74].

Ejemplo 4. Cepas de *E. coli* modificadas genéticamente modificadas por ingeniería genética para la producción potenciada de tiamina

Con el fin de desplazar el equilibrio más hacia la tiamina no fosforilada (THI) en una cepa de producción de tiamina, se combinó la expresión de una TMP fosfatasa con mutaciones en los genes endógenos de tiamina o TMP cinasas, *thiK* y *thiL*. Dado que la TPP es esencial para el crecimiento celular, y dado que la única enzima en *E. coli* que puede producir TPP a partir de TMP está codificada por el gen *thiL*, las cepas completas noqueadas para *thiL* no son viables a menos que se les proporcione un suministro de TPP, el cual no es económico. En cambio, la actividad cinasa expresada en las células huésped disminuyó al introducir una mutación puntual (codón 133 de GGT a GAC) en el gen *thiL* que provoca la sustitución G133D en la enzima codificada. Se sabe que esta mutación puntual disminuye la actividad cinasa codificada por el gen mutante *thiL927* en *Salmonella typhimurium* (Webb E., et al., 1997). Esta mutación se introdujo en células de la cepa huésped de *E. coli* usando modificación automatizada múltiple por ingeniería genética del genoma (MAGE, *Multiplex Automated Genome Engineering*).

4.1. Construcción de una cepa noqueada para *thiK* (BS131)

Se noqueó el gen *thiK* en la cepa huésped DH10B de *E. coli* para producir la cepa BS131 usando métodos de

recombinación con λ -RED (Datsenko K., *et al.*, 2000) mediante las siguientes etapas: se amplificó por PCR un constructo *thiK::Kan* a partir de la cepa JW1092 de la colección Keio (Baba T., *et al.*, 2006). Aunque el gen *thiK* amplificado se denominó originalmente *ycfN*, en la base de datos Keio, este corresponde a *thiK* (Melnick J., *et al.*, 2004). El fragmento amplificado se transformó mediante electroporación en células competentes de DH10B de *E. coli* que portaban los genes de la recombinasa λ Red expresados a partir de un promotor inducible (pKD46) (Datsenko *et al.*, 2000). Los integrantes exitosos se seleccionaron en medio de kanamicina y se confirmaron mediante PCR de colonias usando los cebadores oBS521 y oBS522. Se eliminó el casete de resistencia a kanamicina transformando las células con el plásmido sensible a la temperatura *pcp20*, que expresa la recombinasa FLP. Las colonias resistentes a ampicilina se aislaron a 30°C y luego volvieron a cultivarse en línea de manera no selectiva a 42°C. La pérdida de KanR se confirmó mediante PCR de colonias con los cebadores oBS521 y oBS522.

4.2. Construcción de *thiL* con G133D (cepa de *thiL** (BS182))

El gen *thiL* se inactivó parcialmente mediante mutación en la cepa huésped DH10B de *E. coli* para producir la cepa BS182 usando MAGE, que comprende la sobreexpresión transitoria de ADN adenina metilasa, tal como se describe por Lennen *et al.* (2015). En resumen: se introdujo el plásmido pMA7sacB en células de la cepa DH10B de *E. coli*. Se hizo crecer la cepa resultante en condiciones inductoras (arabinosa al 0,2%) hasta la mitad de la fase logarítmica y se prepararon células electrocompetentes usando condiciones convencionales. Las células competentes se sometieron a electroporación con una mezcla 1:1 de oligonucleótidos oBS445 y oBS446, que introducen ambos una mutación en la copia genómica del gen *thiL* que codifica para una mutación por sustitución G133D. Después del rescate, se repitió este procedimiento una vez más antes de sembrar en placa las células, y los mutantes exitosos se identificaron mediante secuenciación con los cebadores oBS456 y oBS457.

4.3. La producción de tiamina se potencia en *E. coli* modificada genéticamente que tiene actividad tiamina cinasa reducida (*thiL**) combinada con la expresión de la fosfatasa de *A. thaliana* (AT5G32470.1)

La cepa de *E. coli* noqueada para *thiK*, BS131, y la cepa de *E. coli* mutada con *ThiL**, BS182, así como la cepa parental DH10B, se transformaron cada una con el vector pBS140 que comprende los genes *thiMD* y *thiCEFSGH* de la ruta de la tiamina, o bien sola o bien en combinación con el vector pBS93 que codifica para la TMP fosfatasa de *A. thaliana* (AT5G32470.1). Estas cepas modificadas genéticamente se cultivaron tal como se describe en el ejemplo 1.2, y el perfil de fosforilación de la tiamina y el contenido de tiamina de todos los cultivos resultantes se determinó mediante HPLC tal como se describe en el ejemplo 1.3.

Tal como se observa en la figura 5 (y la tabla 3), la expresión de la TMP fosfatasa de *A. thaliana* sola convierte la mayor parte de TMP intracelular y extracelular en tiamina. La mutación de *thiL* sola disminuye significativamente la cantidad de TPP con respecto a TMP y tiamina. Una combinación de estas dos modificaciones genéticas da como resultado que casi toda la tiamina celular se convierta en el estado desfosforilado (el 90% del total, en comparación con el 14% para las cepas de cinasa de tipo natural). Esto corresponde a un aumento de 6,8 veces en la cantidad de THI producida. Además, la cantidad de tiamina no fosforilada extracelular también aumenta en gran medida (7 veces), lo que es muy deseable para facilitar el procesamiento aguas abajo del producto. Finalmente, el título global de THI + TMP + TPP también se aumenta en un 24%. En cepas que expresan la TMP fosfatasa de *A. thaliana*, los niveles intracelulares de compuestos de tiamina se reducen sistemáticamente; lo que a su vez puede reducir la inhibición de retroalimentación potencial en la ruta de la tiamina, explicando de ese modo la producción potenciada de compuestos de amina totales. La delección de *thiK*, sin embargo, sólo potencia ligeramente la producción de THI cuando se combina con la expresión de la TMP fosfatasa de *A. thaliana*. Por consiguiente, la desfosforilación de tiamina catalizada por la TMP fosfatasa parece compensar la fosforilación de tiamina catalizada por la tiamina cinasa (ThiK), de modo que la delección del gen *ThiK* tiene un efecto menor.

Tabla 3. Producción de compuestos de tiamina no fosforilada en cepas de *E. coli* modificadas genéticamente

Cepa de <i>E. coli</i>	Modificación del genoma/vectores	Producción de THI + TMP + TPP [μ M]	Producción de THI [μ M]	Producción de THI + TMP + TPP extracelulares [μ M]	Producción de THI extracelular [μ M]
BS167	Cepa de tipo natural pBS140 pBS92	3,97	0,57	2,46	0,51
BS168	Cepa de tipo natural pBS140 pBS93	4,79	3,39	3,53	3,02
BS233	<i>thiL</i> * pBS140 pBS92	4,02	0,78	3,22	0,71
BS234	<i>thiL</i> * pBS140 pBS93	4,91	4,43	4,26	4,05

Bibliografía

- 5 Baba T, Ara T, Hasegawa M, *et al.* (2006) Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol.;2:2006.0008
- 10 Burroughs, A.M. *et al.*, 2006. Evolutionary Genomics of the HAD Superfamily: Understanding the Structural Adaptations and Catalytic Diversity in a Superfamily of Phosphoesterases and Allied Enzymes. J. Mol. Biol., 361, págs.1003-1034.
- Datsenko KA, Wanner BL. (2000) One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc Natl Acad Sci U S A.;97(12):6640-5.
- 15 Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison C a, Smith HO., (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods., 6(5):343-5.
- Goyer A, Hasnain G, Frelin O, Ralat M a, Gregory JF, Hanson AD. (2013) A cross-kingdom Nudix enzyme that pre-empt damage in thiamin metabolism. Biochem J. 454(3):533-42.
- 20 Guzman, L., Belin, D. & Carson, M.J., 1995. Tight Regulation, Modulation, and High-Level Expression by Vectors Containing the Arabinose P BAD Promoter. Journal of bacteriology, 177(14):4121-4130.
- Hasnain G., *et al.*, (2015) Bacterial and plant HAFD enzymes catalyze a missing phosphatase step in thiamin diphosphatase biosynthesis. Biochemical Journal, 2015 DOI: 10.1042/BJ20150805
- 25 Komeda Y, Tanaka M, Nishimune T. (1988) A th-1 Mutant of Arabidopsis thaliana is Defective for a Thiamin-Phosphate-Synthesizing Enzyme: Thiamin Phosphate Pyrophosphorilase. Plant Physiol. 88(2):248-50.
- 30 Lennen, R.M. *et al.*, (2015) Transient overexpression de DNA adenine methylase enables efficient and mobile genome engineering with reduced off-target effects. Nucleic acids research, págs.1-14.
- Melnick J, Lis E, Park J-H, *et al.* (2004) Identification of the two missing bacterial genes involved in thiamine salvage: thiamine pyrophosphokinase and thiamine kinase. J Bacteriol.:186(11):3660-2
- 35 Norrander J, Tomas Kempe, Joachim Messing, 1983. Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis, Gene, 26(1), págs. 101-106.
- Schyns G, Geng Y, Barbosa TM, Henriques A, Perkins JB. (2005) Isolation and Characterization of New Thiamine-Deregulated Mutants of Bacillus subtilis. J Bacteriol.; 187(23):8127-8136.
- 40 Webb E, Downs D. (1997) Characterization of thiL, Encoding Thiamin-monophosphate Kinase, in Salmonella typhimurium. J Biol Chem.;272(25):15702-15707

Lista de secuencias

- 45 <110> Biosyntia ApS
- <120> Fábrica células bacterianas modificadas genéticamente para la producción de tiamina
- 50 <130> P2050PC00
- <150> documento EP15201200.1
- <151> 18-12-2015
- 55 <160> 157
- <170> PatentIn versión 3.5
- <210> 1
- 60 <211> 1854
- <212> ADN
- <213> *Arabidopsis thaliana*
- <220>
- 65 <221> CDS
- <222> (1)..(1854)

ES 2 806 732 T3

<223> gen de *Arabidopsis thaliana* que codifica para TMP fosfatasa [AT5G32470.1]

<400> 1

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

ES 2 806 732 T3

ttg aaa atg cac gac tca ttt gta cag gat tgg gat tta gac atc aac Leu Lys Met His Asp Ser Phe Val Gln Asp Trp Asp Leu Asp Ile Asn 145 150 155 160	480
aaa gaa gta agt gtt aac tca gca act ttg aga tac act gag ttc ttg Lys Glu Val Ser Val Asn Ser Ala Thr Leu Arg Tyr Thr Glu Phe Leu 165 170 175	528
tta gct aca gca tcc gga aaa gta gaa gga tgc aaa gct ccc ggc atg Leu Ala Thr Ala Ser Gly Lys Val Glu Gly Cys Lys Ala Pro Gly Met 180 185 190	576
ctt gat act cca ttt gaa aaa aca aaa gtt gct gcc tac acg ctt ggt Leu Asp Thr Pro Phe Glu Lys Thr Lys Val Ala Ala Tyr Thr Leu Gly 195 200 205	624
gct gtg aca cct tgc atg cgg ttg tat gcc ttt ctc ggt aag gag ttt Ala Val Thr Pro Cys Met Arg Leu Tyr Ala Phe Leu Gly Lys Glu Phe 210 215 220	672
gga tca ctt ctt gat ctg agt gat gtg aac cat ccc tac aag aaa tgg Gly Ser Leu Leu Asp Leu Ser Asp Val Asn His Pro Tyr Lys Lys Trp 225 230 235 240	720
atc gat aat tat tct agt gat gct ttc cag gca tca gcc aag caa act Ile Asp Asn Tyr Ser Ser Asp Ala Phe Gln Ala Ser Ala Lys Gln Thr 245 250 255	768
gaa gac ttg ctt gag aag ctt agt gtc tct atg act ggt gaa gaa ttg Glu Asp Leu Leu Glu Lys Leu Ser Val Ser Met Thr Gly Glu Glu Leu 260 265 270	816
gac ata att gaa aaa ttg tat caa cag gct atg aaa ctt gaa gta gag Asp Ile Ile Glu Lys Leu Tyr Gln Gln Ala Met Lys Leu Glu Val Glu 275 280 285	864
ttc ttc cat gcc cag cca ctt gcc cag cct acc ata gtt cca ctg ctc Phe Phe His Ala Gln Pro Leu Ala Gln Pro Thr Ile Val Pro Leu Leu 290 295 300	912
aag aac cac tca aaa gat gat ctg gtg atc ttt tct gat ttt gat ctg Lys Asn His Ser Lys Asp Asp Leu Val Ile Phe Ser Asp Phe Asp Leu 305 310 315 320	960
act tgc acc gtt gtg gat tct tct gct att tta gcg gaa ata gca att Thr Cys Thr Val Val Asp Ser Ser Ala Ile Leu Ala Glu Ile Ala Ile 325 330 335	1008
gta act gcc cca aaa gat gaa caa agt cga tct gga caa caa att cat Val Thr Ala Pro Lys Asp Glu Gln Ser Arg Ser Gly Gln Gln Ile His 340 345 350	1056
cgg atg ctc tca tct gac ctt aag aac acc tgg aat cta ctt tct aaa Arg Met Leu Ser Ser Asp Leu Lys Asn Thr Trp Asn Leu Leu Ser Lys 355 360 365	1104
caa tac aca gag cat tat gaa gaa tgc ata gag agt att ctg aat aaa Gln Tyr Thr Glu His Tyr Glu Glu Cys Ile Glu Ser Ile Leu Asn Lys 370 375 380	1152
aag aaa gcg gac aag ttt gac tat gaa ggt tta tgt aaa gca cta gag Lys Lys Ala Asp Lys Phe Asp Tyr Glu Gly Leu Cys Lys Ala Leu Glu 385 390 395	1200

ES 2 806 732 T3

385	390	395	400	
cag ctt tca gat ttt gag aaa gag gca aat aat cga gtg att gag tct				1248
Gln Leu Ser Asp Phe 405	Glu Lys Glu Ala Asn 410	Arg Val Ile Glu Ser 415		
ggt gta ctc aaa ggc ctg aat ctt gaa gac att aag cgc gct ggg gaa				1296
Gly Val Leu Lys Gly Leu Asn Leu Glu Asp Ile Lys Arg Ala Gly Glu	420	425	430	
agg tta atc ctt caa gat gga tgc atc aat gtc ttc cag aaa att tta				1344
Arg Leu Ile Leu Gln Asp Gly Cys Ile Asn Val Phe Gln Lys Ile Leu	435	440	445	
aag act gag aat ctg aat gca gaa ctt cat gtg ctt tcc tat tgt tgg				1392
Lys Thr Glu Asn Leu Asn Ala Glu Leu His Val Leu Ser Tyr Cys Trp	450	455	460	
tgt ggt gac ctc atc agg gca gcc ttt tct gca ggc gga gta gat gca				1440
Cys Gly Asp Leu Ile Arg Ala Ala Phe Ser Ala Gly Gly Val Asp Ala	465	470	475	480
gtg gaa gta cat gca aat gaa ttc aca ttt gag gaa tcc atc tcg act				1488
Val Glu Val His Ala Asn Glu Phe Thr Phe Glu Glu Ser Ile Ser Thr	485	490	495	
ggt gag atc gaa aga aag gtg gaa tcc cca att aac aaa gct caa cag				1536
Gly Glu Ile Glu Arg Lys Val Glu Ser Pro Ile Asn Lys Ala Gln Gln	500	505	510	
ttc aaa agt atc cta caa aac aga aag aat gag aac aat aag aaa agt				1584
Phe Lys Ser Ile Leu Gln Asn Arg Lys Asn Glu Asn Asn Lys Lys Ser	515	520	525	
ttc ttg agt gtg tat att gga gat tgc gta ggt gac ttg ctg tgt ctc				1632
Phe Leu Ser Val Tyr Ile Gly Asp Ser Val Gly Asp Leu Leu Cys Leu	530	535	540	
ctc gaa gca gat ata gga ata gtg gtt agc tct agc tgc agt ctc agg				1680
Leu Glu Ala Asp Ile Gly Ile Val Val Ser Ser Ser Ser Ser Leu Arg	545	550	555	560
aga gtt gga agc cat ttt ggg gtc tca ttt gtg cct ttg ttt tct gga				1728
Arg Val Gly Ser His Phe Gly Val Ser Phe Val Pro Leu Phe Ser Gly	565	570	575	
atc gtc cag aaa cag aaa caa cac act gaa gaa tca tca tca tca gca				1776
Ile Val Gln Lys Gln Lys Gln His Thr Glu Glu Ser Ser Ser Ser Ala	580	585	590	
tgg aaa gga ctc tct ggc aca ctt tac aca gtt tca agc tgg gcc gaa				1824
Trp Lys Gly Leu Ser Gly Thr Leu Tyr Thr Val Ser Ser Trp Ala Glu	595	600	605	
att cat tca ttc gct ctt gga tgg gag taa				1854
Ile His Ser Phe Ala Leu Gly Trp Glu	610	615		

<210> 2

<211> 617

5 <212> PRT

<213> *Arabidopsis thaliana*

<400> 2

ES 2 806 732 T3

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

ES 2 806 732 T3

Ile	Asp	Asn	Tyr	Ser	Ser	Asp	Ala	Phe	Gln	Ala	Ser	Ala	Lys	Gln	Thr	245	250	255
Glu	Asp	Leu	Leu	Glu	Lys	Leu	Ser	Val	Ser	Met	Thr	Gly	Glu	Glu	Leu	260	265	270
Asp	Ile	Ile	Glu	Lys	Leu	Tyr	Gln	Gln	Ala	Met	Lys	Leu	Glu	Val	Glu	275	280	285
Phe	Phe	His	Ala	Gln	Pro	Leu	Ala	Gln	Pro	Thr	Ile	Val	Pro	Leu	Leu	290	295	300
Lys	Asn	His	Ser	Lys	Asp	Asp	Leu	Val	Ile	Phe	Ser	Asp	Phe	Asp	Leu	305	310	315
Thr	Cys	Thr	Val	Val	Asp	Ser	Ser	Ala	Ile	Leu	Ala	Glu	Ile	Ala	Ile	325	330	335
Val	Thr	Ala	Pro	Lys	Asp	Glu	Gln	Ser	Arg	Ser	Gly	Gln	Gln	Ile	His	340	345	350
Arg	Met	Leu	Ser	Ser	Asp	Leu	Lys	Asn	Thr	Trp	Asn	Leu	Leu	Ser	Lys	355	360	365
Gln	Tyr	Thr	Glu	His	Tyr	Glu	Glu	Cys	Ile	Glu	Ser	Ile	Leu	Asn	Lys	370	375	380
Lys	Lys	Ala	Asp	Lys	Phe	Asp	Tyr	Glu	Gly	Leu	Cys	Lys	Ala	Leu	Glu	385	390	395
Gln	Leu	Ser	Asp	Phe	Glu	Lys	Glu	Ala	Asn	Asn	Arg	Val	Ile	Glu	Ser	405	410	415
Gly	Val	Leu	Lys	Gly	Leu	Asn	Leu	Glu	Asp	Ile	Lys	Arg	Ala	Gly	Glu	420	425	430
Arg	Leu	Ile	Leu	Gln	Asp	Gly	Cys	Ile	Asn	Val	Phe	Gln	Lys	Ile	Leu	435	440	445
Lys	Thr	Glu	Asn	Leu	Asn	Ala	Glu	Leu	His	Val	Leu	Ser	Tyr	Cys	Trp	450	455	460
Cys	Gly	Asp	Leu	Ile	Arg	Ala	Ala	Phe	Ser	Ala	Gly	Gly	Val	Asp	Ala	465	470	475
Val	Glu	Val	His	Ala	Asn	Glu	Phe	Thr	Phe	Glu	Glu	Ser	Ile	Ser	Thr	485	490	495

ES 2 806 732 T3

Gly Glu Ile Glu Arg Lys Val Glu Ser Pro Ile Asn Lys Ala Gln Gln
500 505 510

Phe Lys Ser Ile Leu Gln Asn Arg Lys Asn Glu Asn Asn Lys Lys Ser
515 520 525

Phe Leu Ser Val Tyr Ile Gly Asp Ser Val Gly Asp Leu Leu Cys Leu
530 535 540

Leu Glu Ala Asp Ile Gly Ile Val Val Ser Ser Ser Ser Ser Leu Arg
545 550 555 560

Arg Val Gly Ser His Phe Gly Val Ser Phe Val Pro Leu Phe Ser Gly
565 570 575

Ile Val Gln Lys Gln Lys Gln His Thr Glu Glu Ser Ser Ser Ser Ala
580 585 590

Trp Lys Gly Leu Ser Gly Thr Leu Tyr Thr Val Ser Ser Trp Ala Glu
595 600 605

Ile His Ser Phe Ala Leu Gly Trp Glu
610 615

<210> 3

<211> 1806

5 <212> ADN

<213> *Pyrus x bretschneideri*

<220>

<221> CDS

10 <222> (1)..(1806)

<223> gen de *Pyrus x bretschneideri* que codifica para TMP fosfatasa [XP_009379735.1]

<400> 3

atg cgc ata ctc ttc ccc cca aac cca atc aaa acc cca act ctc ttc 48
Met Arg Ile Leu Phe Pro Pro Asn Pro Ile Lys Thr Pro Thr Leu Phe
1 5 10 15

aac tcc ctc cgt ctg cga ttc aac tcg ctc cga tcc cac tgt gcc aac 96
Asn Ser Leu Arg Leu Arg Phe Asn Ser Leu Arg Ser His Cys Ala Asn
20 25 30

tca atg gcc gta cct ccg ccg aag tca gcc atg gct tcc gcc gtc gtc 144
Ser Met Ala Val Pro Pro Pro Lys Ser Ala Met Ala Ser Ala Val Val
35 40 45

ggc aac gag gtg ggt ctc gcc cgc cgc ttc tgg atc aag ttc aag cga 192
Gly Asn Glu Val Gly Leu Ala Arg Arg Phe Trp Ile Lys Phe Lys Arg
50 55 60

gaa tcg att ttc gct atg tac act ccc ttc acg ctc tgt ttg gct gct 240

ES 2 806 732 T3

Glu 65	Ser	Ile	Phe	Ala	Met 70	Tyr	Thr	Pro	Phe	Thr 75	Leu	Cys	Leu	Ala	Ala 80		
ggg Gly	aat Asn	ctc Leu	aag Lys	att Ile	gaa Glu	act Thr	ttc Phe	cgc Arg	gat Asp	tat Tyr	att Ile	gcc Ala	caa Gln	gat Asp	gtt Val	288	
cac His	ttt Phe	ctc Leu	aag Lys	gcc Ala	ttc Phe	gct Ala	cac His	gcg Ala	tat Tyr	gaa Glu	ttg Leu	gca Ala	gaa Glu	gat Asp	tgt Cys	336	
gca Ala	gac Asp	gat Asp	gat Asp	gat Asp	gca Ala	aag Lys	ccc Pro	gtg Val	att Ile	tct Ser	gag Glu	ttg Leu	agg Arg	agg Arg	gca Ala	384	
gtt Val	ctg Leu	cag Gln	gag Glu	ctg Leu	aaa Lys	atg Met	cat His	gat Asp	tca Ser	ttt Phe	gtg Val	aag Lys	gaa Glu	tgg Trp	ggg Gly	432	
tta Leu	cag Gln	ggt Gly	gct Ala	aaa Lys	gag Glu	acc Thr	cct Pro	atc Ile	aac Asn	tcc Ser	gct Ala	gcg Ala	gtg Val	aag Lys	tac Tyr	480	
aca Thr	gat Asp	ttc Phe	tta Leu	ttg Ala	gca Ala	aca Thr	gcc Ala	tct Ser	gga Gly	aaa Lys	gtt Val	gaa Glu	gga Gly	gtc Val	aag Lys	528	
gga Gly	cct Pro	ggt Gly	aaa Lys	ctt Leu	gca Ala	act Thr	cca Pro	ttt Phe	gaa Glu	aga Arg	acc Thr	aaa Lys	gtg Val	gct Ala	gct Ala	576	
tac Tyr	acc Thr	ctt Leu	ggc Gly	gct Ala	atg Met	act Thr	cct Pro	tgc Cys	atg Met	aga Arg	ctg Leu	tat Tyr	gcc Ala	ttt Phe	ctt Leu	624	
ggt Gly	aag Lys	gag Glu	ttc Phe	aag Lys	gca Ala	ctt Leu	cta Leu	gat Asp	ccc Pro	agc Ser	gaa Glu	ggc Gly	agt Ser	cac His	ccg Pro	672	
tac Tyr	ttg Leu	aag Lys	tgg Trp	att Ile	gac Asp	agt Ser	tat Tyr	tct Ser	tct Ser	aaa Lys	agt Ser	ttt Phe	cag Gln	gca Ala	tca Ser	720	
gct Ala	gtg Val	caa Gln	atc Ile	gaa Glu	gag Glu	ttg Leu	ctg Leu	gat Asp	aaa Lys	cta Leu	agt Ser	gtc Val	tct Ser	ttg Leu	aca Thr	768	
ggc Gly	gag Glu	gag Glu	ctt Leu	gac Asp	atc Ile	atc Ile	gaa Glu	aag Lys	ctt Leu	tac Tyr	cac His	caa Gln	gca Ala	atg Met	aaa Lys	816	
ctt Leu	gag Glu	atc Ile	gag Glu	ttc Phe	ttc Phe	tct Ser	gct Ala	cag Gln	tct Ser	ctt Leu	gtt Val	cag Gln	cca Pro	act Thr	gta Val	864	
gtt Val	cct Pro	ctg Leu	atc Ile	aga Arg	gaa Glu	cat His	aac Asn	cct Pro	gca Ala	gaa Glu	gat Asp	cgg Arg	ctc Leu	atg Met	ata Ile	912	
ttt Phe	tct Ser	gat Asp	ttt Phe	gat Asp	ttg Leu	act Thr	tgt Cys	aca Thr	gtc Val	gtt Val	gat Asp	tca Ser	tct Ser	gcc Ala	att Ile	960	

ttg gct gaa att gca ata gta aca gca cca aaa tct gat caa cat caa Leu Ala Glu Ile Ala Ile Val Thr Ala Pro Lys Ser Asp Gln His Gln 325 330 335	1008
ccc gaa aat cag att gct cgg atg tct tgc gct gat ctc agg aat aca Pro Glu Asn Gln Ile Ala Arg Met Ser Ser Ala Asp Leu Arg Asn Thr 340 345 350	1056
tgg ggt ctt ctt tcc agg cag tac aca gaa gag tat gag caa tgc ata Trp Gly Leu Leu Ser Arg Gln Tyr Thr Glu Glu Tyr Glu Gln Cys Ile 355 360 365	1104
gaa agc att gtt ccc act gaa aaa gca gtg ttt gac tat gaa aat ttg Glu Ser Ile Val Pro Thr Glu Lys Ala Val Phe Asp Tyr Glu Asn Leu 370 375 380	1152
ctt aaa gca cta gag aaa ctt tca gat ttt gag agg aag gca aac aat Leu Lys Ala Leu Glu Lys Leu Ser Asp Phe Glu Arg Lys Ala Asn Asn 385 390 395 400	1200
aga gtc acg aag tct gaa gta ctc aag ggt ctt aat ctc gaa gat ata Arg Val Thr Lys Ser Glu Val Leu Lys Gly Leu Asn Leu Glu Asp Ile 405 410 415	1248
aaa aga gct ggt gaa cgt ctc att ctt caa gat ggc tgt att aat ttc Lys Arg Ala Gly Glu Arg Leu Ile Leu Gln Asp Gly Cys Ile Asn Phe 420 425 430	1296
ttt cag aaa att gcc aag agt gaa aac ttg aat gca aat gtt cat gtt Phe Gln Lys Ile Ala Lys Ser Glu Asn Leu Asn Ala Asn Val His Val 435 440 445	1344
ctt tca tac tgt tgg tgt ggt gat ctc ata aga tgc gcc ttt tca tca Leu Ser Tyr Cys Trp Cys Gly Asp Leu Ile Arg Ser Ala Phe Ser Ser 450 455 460	1392
ggg ggt tta aac gag ctg gat gta cat gca aat gag ttt acc ttc gag Gly Gly Leu Asn Glu Leu Asp Val His Ala Asn Glu Phe Thr Phe Glu 465 470 475 480	1440
gaa tcc atc tcc aca ggt gat att gtt aag aag gtg gag tcc cct att Glu Ser Ile Ser Thr Gly Asp Ile Val Lys Lys Val Glu Ser Pro Ile 485 490 495	1488
gac aag gtt aaa tct ttt aaa gat att ttg aaa aat tgc agc aat gac Asp Lys Val Lys Ser Phe Lys Asp Ile Leu Lys Asn Cys Ser Asn Asp 500 505 510	1536
aga aag aac ttg act gtt tac att gga gac tgc gtg ggt gac tta ctt Arg Lys Asn Leu Thr Val Tyr Ile Gly Asp Ser Val Gly Asp Leu Leu 515 520 525	1584
tgt ctg ctg gag gcg gat att gga atc gta att ggg tca agt tca agc Cys Leu Leu Glu Ala Asp Ile Gly Ile Val Ile Gly Ser Ser Ser Ser 530 535 540	1632
ctt agg aga gtg gcg act cag ttt ggg gta tct ttt gtt ccg ttg ttc Leu Arg Arg Val Ala Thr Gln Phe Gly Val Ser Phe Val Pro Leu Phe 545 550 555 560	1680
ccg ggt tta gtt aag aaa cag aaa gaa tgc aca gat gga agg tct cct Pro Gly Leu Val Lys Lys Gln Lys Glu Cys Thr Asp Gly Arg Ser Pro 565 570 575	1728
agt tgg aaa ggg tta act ggt att ctt tac aca gtg aat agt tgg gcg Ser Trp Lys Gly Leu Thr Gly Ile Leu Tyr Thr Val Asn Ser Trp Ala 580 585 590	1776
gaa ata cat gcc ttc att ttg ggg tgt taa Glu Ile His Ala Phe Ile Leu Gly Cys 595 600	1806

<210> 4

5 <211> 601

<212> PRT

<213> *Pyrus x bretschneideri*

ES 2 806 732 T3

<400> 4

Met Arg Ile Leu Phe Pro Pro Asn Pro Ile Lys Thr Pro Thr Leu Phe
1 5 10 15

Asn Ser Leu Arg Leu Arg Phe Asn Ser Leu Arg Ser His Cys Ala Asn
20 25 30

Ser Met Ala Val Pro Pro Pro Lys Ser Ala Met Ala Ser Ala Val Val
35 40 45

Gly Asn Glu Val Gly Leu Ala Arg Arg Phe Trp Ile Lys Phe Lys Arg
50 55 60

Glu Ser Ile Phe Ala Met Tyr Thr Pro Phe Thr Leu Cys Leu Ala Ala
65 70 75 80

Gly Asn Leu Lys Ile Glu Thr Phe Arg Asp Tyr Ile Ala Gln Asp Val
85 90 95

His Phe Leu Lys Ala Phe Ala His Ala Tyr Glu Leu Ala Glu Asp Cys
100 105 110

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

Val Leu Gln Glu Leu Lys Met His Asp Ser Phe Val Lys Glu Trp Gly
130 135 140

Leu Gln Gly Ala Lys Glu Thr Pro Ile Asn Ser Ala Ala Val Lys Tyr
145 150 155 160

Thr Asp Phe Leu Leu Ala Thr Ala Ser Gly Lys Val Glu Gly Val Lys
165 170 175

Gly Pro Gly Lys Leu Ala Thr Pro Phe Glu Arg Thr Lys Val Ala Ala
180 185 190

ES 2 806 732 T3

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

ES 2 806 732 T3

435 440 445
 Leu Ser Tyr Cys Trp Cys Gly Asp Leu Ile Arg Ser Ala Phe Ser Ser
 450 455 460
 Gly Gly Leu Asn Glu Leu Asp Val His Ala Asn Glu Phe Thr Phe Glu
 465 470 475 480
 Glu Ser Ile Ser Thr Gly Asp Ile Val Lys Lys Val Glu Ser Pro Ile
 485 490 495
 Asp Lys Val Lys Ser Phe Lys Asp Ile Leu Lys Asn Cys Ser Asn Asp
 500 505 510
 Arg Lys Asn Leu Thr Val Tyr Ile Gly Asp Ser Val Gly Asp Leu Leu
 515 520 525
 Cys Leu Leu Glu Ala Asp Ile Gly Ile Val Ile Gly Ser Ser Ser Ser
 530 535 540
 Leu Arg Arg Val Ala Thr Gln Phe Gly Val Ser Phe Val Pro Leu Phe
 545 550 555 560
 Pro Gly Leu Val Lys Lys Gln Lys Glu Cys Thr Asp Gly Arg Ser Pro
 565 570 575
 Ser Trp Lys Gly Leu Thr Gly Ile Leu Tyr Thr Val Asn Ser Trp Ala
 580 585 590
 Glu Ile His Ala Phe Ile Leu Gly Cys
 595 600

<210> 5

<211> 1797

5 <212> ADN

<213> *Brassica napus*

<220>

<221> CDS

10 <222> (1)..(1797)

<223> gen de *Brassica napus* que codifica para TMP fosfatasa [CDY62623.1]

<400> 5

atg cgc atc ctc aac aac tcg ctc gcc ctt ctc cga tcg ccc cgc gcc 48
 Met Arg Ile Leu Asn Asn Ser Leu Ala Leu Leu Arg Ser Pro Arg Ala
 1 5 10 15

gcc gcc ccg atc cgt tct cta ctg ttc ggc agc aag aag tct tcc gtc 96
 Ala Ala Pro Ile Arg Ser Leu Leu Phe Gly Ser Lys Lys Ser Ser Val
 20 25 30

ES 2 806 732 T3

tcc cga tcg gcg gcc gcc ttc tct tcg gcg atg tcg att cct cct cct	144
Ser Arg Ser Ala Ala Ala Phe Ser Ser Ala Met Ser Ile Pro Pro Pro	
35 40 45	
agc ata tcc acc tcg gaa gaa gct ctg gcg ggg agg ctg tgg atc aag	192
Ser Ile Ser Thr Ser Glu Glu Ala Leu Ala Gly Arg Leu Trp Ile Lys	
50 55 60	
ttc aac aga gag tgc ctc ttc tct atg tac agc ccc ttc gcc gtt tct	240
Phe Asn Arg Glu Cys Leu Phe Ser Met Tyr Ser Pro Phe Ala Val Ser	
65 70 75 80	
ttg gcc gcc ggc aat ctc aag atc gag acc ttc cgg cag tat att gct	288
Leu Ala Ala Gly Asn Leu Lys Ile Glu Thr Phe Arg Gln Tyr Ile Ala	
85 90 95	
cag gat gtt cat ttc ctc aag gcc ttt gct cac gcg tat gag ttg gcc	336
Gln Asp Val His Phe Leu Lys Ala Phe Ala His Ala Tyr Glu Leu Ala	
100 105 110	
gca gag tgt gct gat gat gat gat gat aag ttg gca att tct gac ttg	384
Ala Glu Cys Ala Asp Asp Asp Asp Lys Leu Ala Ile Ser Asp Leu	
115 120 125	
agg aaa agc gtc atg gat gag ttg aaa atg cac aac tca ttt gta cag	432
Arg Lys Ser Val Met Asp Glu Leu Lys Met His Asn Ser Phe Val Gln	
130 135 140	
gat tgg gat tta gac atc agc aaa gaa gta agt gtt aac tca gca aca	480
Asp Trp Asp Leu Asp Ile Ser Lys Glu Val Ser Val Asn Ser Ala Thr	
145 150 155 160	
ttg aga tac acc gag ttc tta tta gct aca tca tcc gga aaa gta gaa	528
Leu Arg Tyr Thr Glu Phe Leu Leu Ala Thr Ser Ser Gly Lys Val Glu	
165 170 175	
gga ctc aaa gct ccc ggc atg ctt gat act cca ttt gag aaa acc aaa	576
Gly Leu Lys Ala Pro Gly Met Leu Asp Thr Pro Phe Glu Lys Thr Lys	
180 185 190	
gtg gcc gcc tac acg ctt ggt gct gtg aca cct tgc atg aag ctg tat	624
Val Ala Ala Tyr Thr Leu Gly Ala Val Thr Pro Cys Met Lys Leu Tyr	
195 200 205	
gcc ttt ctt ggt aag gag ttt gga gcg ctt cta gat tcg agt gaa gcg	672
Ala Phe Leu Gly Lys Glu Phe Gly Ala Leu Leu Asp Ser Ser Glu Ala	
210 215 220	
aac cat ccc tac aag aaa tgg atc gaa aat tat tct agt gat gca ttc	720
Asn His Pro Tyr Lys Lys Trp Ile Glu Asn Tyr Ser Ser Asp Ala Phe	
225 230 235 240	
cag gca tca gct aag caa act gaa gac ttg ctt gag aag ctt agt gtg	768
Gln Ala Ser Ala Lys Gln Thr Glu Asp Leu Leu Glu Lys Leu Ser Val	
245 250 255	
tgt atg act ggc gaa gag ctg gac atc att gaa aaa ctg tat caa cag	816
Cys Met Thr Gly Glu Glu Leu Asp Ile Ile Glu Lys Leu Tyr Gln Gln	
260 265 270	
gca atg aaa ctt gaa gta gag ttc ttc cac gca caa ccg ttt gct cag	864
Ala Met Lys Leu Glu Val Glu Phe Phe His Ala Gln Pro Phe Ala Gln	
275 280 285	

ES 2 806 732 T3

cct acc ata gtt ccg ctg ctg aag aac cat tca aaa gat gag ctg atg Pro Thr Ile Val Pro Leu Leu Lys Asn His Ser Lys Asp Glu Leu Met 290 295 300	912
ata ttt tct gat ttt gat ctg act tgc acc gtt gtt gat tct tct gct Ile Phe Ser Asp Phe Asp Leu Thr Cys Thr Val Val Asp Ser Ser Ala 305 310 315 320	960
att tta gcc gaa att gca atc gta act gcc ccg aaa gat gat cag ggt Ile Leu Ala Glu Ile Ala Ile Val Thr Ala Pro Lys Asp Asp Gln Gly 325 330 335	1008
caa caa att aat ccg atg ctt tcg gct gac ctt aag aac acc tgg agt Gln Gln Ile Asn Arg Met Leu Ser Ala Asp Leu Lys Asn Thr Trp Ser 340 345 350	1056
cta ctt tcc aaa cag tat aca gag cac tat gaa gag tgc ata gag agt Leu Leu Ser Lys Gln Tyr Thr Glu His Tyr Glu Glu Cys Ile Glu Ser 355 360 365	1104
att ctg aat aag gaa aaa gcg gac aag ttt gac tac gag ggt ttg tgt Ile Leu Asn Lys Glu Lys Ala Asp Lys Phe Asp Tyr Glu Gly Leu Cys 370 375 380	1152
gaa gca cta gag cag ctg tca gag ttt gag aag aaa gca aac gac cga Glu Ala Leu Glu Gln Leu Ser Glu Phe Glu Lys Lys Ala Asn Asp Arg 385 390 395 400	1200
gtg ata gag tct ggt gta ctc aag ggc ctg aat ctc gat gac atc aag Val Ile Glu Ser Gly Val Leu Lys Gly Leu Asn Leu Asp Asp Ile Lys 405 410 415	1248
cga gct ggg gaa agg ttg att ctt caa gat ggc tgc atc aat gtc ttc Arg Ala Gly Glu Arg Leu Ile Leu Gln Asp Gly Cys Ile Asn Val Phe 420 425 430	1296
cag aaa att ttg aag act cag gat gtg aat gca aaa ctc cac gtg ctt Gln Lys Ile Leu Lys Thr Gln Asp Val Asn Ala Lys Leu His Val Leu 435 440 445	1344
tcg tat tgt tgg tgt ggt gac ctc atc aga gca gcc ttt tct gca ccg Ser Tyr Cys Trp Cys Gly Asp Leu Ile Arg Ala Ala Phe Ser Ala Arg 450 455 460	1392
gga gta gat gca gtg gaa gta cat gca aat gaa ttc aca ttc gag gaa Gly Val Asp Ala Val Glu Val His Ala Asn Glu Phe Thr Phe Glu Glu 465 470 475 480	1440
tcc atc tct act gga gaa ata gaa aga aaa gtg gaa tcc cca atc gac Ser Ile Ser Thr Gly Glu Ile Glu Arg Lys Val Glu Ser Pro Ile Asp 485 490 495	1488
aag gct caa cag ttc aag agc atc cta caa aac aga aag aag gat gag Lys Ala Gln Gln Phe Lys Ser Ile Leu Gln Asn Arg Lys Lys Asp Glu 500 505 510	1536
gag aaa agc atc ctc act gtt tac att gga gat tca gta ggt gac ttg Glu Lys Ser Ile Leu Thr Val Tyr Ile Gly Asp Ser Val Gly Asp Leu 515 520 525	1584
ctc tgt ctc ctg gag gca gac att gga ata gtg gtc gcc tct agc tcg Leu Cys Leu Leu Glu Ala Asp Ile Gly Ile Val Val Ala Ser Ser Ser 1632	

ES 2 806 732 T3

530 535 540

agc ctc agg aga gtg gga agc cat ttc ggg gtc tca ttt gtg cct ttg 1680
 Ser Leu Arg Arg Val Gly Ser His Phe Gly Val Ser Phe Val Pro Leu
 545 550 555 560

ttc tct gga att gtg caa aaa cag aaa caa gaa gaa acc tgg aag ggg 1728
 Phe Ser Gly Ile Val Gln Lys Gln Lys Gln Glu Glu Thr Trp Lys Gly
 565 570 575

ctc tct ggc aca ctt tac acg gta tca agc tgg gct gaa ata cat tcc 1776
 Leu Ser Gly Thr Leu Tyr Thr Val Ser Ser Trp Ala Glu Ile His Ser
 580 585 590

ttc gct ctt gga tgg gag taa 1797
 Phe Ala Leu Gly Trp Glu
 595

<210> 6
 <211> 598
 5 <212> PRT
 <213> *Brassica napus*

<400> 6
 Met Arg Ile Leu Asn Asn Ser Leu Ala Leu Leu Arg Ser Pro Arg Ala
 1 5 10 15

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

Ser Arg Ser Ala Ala Ala Phe Ser Ser Ala Met Ser Ile Pro Pro Pro
 35 40 45

Ser Ile Ser Thr Ser Glu Glu Ala Leu Ala Gly Arg Leu Trp Ile Lys
 50 55 60

Phe Asn Arg Glu Cys Leu Phe Ser Met Tyr Ser Pro Phe Ala Val Ser
 65 70 75 80

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

Gln Asp Val His Phe Leu Lys Ala Phe Ala His Ala Tyr Glu Leu Ala
 100 105 110

Ala Glu Cys Ala Asp Asp Asp Asp Lys Leu Ala Ile Ser Asp Leu
 115 120 125

Arg Lys Ser Val Met Asp Glu Leu Lys Met His Asn Ser Phe Val Gln
 130 135 140

Asp Trp Asp Leu Asp Ile Ser Lys Glu Val Ser Val Asn Ser Ala Thr

ES 2 806 732 T3

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

ES 2 806 732 T3

Val Ile Glu Ser Gly Val Leu Lys Gly Leu Asn Leu Asp Asp Ile Lys
405 410 415

Arg Ala Gly Glu Arg Leu Ile Leu Gln Asp Gly Cys Ile Asn Val Phe
420 425 430

Gln Lys Ile Leu Lys Thr Gln Asp Val Asn Ala Lys Leu His Val Leu
435 440 445

Ser Tyr Cys Trp Cys Gly Asp Leu Ile Arg Ala Ala Phe Ser Ala Arg
450 455 460

Gly Val Asp Ala Val Glu Val His Ala Asn Glu Phe Thr Phe Glu Glu
465 470 475 480

Ser Ile Ser Thr Gly Glu Ile Glu Arg Lys Val Glu Ser Pro Ile Asp
485 490 495

Lys Ala Gln Gln Phe Lys Ser Ile Leu Gln Asn Arg Lys Lys Asp Glu
500 505 510

Glu Lys Ser Ile Leu Thr Val Tyr Ile Gly Asp Ser Val Gly Asp Leu
515 520 525

Leu Cys Leu Leu Glu Ala Asp Ile Gly Ile Val Val Ala Ser Ser Ser
530 535 540

Ser Leu Arg Arg Val Gly Ser His Phe Gly Val Ser Phe Val Pro Leu
545 550 555 560

Phe Ser Gly Ile Val Gln Lys Gln Lys Gln Glu Glu Thr Trp Lys Gly
565 570 575

Leu Ser Gly Thr Leu Tyr Thr Val Ser Ser Trp Ala Glu Ile His Ser
580 585 590

Phe Ala Leu Gly Trp Glu
595

<210> 7

<211> 1815

5 <212> ADN

<213> *Glycine max*

<220>

<221> CDS

10 <222> (1)..(1815)

<223> gen de *Glycine max* que codifica para TMP fosfatasa [XP_003536133.1]

<400> 7

ES 2 806 732 T3

atg cgc atg cgg tgg ttc ctc cga agc cca atc atc aaa acc tcg ctg Met Arg Met Arg Trp Phe Leu Arg Ser Pro Ile Ile Lys Thr Ser Leu 1 5 10 15	48
ctg aat ctg agc cct cca att tcg ttt aga cct cac tgg gcg agg agg Leu Asn Leu Ser Pro Pro Ile Ser Phe Arg Pro His Trp Ala Arg Arg 20 25 30	96
acc ttc act tct tcg aga ttg tca atg gcg gcc atc cac aac cac agc Thr Phe Thr Ser Ser Arg Leu Ser Met Ala Ala Ile His Asn His Ser 35 40 45	144
aac agc aac agc gaa acc gga ctc gcg aga cgg ttt tgg atc aag ttc Asn Ser Asn Ser Glu Thr Gly Leu Ala Arg Arg Phe Trp Ile Lys Phe 50 55 60	192
act cgt gaa tcc atc ttc gcc atg tac act ccc ttc gcc atc gcc ttg Thr Arg Glu Ser Ile Phe Ala Met Tyr Thr Pro Phe Ala Ile Ala Leu 65 70 75 80	240
gcc tcc ggt aat ttg cac att gat tcc ttc cac cat tac atc gcc caa Ala Ser Gly Asn Leu His Ile Asp Ser Phe His His Tyr Ile Ala Gln 85 90 95	288
gac gtt cat ttc cta cgc gcc ttt gct caa gcg tat gag ttg gct gaa Asp Val His Phe Leu Arg Ala Phe Ala Gln Ala Tyr Glu Leu Ala Glu 100 105 110	336
gag tgt gct gat gac gac gat gcg aaa ctt gga atc tgt gag ttg agg Glu Cys Ala Asp Asp Asp Ala Lys Leu Gly Ile Cys Glu Leu Arg 115 120 125	384
aag gca gtt cta gag gag ctg aag atg cac aac ttg ctg gta cag gaa Lys Ala Val Leu Glu Glu Leu Lys Met His Asn Leu Leu Val Gln Glu 130 135 140	432
cgg gag ttg gac ctt gcc aaa gag cat ggt att aat tct gca act gtt Arg Glu Leu Asp Leu Ala Lys Glu His Gly Ile Asn Ser Ala Thr Val 145 150 155 160	480
aag tac aca gag ttc ctg ctg gct aca gcc tct ggg aag att gaa gga Lys Tyr Thr Glu Phe Leu Leu Ala Thr Ala Ser Gly Lys Ile Glu Gly 165 170 175	528
cta aaa ggt cct ggt aaa ctt gct aca cca ttt gag aaa aca aaa att Leu Lys Gly Pro Gly Lys Leu Ala Thr Pro Phe Glu Lys Thr Lys Ile 180 185 190	576
gct gct tat act tta ggt gcc atg act cct tgc atg agg ctt tat gcc Ala Ala Tyr Thr Leu Gly Ala Met Thr Pro Cys Met Arg Leu Tyr Ala 195 200 205	624
gtt atg gga aag aag ttc cag gaa ctt ttg gat tcc aat gaa agt act Val Met Gly Lys Lys Phe Gln Glu Leu Leu Asp Ser Asn Glu Ser Thr 210 215 220	672
cac cca tat aac aag tgg atc aac aac tat tcc tct gat ggt ttc cag His Pro Tyr Asn Lys Trp Ile Asn Asn Tyr Ser Ser Asp Gly Phe Gln 225 230 235 240	720
gct act act ctg caa act gaa gat ttg ctc gac aaa cta agt gtc tct Ala Thr Thr Leu Gln Thr Glu Asp Leu Leu Asp Lys Leu Ser Val Ser	768

ES 2 806 732 T3

245	250	255	
ttg act ggt gaa gaa ctt gat gtc att gaa aag ctt tat tac caa gca Leu Thr Gly Glu Glu Leu Asp Val Ile Glu Lys Leu Tyr Tyr Gln Ala 260 265 270			816
atg aag ctt gaa ata gag ttc ttc tct gct cag cca ctc ttc cag cca Met Lys Leu Glu Ile Glu Phe Phe Ser Ala Gln Pro Leu Phe Gln Pro 275 280 285			864
act ata gta ccc ttg act aaa gga cat aag cct gtg gaa gat cat ctc Thr Ile Val Pro Leu Thr Lys Gly His Lys Pro Val Glu Asp His Leu 290 295 300			912
att att ttt tct gat ttt gat tta aca tgc acc gta gtt gat tcg tcc Ile Ile Phe Ser Asp Phe Asp Leu Thr Cys Thr Val Val Asp Ser Ser 305 310 315 320			960
gcc atc ttg gct gaa att gcc ata gtg acg gca cca aaa tct gat cag Ala Ile Leu Ala Glu Ile Ala Ile Val Thr Ala Pro Lys Ser Asp Gln 325 330 335			1008
aat cag cct gaa gat caa att gtt cgg atg tta tct tct gac ctc agg Asn Gln Pro Glu Asp Gln Ile Val Arg Met Leu Ser Ser Asp Leu Arg 340 345 350			1056
aat aca tgg ggt ttt cta tct aaa cag tat acg gag gag tat gag caa Asn Thr Trp Gly Phe Leu Ser Lys Gln Tyr Thr Glu Glu Tyr Glu Gln 355 360 365			1104
tgt ata gaa agc att atg cct ccc gat aga ttg aac aat ttc gat tac Cys Ile Glu Ser Ile Met Pro Pro Asp Arg Leu Asn Asn Phe Asp Tyr 370 375 380			1152
aaa gaa ttg tcg atg gcc ctt gag caa ctt tca aaa ttt gag aac act Lys Glu Leu Ser Met Ala Leu Glu Gln Leu Ser Lys Phe Glu Asn Thr 385 390 395 400			1200
gca aat aat agg gtt atc gag tca ggg gta ctc aag ggt ata agt cta Ala Asn Asn Arg Val Ile Glu Ser Gly Val Leu Lys Gly Ile Ser Leu 405 410 415			1248
gaa gat ata aag cgt gct gga gag cgt ctg ata cta caa gat ggt tgc Glu Asp Ile Lys Arg Ala Gly Glu Arg Leu Ile Leu Gln Asp Gly Cys 420 425 430			1296
cct aac ttc ttt cag agc att gtt aag aat gaa aat ttg aat gcc aac Pro Asn Phe Phe Gln Ser Ile Val Lys Asn Glu Asn Leu Asn Ala Asn 435 440 445			1344
gtg cat gtt ctt tca tac tgc tgg tgt ggt gac ctc att agg tct act Val His Val Leu Ser Tyr Cys Trp Cys Gly Asp Leu Ile Arg Ser Thr 450 455 460			1392
ttc tct tcc gct gat tta aat gag ttg aat gtt cat gct aat gag ttc Phe Ser Ser Ala Asp Leu Asn Glu Leu Asn Val His Ala Asn Glu Phe 465 470 475 480			1440
act tat gag gga tct gtt tcc acg ggt gaa att gtt aag aaa gtg gag Thr Tyr Glu Gly Ser Val Ser Thr Gly Glu Ile Val Lys Lys Val Glu 485 490 495			1488
tct ccc att gac aag gtt gaa gct ttt cgt aac ata ttg aaa aat tgc			1536

ES 2 806 732 T3

Ser	Pro	Ile	Asp	Lys	Val	Glu	Ala	Phe	Arg	Asn	Ile	Leu	Lys	Asn	Cys	
			500					505					510			
aat	gat	gac	aaa	aag	aaa	tta	act	gtt	tac	att	ggc	gat	tca	gtg	ggt	1584
Asn	Asp	Asp	Lys	Lys	Lys	Leu	Thr	Val	Tyr	Ile	Gly	Asp	Ser	Val	Gly	
			515				520					525				
gat	tta	ctt	tgc	cta	ctt	gaa	gct	gat	gta	gga	att	gtg	att	ggt	tca	1632
Asp	Leu	Leu	Cys	Leu	Leu	Glu	Ala	Asp	Val	Gly	Ile	Val	Ile	Gly	Ser	
			530			535					540					
agt	tca	agc	ctt	aga	agt	gta	ggg	acg	cag	ttt	ggt	att	tca	ttt	gtc	1680
Ser	Ser	Ser	Leu	Arg	Ser	Val	Gly	Thr	Gln	Phe	Gly	Ile	Ser	Phe	Val	
			545			550			555						560	
cca	ttg	tat	tct	ggc	ttg	gtt	aag	aaa	cag	aaa	gaa	tat	gtt	gaa	gga	1728
Pro	Leu	Tyr	Ser	Gly	Leu	Val	Lys	Lys	Gln	Lys	Glu	Tyr	Val	Glu	Gly	
				565					570					575		
agc	act	tct	gat	tgg	aag	ggt	tta	tct	ggc	att	ctt	tac	aca	gtc	tct	1776
Ser	Thr	Ser	Asp	Trp	Lys	Gly	Leu	Ser	Gly	Ile	Leu	Tyr	Thr	Val	Ser	
			580				585						590			
agt	tgg	gct	gaa	gtg	cat	gct	ttt	att	ttg	ggt	tgc	tag				1815
Ser	Trp	Ala	Glu	Val	His	Ala	Phe	Ile	Leu	Gly	Cys					
			595				600									

<210> 8
 <211> 604
 <212> PRT
 <213> Glycine max

<400> 8

Met	Arg	Met	Arg	Trp	Phe	Leu	Arg	Ser	Pro	Ile	Ile	Lys	Thr	Ser	Leu	
1				5					10					15		
Leu	Asn	Leu	Ser	Pro	Pro	Ile	Ser	Phe	Arg	Pro	His	Trp	Ala	Arg	Arg	
			20					25					30			
Thr	Phe	Thr	Ser	Ser	Arg	Leu	Ser	Met	Ala	Ala	Ile	His	Asn	His	Ser	
			35				40					45				
Asn	Ser	Asn	Ser	Glu	Thr	Gly	Leu	Ala	Arg	Arg	Phe	Trp	Ile	Lys	Phe	
			50			55					60					
Thr	Arg	Glu	Ser	Ile	Phe	Ala	Met	Tyr	Thr	Pro	Phe	Ala	Ile	Ala	Leu	
			65		70				75						80	
Ala	Ser	Gly	Asn	Leu	His	Ile	Asp	Ser	Phe	His	His	Tyr	Ile	Ala	Gln	
			85					90						95		
Asp	Val	His	Phe	Leu	Arg	Ala	Phe	Ala	Gln	Ala	Tyr	Glu	Leu	Ala	Glu	
			100				105						110			

ES 2 806 732 T3

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

ES 2 806 732 T3

Cys Ile Glu Ser Ile Met Pro Pro Asp Arg Leu Asn Asn Phe Asp Tyr
370 375 380

Lys Glu Leu Ser Met Ala Leu Glu Gln Leu Ser Lys Phe Glu Asn Thr
385 390 395 400

Ala Asn Asn Arg Val Ile Glu Ser Gly Val Leu Lys Gly Ile Ser Leu
405 410 415

Glu Asp Ile Lys Arg Ala Gly Glu Arg Leu Ile Leu Gln Asp Gly Cys
420 425 430

Pro Asn Phe Phe Gln Ser Ile Val Lys Asn Glu Asn Leu Asn Ala Asn
435 440 445

Val His Val Leu Ser Tyr Cys Trp Cys Gly Asp Leu Ile Arg Ser Thr
450 455 460

Phe Ser Ser Ala Asp Leu Asn Glu Leu Asn Val His Ala Asn Glu Phe
465 470 475 480

Thr Tyr Glu Gly Ser Val Ser Thr Gly Glu Ile Val Lys Lys Val Glu
485 490 495

Ser Pro Ile Asp Lys Val Glu Ala Phe Arg Asn Ile Leu Lys Asn Cys
500 505 510

Asn Asp Asp Lys Lys Lys Leu Thr Val Tyr Ile Gly Asp Ser Val Gly
515 520 525

Asp Leu Leu Cys Leu Leu Glu Ala Asp Val Gly Ile Val Ile Gly Ser
530 535 540

Ser Ser Ser Leu Arg Ser Val Gly Thr Gln Phe Gly Ile Ser Phe Val
545 550 555 560

Pro Leu Tyr Ser Gly Leu Val Lys Lys Gln Lys Glu Tyr Val Glu Gly
565 570 575

Ser Thr Ser Asp Trp Lys Gly Leu Ser Gly Ile Leu Tyr Thr Val Ser
580 585 590

Ser Trp Ala Glu Val His Ala Phe Ile Leu Gly Cys
595 600

<210> 9

<211> 1845

5 <212> ADN

<213> *Nicotiana tomentosiformis*

<220>

<221> CDS

10 <222> (1)..(1845)

<223> gen de *Nicotiana tomentosiformis* que codifica para TMP fosfatasa [XP_009615535.1]

<400> 9

ES 2 806 732 T3

atg cgc ttc tca tta tta tcg ccc ctt gtt ctt aac cca gtc atc aga Met Arg Phe Ser Leu Leu Ser Pro Leu Val Leu Asn Pro Val Ile Arg 1 5 10 15	48
ttc tcc aat tcc aac gcg ctt ttt ggg tta cga ttc caa tta tac cct Phe Ser Asn Ser Asn Ala Leu Phe Gly Leu Arg Phe Gln Leu Tyr Pro 20 25 30	96
cgt tac tct cgg tat tta cga tcg ccc gtt aca atg gcg tcg gcg aaa Arg Tyr Ser Arg Tyr Leu Arg Ser Pro Val Thr Met Ala Ser Ala Lys 35 40 45	144
cca aag ccg gcg gcg gcg gtg aac aag ttt ccg gta gag gag gaa tgt Pro Lys Pro Ala Ala Ala Val Asn Lys Phe Pro Val Glu Glu Glu Cys 50 55 60	192
gtg ggt ata gcg agg aag tgt tgg atc aag ttc aag aga gag tct act Val Gly Ile Ala Arg Lys Cys Trp Ile Lys Phe Lys Arg Glu Ser Thr 65 70 75 80	240
ttc gct ctg tac act ccg ttt gtg gtt agt ttg gca tca gga acc cta Phe Ala Leu Tyr Thr Pro Phe Val Val Ser Leu Ala Ser Gly Thr Leu 85 90 95	288
aat ctg gac act ttc cgc cat tac att gct cag gat gtt cac ttc ctc Asn Leu Asp Thr Phe Arg His Tyr Ile Ala Gln Asp Val His Phe Leu 100 105 110	336
aaa tcc ttc gct caa gcg tat gaa gct gca gaa gag tgt act gac gat Lys Ser Phe Ala Gln Ala Tyr Glu Ala Ala Glu Glu Cys Thr Asp Asp 115 120 125	384
gac gat gcg aag gtt ggc att agt gag ttg cgg aag aat gtt att gaa Asp Asp Ala Lys Val Gly Ile Ser Glu Leu Arg Lys Asn Val Ile Glu 130 135 140	432
gaa ctt aaa atg cat gat gca gtt tta aaa gag tgg ggc att gat ctg Glu Leu Lys Met His Asp Ala Val Leu Lys Glu Trp Gly Ile Asp Leu 145 150 155 160	480
gtc aaa gag tcc agt ctt aac cct gca acg gcc aag tac aca gat ttt Val Lys Glu Ser Ser Leu Asn Pro Ala Thr Ala Lys Tyr Thr Asp Phe 165 170 175	528
tta tca gct aca gct tca gga aag gtg gaa gga gta aaa gct gct aaa Leu Ser Ala Thr Ala Ser Gly Lys Val Glu Gly Val Lys Ala Ala Lys 180 185 190	576
ctt gcc aca cca ttt gag aga acg aag ttg gca gct tat act cta ggt Leu Ala Thr Pro Phe Glu Arg Thr Lys Leu Ala Ala Tyr Thr Leu Gly 195 200 205	624

ES 2 806 732 T3

gct atg act cct tgc atg agg ctt tac gcc tac att ggt aag gag ctg Ala Met Thr Pro Cys Met Arg Leu Tyr Ala Tyr Ile Gly Lys Glu Leu 210 215 220	672
caa gtg ttc ctc gag gga gag aaa att cat cca tac aag aag tgg att Gln Val Phe Leu Glu Gly Glu Lys Ile His Pro Tyr Lys Lys Trp Ile 225 230 235 240	720
gac agt tat gcc tct gaa agt ttc cag gca tca gct ctt caa acc gag Asp Ser Tyr Ala Ser Glu Ser Phe Gln Ala Ser Ala Leu Gln Thr Glu 245 250 255	768
gac ttg ttg gat aaa ctg agt gtc cct ttg aca ggc gag gag ctt gac Asp Leu Leu Asp Lys Leu Ser Val Pro Leu Thr Gly Glu Glu Leu Asp 260 265 270	816
atc att gaa aag ctt tat cat caa gca atg aaa ctt gaa att gat ttc Ile Ile Glu Lys Leu Tyr His Gln Ala Met Lys Leu Glu Ile Asp Phe 275 280 285	864
ttc tta acc cag cca ctt gtt cag aaa gct gtc atc cct ttg tca aaa Phe Leu Thr Gln Pro Leu Val Gln Lys Ala Val Ile Pro Leu Ser Lys 290 295 300	912
gat cac aac cct gct gaa cac cgg ctt aca ata ttt tct gat ttc gat Asp His Asn Pro Ala Glu His Arg Leu Thr Ile Phe Ser Asp Phe Asp 305 310 315 320	960
ttg acg tgc act gtt gtt gat tct tct gcc atc ttg gct gaa att gca Leu Thr Cys Thr Val Val Asp Ser Ser Ala Ile Leu Ala Glu Ile Ala 325 330 335	1008
att ata aca gca ccg aga tct gat caa aat cga cca gag aat caa att Ile Ile Thr Ala Pro Arg Ser Asp Gln Asn Arg Pro Glu Asn Gln Ile 340 345 350	1056
gcg cgg atg ttg tgc gct gat ttg agg aat aca tgg gga gat ctc tct Ala Arg Met Leu Ser Ala Asp Leu Arg Asn Thr Trp Gly Asp Leu Ser 355 360 365	1104
aag cag tac act gaa gag tat gag caa tgt ata gag aag atg tta ctt Lys Gln Tyr Thr Glu Glu Tyr Glu Gln Cys Ile Glu Lys Met Leu Leu 370 375 380	1152
act gaa aaa gcg gaa aaa ttt gat tat gaa aga ctg cat aaa aca ctt Thr Glu Lys Ala Glu Lys Phe Asp Tyr Glu Arg Leu His Lys Thr Leu 385 390 395 400	1200
gag gaa ctt tct gat ttt gag aaa aga gca aat act agg gtg act gaa Glu Glu Leu Ser Asp Phe Glu Lys Arg Ala Asn Thr Arg Val Thr Glu 405 410 415	1248
tct ggg gta ctg aaa ggt tta aac ctt gaa gac ata aaa cga gct ggg Ser Gly Val Leu Lys Gly Leu Asn Leu Glu Asp Ile Lys Arg Ala Gly 420 425 430	1296
cag cga ttg att ctc cag gat ggt tgc acc aac ttc ttc cag agc ata Gln Arg Leu Ile Leu Gln Asp Gly Cys Thr Asn Phe Phe Gln Ser Ile 435 440 445	1344
ata aga aat gaa aat ctg aac gca gac att cat gtc ctc tcc tat tgc Ile Arg Asn Glu Asn Leu Asn Ala Asp Ile His Val Leu Ser Tyr Cys 450 455 460	1392

ES 2 806 732 T3

tgg tgt ggc gac ctt att agg tct tcc ttt tca tca ggg ggt ata gac	1440
Trp Cys Gly Asp Leu Ile Arg Ser Ser Phe Ser Ser Gly Gly Ile Asp	
465 470 475 480	
gct ctg aat gtg cat gcc aat gag ttt atg ttt caa gaa tct cta tcc	1488
Ala Leu Asn Val His Ala Asn Glu Phe Met Phe Gln Glu Ser Leu Ser	
485 490 495	
act ggt gaa att gtt aag aaa gtt gaa tcc ccc att gac aag gtt caa	1536
Thr Gly Glu Ile Val Lys Lys Val Glu Ser Pro Ile Asp Lys Val Gln	
500 505 510	
gca ttc agt aaa att cga atg aac tgt ggc aat gac caa aaa aat ctg	1584
Ala Phe Ser Lys Ile Arg Met Asn Cys Gly Asn Asp Gln Lys Asn Leu	
515 520 525	
act ctt tat att ggg gat tca gtc ggc gac tta ctt tgc ttg ctt gaa	1632
Thr Leu Tyr Ile Gly Asp Ser Val Gly Asp Leu Leu Cys Leu Leu Glu	
530 535 540	
gca gat gtt ggc ata gtg ctt ggt acg agc tca agt cta agg acg gtg	1680
Ala Asp Val Gly Ile Val Leu Gly Thr Ser Ser Ser Leu Arg Thr Val	
545 550 555 560	
ggg aat cat ttt ggt gtt tct ttt gtt cct ctg ttt cca ggt gtt gtc	1728
Gly Asn His Phe Gly Val Ser Phe Val Pro Leu Phe Pro Gly Val Val	
565 570 575	
cag aaa cag aag atg tgc act ggg gta gac tcg tca agt tgt tgg aag	1776
Gln Lys Gln Lys Met Cys Thr Gly Val Asp Ser Ser Ser Cys Trp Lys	
580 585 590	
gga cta tct ggt gtt ctc tat act gcc tct agc tgg gct gag ata cat	1824
Gly Leu Ser Gly Val Leu Tyr Thr Ala Ser Ser Trp Ala Glu Ile His	
595 600 605	
gct ttt gta ttg ggg tca tga	1845
Ala Phe Val Leu Gly Ser	
610	

<210> 10

<211> 614

5 <212> PRT

<213> *Nicotiana tomentosiformis*

<400> 10

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

Phe Ser Asn Ser Asn Ala Leu Phe Gly Leu Arg Phe Gln Leu Tyr Pro
20 25 30

Arg Tyr Ser Arg Tyr Leu Arg Ser Pro Val Thr Met Ala Ser Ala Lys
35 40 45

Pro Lys Pro Ala Ala Ala Val Asn Lys Phe Pro Val Glu Glu Glu Cys
50 55 60

10

ES 2 806 732 T3

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

ES 2 806 732 T3

305		310		315		320
Leu Thr Cys Thr Val Val Asp Ser Ser Ala Ile Leu Ala Glu Ile Ala						
		325		330		335
Ile Ile Thr Ala Pro Arg Ser Asp Gln Asn Arg Pro Glu Asn Gln Ile						
		340		345		350
Ala Arg Met Leu Ser Ala Asp Leu Arg Asn Thr Trp Gly Asp Leu Ser						
		355		360		365
Lys Gln Tyr Thr Glu Glu Tyr Glu Gln Cys Ile Glu Lys Met Leu Leu						
		370		375		380
Thr Glu Lys Ala Glu Lys Phe Asp Tyr Glu Arg Leu His Lys Thr Leu						
		385		390		400
Glu Glu Leu Ser Asp Phe Glu Lys Arg Ala Asn Thr Arg Val Thr Glu						
		405		410		415
Ser Gly Val Leu Lys Gly Leu Asn Leu Glu Asp Ile Lys Arg Ala Gly						
		420		425		430
Gln Arg Leu Ile Leu Gln Asp Gly Cys Thr Asn Phe Phe Gln Ser Ile						
		435		440		445
Ile Arg Asn Glu Asn Leu Asn Ala Asp Ile His Val Leu Ser Tyr Cys						
		450		455		460
Trp Cys Gly Asp Leu Ile Arg Ser Ser Phe Ser Ser Gly Gly Ile Asp						
		465		470		480
Ala Leu Asn Val His Ala Asn Glu Phe Met Phe Gln Glu Ser Leu Ser						
		485		490		495
Thr Gly Glu Ile Val Lys Lys Val Glu Ser Pro Ile Asp Lys Val Gln						
		500		505		510
Ala Phe Ser Lys Ile Arg Met Asn Cys Gly Asn Asp Gln Lys Asn Leu						
		515		520		525
Thr Leu Tyr Ile Gly Asp Ser Val Gly Asp Leu Leu Cys Leu Leu Glu						
		530		535		540
Ala Asp Val Gly Ile Val Leu Gly Thr Ser Ser Ser Leu Arg Thr Val						
		545		550		560

ES 2 806 732 T3

Gly Asn His Phe Gly Val Ser Phe Val Pro Leu Phe Pro Gly Val Val
565 570 575

Gln Lys Gln Lys Met Cys Thr Gly Val Asp Ser Ser Ser Cys Trp Lys
580 585 590

Gly Leu Ser Gly Val Leu Tyr Thr Ala Ser Ser Trp Ala Glu Ile His
595 600 605

Ala Phe Val Leu Gly Ser
610

<210> 11

<211> 1779

<212> ADN

<213> *Populus trichocarpa*

<220>

<221> CDS

<222> (1)..(1779)

<223> gen de *Populus trichocarpa* que codifica para TMP fosfatasa [XP_002325785.2]

<400> 11

atg cgc cta ctc ttg ttt act tct cca aac cca atc aaa acc tct tca 48
Met Arg Leu Leu Leu Phe Thr Ser Pro Asn Pro Ile Lys Thr Ser Ser
1 5 10 15

tca cta tat ttc ctc aac tcg ctc cga tcc aac tta acc aaa cgc acc 96
Ser Leu Tyr Phe Leu Asn Ser Leu Arg Ser Asn Leu Thr Lys Arg Thr
20 25 30

ttg cca act cgg aga tct ttc atc cct gca aga atg gca atc cct cca 144
Leu Pro Thr Arg Arg Ser Phe Ile Pro Ala Arg Met Ala Ile Pro Pro
35 40 45

cga tca ata gca tca gcg cca tct tgc act aca aca tca ggc aga agt 192
Arg Ser Ile Ala Ser Ala Pro Ser Cys Thr Thr Ser Gly Arg Ser
50 55 60

aac atc aac att gaa gag ggt ctt gct agt aaa ttc tgg atc aag ttt 240
Asn Ile Asn Ile Glu Glu Gly Leu Ala Ser Lys Phe Trp Ile Lys Phe
65 70 75 80

aga aga gaa tcc gtt ttt gct atg tac act cct ttt gtc atc tct ttg 288
Arg Arg Glu Ser Val Phe Ala Met Tyr Thr Pro Phe Val Ile Ser Leu
85 90 95

gct tct ggc act ctc aag att gat tct ttc agg cat tat atc tct caa 336
Ala Ser Gly Thr Leu Lys Ile Asp Ser Phe Arg His Tyr Ile Ser Gln
100 105 110

gat tct cac ttt ctc aaa tct ttt gct cat gcg ttt gaa tta gcg gaa 384
Asp Ser His Phe Leu Lys Ser Phe Ala His Ala Phe Glu Leu Ala Glu
115 120 125

gag tgt gct gat gat gat gaa gca aag cta gca atc tcc gag ttg agg 432
Glu Cys Ala Asp Asp Asp Glu Ala Lys Leu Ala Ile Ser Glu Leu Arg

ES 2 806 732 T3

130	135	140	
aag ggt gtc tta gag gag ctg aag atg cac aat tca ttt gta cag gaa Lys Gly Val Leu Glu Glu Leu Lys Met His Asn Ser Phe Val Gln Glu 145 150 155 160			480
tgg ggt ata gac cca ggt aaa gag ggg act atc aat tct gct act gta Trp Gly Ile Asp Pro Gly Lys Glu Gly Thr Ile Asn Ser Ala Thr Val 165 170 175			528
aaa tac aca gat ttc ttg ttg gct aca gct tct ggg aag gtt gaa gga Lys Tyr Thr Asp Phe Leu Leu Ala Thr Ala Ser Gly Lys Val Glu Gly 180 185 190			576
gtg aaa ggt ctt ggt aaa ctt gca act cct ttt gaa aga aca aaa gtt Val Lys Gly Leu Gly Lys Leu Ala Thr Pro Phe Glu Arg Thr Lys Val 195 200 205			624
gca gcc tat act ctg ggt gcc atg aca cct tgc atg cgg ctg tat tcc Ala Ala Tyr Thr Leu Gly Ala Met Thr Pro Cys Met Arg Leu Tyr Ser 210 215 220			672
ttt cta ggc aag gaa ctc cag gca gtt tta gat ccg gag gaa gat ggg Phe Leu Gly Lys Glu Leu Gln Ala Val Leu Asp Pro Glu Glu Asp Gly 225 230 235 240			720
cac cct tac aag aag tgg att gac agt tat tcg tct gag agt ttt cag His Pro Tyr Lys Lys Trp Ile Asp Ser Tyr Ser Ser Glu Ser Phe Gln 245 250 255			768
gca tca gct ctg caa act gaa gac ttg ctg gat aaa ctt agt gtc tcc Ala Ser Ala Leu Gln Thr Glu Asp Leu Leu Asp Lys Leu Ser Val Ser 260 265 270			816
ttg aca ggc gag gag ctt gac atc att gaa aag ctt tat cac cag gcc Leu Thr Gly Glu Glu Leu Asp Ile Ile Glu Lys Leu Tyr His Gln Ala 275 280 285			864
atg aaa ctt gaa ata gaa ttc ttc ctt gct cag cca att gct cag aca Met Lys Leu Glu Ile Glu Phe Phe Leu Ala Gln Pro Ile Ala Gln Thr 290 295 300			912
act tta gct ccc ctg aca aaa ggg cat aac cct gaa gaa gac cgg ctt Thr Leu Ala Pro Leu Thr Lys Gly His Asn Pro Glu Glu Asp Arg Leu 305 310 315 320			960
gtc ata ttt tct gat ttt gat ttg aca tgc act gtt gtt gac tct tct Val Ile Phe Ser Asp Phe Asp Leu Thr Cys Thr Val Val Asp Ser Ser 325 330 335			1008
gcc att ttg gca gaa att gca ata cta aca gca cca aaa tct gat gtg Ala Ile Leu Ala Glu Ile Ala Ile Leu Thr Ala Pro Lys Ser Asp Val 340 345 350			1056
gtt caa cct gag act caa att gct cga atg tca tca gct gat ctg agg Val Gln Pro Glu Thr Gln Ile Ala Arg Met Ser Ser Ala Asp Leu Arg 355 360 365			1104
aac aca tgg ggt ctt ctt tct gga cag tac acg gaa gag tat gaa caa Asn Thr Trp Gly Leu Leu Ser Gly Gln Tyr Thr Glu Glu Tyr Glu Gln 370 375 380			1152
tgt att gaa agc att atg cca tct gca aaa gtg gaa ttc aac tat gaa			1200

ES 2 806 732 T3

Cys	Ile	Glu	Ser	Ile	Met	Pro	Ser	Ala	Lys	Val	Glu	Phe	Asn	Tyr	Glu		
385					390					395					400		
gct	ctt	tgt	aaa	gca	ctt	gaa	caa	ctt	tca	gac	ttt	gag	cga	agg	gca		1248
Ala	Leu	Cys	Lys	Ala	Leu	Glu	Gln	Leu	Ser	Asp	Phe	Glu	Arg	Arg	Ala		
				405				410					415				
aat	tct	aga	gtg	att	gat	tct	gga	gtt	ctc	aaa	ggt	ttg	aat	ctt	gaa		1296
Asn	Ser	Arg		Val	Ile	Asp	Ser	Gly	Val	Leu	Lys	Gly	Leu	Asn	Leu	Glu	
			420					425					430				
gat	gta	aaa	cga	gcg	ggt	gaa	cgt	ttg	att	ctt	cag	gat	ggt	tgc	att		1344
Asp	Val	Lys	Arg	Ala	Gly	Glu	Arg	Leu	Ile	Leu	Gln	Asp	Gly	Cys	Ile		
		435					440					445					
ggt	ttc	ttt	cag	aaa	att	gtg	aag	aat	gaa	aat	ttg	aac	act	aat	gtc		1392
Gly	Phe	Phe	Gln	Lys	Ile	Val	Lys	Asn	Glu	Asn	Leu	Asn	Thr	Asn	Val		
	450					455					460						
cat	gtg	ctc	tca	tac	tgc	tgg	tgt	ggt	gat	ctc	atc	aga	tca	gct	ttc		1440
His	Val	Leu	Ser	Tyr	Cys	Trp	Cys	Gly	Asp	Leu	Ile	Arg	Ser	Ala	Phe		
465					470					475					480		
tcc	tca	ggg	ggt	ttg	gat	gct	cta	aat	att	cat	gca	aat	gag	tta	att		1488
Ser	Ser	Gly	Gly	Leu	Asp	Ala	Leu	Asn	Ile	His	Ala	Asn	Glu	Leu	Ile		
				485					490					495			
ttt	gaa	gaa	tca	atc	tcc	acg	gga	gag	att	aac	ttg	act	gtt	tac	att		1536
Phe	Glu	Glu	Ser	Ile	Ser	Thr	Gly	Glu	Ile	Asn	Leu	Thr	Val	Tyr	Ile		
			500					505					510				
gga	gat	tca	gtt	ggt	gac	ttg	ctt	tgt	cta	ctt	cag	gca	gat	att	ggt		1584
Gly	Asp	Ser	Val	Gly	Asp	Leu	Leu	Cys	Leu	Leu	Gln	Ala	Asp	Ile	Gly		
		515				520						525					
att	gta	gtt	gga	tct	agt	gca	agc	tta	agg	agc	gtg	gga	agt	caa	tat		1632
Ile	Val	Val	Gly	Ser	Ser	Ala	Ser	Leu	Arg	Ser	Val	Gly	Ser	Gln	Tyr		
	530					535					540						
ggt	gtt	tct	ttt	gta	cca	ctg	ttc	cct	ggc	ttg	gta	aga	aaa	cag	aaa		1680
Gly	Val	Ser	Phe	Val	Pro	Leu	Phe	Pro	Gly	Leu	Val	Arg	Lys	Gln	Lys		
545					550				555					560			
gaa	tct	gat	gga	gaa	tct	cct	aat	tgg	aaa	ggg	cta	tct	ggc	ata	cta		1728
Glu	Ser	Asp	Gly	Glu	Ser	Pro	Asn	Trp	Lys	Gly	Leu	Ser	Gly	Ile	Leu		
				565					570					575			
tat	aca	gtc	tcc	agt	tgg	tca	gaa	ata	cat	gcc	ttc	att	ttg	ggg	tgg		1776
Tyr	Thr	Val	Ser	Ser	Trp	Ser	Glu	Ile	His	Ala	Phe	Ile	Leu	Gly	Trp		
			580					585					590				
tag																	1779

<210> 12

<211> 592

5 <212> PRT

<213> *Populus trichocarpa*

<400> 12

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

ES 2 806 732 T3

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

ES 2 806 732 T3

260	265	270
Leu Thr Gly Glu Glu Leu Asp Ile Ile Glu Lys Leu Tyr His Gln Ala 275 280 285		
Met Lys Leu Glu Ile Glu Phe Phe Leu Ala Gln Pro Ile Ala Gln Thr 290 295 300		
Thr Leu Ala Pro Leu Thr Lys Gly His Asn Pro Glu Glu Asp Arg Leu 305 310 315 320		
Val Ile Phe Ser Asp Phe Asp Leu Thr Cys Thr Val Val Asp Ser Ser 325 330 335		
Ala Ile Leu Ala Glu Ile Ala Ile Leu Thr Ala Pro Lys Ser Asp Val 340 345 350		
Val Gln Pro Glu Thr Gln Ile Ala Arg Met Ser Ser Ala Asp Leu Arg 355 360 365		
Asn Thr Trp Gly Leu Leu Ser Gly Gln Tyr Thr Glu Glu Tyr Glu Gln 370 375 380		
Cys Ile Glu Ser Ile Met Pro Ser Ala Lys Val Glu Phe Asn Tyr Glu 385 390 395 400		
Ala Leu Cys Lys Ala Leu Glu Gln Leu Ser Asp Phe Glu Arg Arg Ala 405 410 415		
Asn Ser Arg Val Ile Asp Ser Gly Val Leu Lys Gly Leu Asn Leu Glu 420 425 430		
Asp Val Lys Arg Ala Gly Glu Arg Leu Ile Leu Gln Asp Gly Cys Ile 435 440 445		
Gly Phe Phe Gln Lys Ile Val Lys Asn Glu Asn Leu Asn Thr Asn Val 450 455 460		
His Val Leu Ser Tyr Cys Trp Cys Gly Asp Leu Ile Arg Ser Ala Phe 465 470 475 480		
Ser Ser Gly Gly Leu Asp Ala Leu Asn Ile His Ala Asn Glu Leu Ile 485 490 495		
Phe Glu Glu Ser Ile Ser Thr Gly Glu Ile Asn Leu Thr Val Tyr Ile 500 505 510		

ES 2 806 732 T3

Gly Asp Ser Val Gly Asp Leu Leu Cys Leu Leu Gln Ala Asp Ile Gly
515 520 525

Ile Val Val Gly Ser Ser Ala Ser Leu Arg Ser Val Gly Ser Gln Tyr
530 535 540

Gly Val Ser Phe Val Pro Leu Phe Pro Gly Leu Val Arg Lys Gln Lys
545 550 555 560

Glu Ser Asp Gly Glu Ser Pro Asn Trp Lys Gly Leu Ser Gly Ile Leu
565 570 575

Tyr Thr Val Ser Ser Trp Ser Glu Ile His Ala Phe Ile Leu Gly Trp
580 585 590

<210> 13

<211> 1860

5 <212> ADN

<213> *Jatropha curcas*

<220>

<221> CDS

10 <222> (1)..(1860)

<223> gen de *Jatropha curcas* que codifica para TMP fosfatasa [KDP23738.1]

<400> 13

atg gcg atc cct cca aag cta gct tcc tca tcg tct tcc atg gcc gcc 48
Met Ala Ile Pro Pro Lys Leu Ala Ser Ser Ser Ser Met Ala Ala
1 5 10 15

tcc cct act tct gct ggt gga acc aac gag gaa ggc ctc gct agt aaa 96
Ser Pro Thr Ser Ala Gly Gly Thr Asn Glu Glu Gly Leu Ala Ser Lys
20 25 30

ttc tgg atc aag ttt cgc cga gaa tcg gtt ctc gct atg tac act cct 144
Phe Trp Ile Lys Phe Arg Arg Glu Ser Val Leu Ala Met Tyr Thr Pro
35 40 45

ttc gtc gtc tct ttt gcc gcc ggc aac ctc aag att gag agt ttt agg 192
Phe Val Val Ser Phe Ala Ala Gly Asn Leu Lys Ile Glu Ser Phe Arg
50 55 60

cat tac atc gct cag gat ttt cac ttc ctc aaa gcc ttc gct cac gcg 240
His Tyr Ile Ala Gln Asp Phe His Phe Leu Lys Ala Phe Ala His Ala
65 70 75 80

tat gaa ttg gca gaa gag tgt gct gat gat gat gcc aag cta gct 288
Tyr Glu Leu Ala Glu Glu Cys Ala Asp Asp Asp Ala Lys Leu Ala
85 90 95

att gcc gcg ttg agg aag ggg gtc tta gag gag ctg aag ttg cat aaa 336
Ile Ala Ala Leu Arg Lys Gly Val Leu Glu Glu Leu Lys Leu His Lys
100 105 110

tca ttt gta cag gaa tgg ggt atg gac cct tcc aaa gag gtg act atc 384
Ser Phe Val Gln Glu Trp Gly Met Asp Pro Ser Lys Glu Val Thr Ile
115 120 125

15

ES 2 806 732 T3

aat tct gca act gca aaa tac aca gat ttc ttg ttg gct aca gct tct Asn Ser Ala Thr Ala Lys Tyr Thr Asp Phe Leu Leu Ala Thr Ala Ser 130 135 140	432
gga aag gtt gaa gga gtg aaa ggt cct ggt aaa ctt gca act cct ttt Gly Lys Val Glu Gly Val Lys Gly Pro Gly Lys Leu Ala Thr Pro Phe 145 150 155 160	480
gaa aga aca aaa gtt gca gct tac act ctt ggt acc atg aca ccc tgt Glu Arg Thr Lys Val Ala Ala Tyr Thr Leu Gly Thr Met Thr Pro Cys 165 170 175	528
atg agg ttg tat gcc ttt cta gct aag gag ctg caa gca cta ata gat Met Arg Leu Tyr Ala Phe Leu Ala Lys Glu Leu Gln Ala Leu Ile Asp 180 185 190	576
gca gaa gct ggt att cat cct tac cag aag tgg att gac aat tac tca Ala Glu Ala Gly Ile His Pro Tyr Gln Lys Trp Ile Asp Asn Tyr Ser 195 200 205	624
tct gag agt ttt cag gca tca gct ctg caa act gaa gac ttg ctg gat Ser Glu Ser Phe Gln Ala Ser Ala Leu Gln Thr Glu Asp Leu Leu Asp 210 215 220	672
aaa ctt agt gtc cct ttg aca ggc gaa gag ctt gac atc att gaa aag Lys Leu Ser Val Pro Leu Thr Gly Glu Glu Leu Asp Ile Ile Glu Lys 225 230 235 240	720
ctt tat cac caa gcc atg aaa ctt gaa ata gag ttc ttc aat gcg cag Leu Tyr His Gln Ala Met Lys Leu Glu Ile Glu Phe Phe Asn Ala Gln 245 250 255	768
cca ctt gat cag ccc act gtg gtt cct ctg aca aaa gag cat aac cct Pro Leu Asp Gln Pro Thr Val Val Pro Leu Thr Lys Glu His Asn Pro 260 265 270	816
cta gaa gat cgc ctc gtg ata ttt tct gat ttt gat ttg aca tgc aca Leu Glu Asp Arg Leu Val Ile Phe Ser Asp Phe Asp Leu Thr Cys Thr 275 280 285	864
gtt gtt gat tcc tct gcc att ttg gca gag att gca att tta aca gca Val Val Asp Ser Ser Ala Ile Leu Ala Glu Ile Ala Ile Leu Thr Ala 290 295 300	912
tca aaa tct gat cag tca caa tct gat aat caa aat gct agg atg tca Ser Lys Ser Asp Gln Ser Gln Ser Asp Asn Gln Asn Ala Arg Met Ser 305 310 315 320	960
tca act gag cta agg aac aca tgg gtt ctt ctc tct gga cag tat act Ser Thr Glu Leu Arg Asn Thr Trp Val Leu Leu Ser Gly Gln Tyr Thr 325 330 335	1008
gaa gaa tat gag caa tgc att gaa agc att ctg ccc tct gaa aaa atg Glu Glu Tyr Glu Gln Cys Ile Glu Ser Ile Leu Pro Ser Glu Lys Met 340 345 350	1056
gag ttc aac ttt gaa gct ttg tgt aaa gca ctc gaa caa ctc tca gac Glu Phe Asn Phe Glu Ala Leu Cys Lys Ala Leu Glu Gln Leu Ser Asp 355 360 365	1104
ttt gag cga agg gca aat gct aga gtt atc aaa tct gga gtt ctt aag Phe Glu Arg Arg Ala Asn Ala Arg Val Ile Lys Ser Gly Val Leu Lys	1152

ES 2 806 732 T3

370	375	380	
ggt ttg aat ctt gaa gac ata aaa cga gct gtg gag ttc aac ttt gaa Gly Leu Asn Leu Glu Asp Ile Lys Arg Ala Val Glu Phe Asn Phe Glu 385 390 395 400			1200
gct ttg tgt aaa gca ctc gaa caa ctc tca gac ttt gag cga agg gca Ala Leu Cys Lys Ala Leu Glu Gln Leu Ser Asp Phe Glu Arg Arg Ala 405 410 415			1248
aat gct aga gtt atc aaa tct gga gtt ctt aag ggt ttg aat ctt gaa Asn Ala Arg Val Ile Lys Ser Gly Val Leu Lys Gly Leu Asn Leu Glu 420 425 430			1296
gac ata aaa cga gct ggt gaa aga ctg att ctt caa gat ggc tgc acc Asp Ile Lys Arg Ala Gly Glu Arg Leu Ile Leu Gln Asp Gly Cys Thr 435 440 445			1344
agt ttt ttt cag aaa atc tcg aag aat gaa aat ctg aat gct aat ata Ser Phe Phe Gln Lys Ile Ser Lys Asn Glu Asn Leu Asn Ala Asn Ile 450 455 460			1392
cat ttc ctc tca tat tgt tgg tgt gct gat ctg atc aga tct gct ttc His Phe Leu Ser Tyr Cys Trp Cys Ala Asp Leu Ile Arg Ser Ala Phe 465 470 475 480			1440
tca tca ggg ggt ttg gat gtt ctg aat ata cat gcg aat gag ttt gat Ser Ser Gly Gly Leu Asp Val Leu Asn Ile His Ala Asn Glu Phe Asp 485 490 495			1488
ttc gta gaa tca att tca acg ggt gag att att atg aag gtg gaa acc Phe Val Glu Ser Ile Ser Thr Gly Glu Ile Ile Met Lys Val Glu Thr 500 505 510			1536
cct aca gac aaa gcc caa gct ttt aat aat att tta atg aac tac agc Pro Thr Asp Lys Ala Gln Ala Phe Asn Asn Ile Leu Met Asn Tyr Ser 515 520 525			1584
cct gac aaa aag aat ttg act gtt tat att gga gac tca gtt ggg gac Pro Asp Lys Lys Asn Leu Thr Val Tyr Ile Gly Asp Ser Val Gly Asp 530 535 540			1632
ttg ctt tgt ctg ctt gcg gca gat ata ggc atc gtg atc gga tca agc Leu Leu Cys Leu Leu Ala Ala Asp Ile Gly Ile Val Ile Gly Ser Ser 545 550 555 560			1680
tcc agc cta agg aga gtc gga agt cag ttt ggt gta aca ttt tta cca Ser Ser Leu Arg Arg Val Gly Ser Gln Phe Gly Val Thr Phe Leu Pro 565 570 575			1728
ttg tat cct ggc ttg gtt aaa aaa cag aga gag tat act gaa gga agc Leu Tyr Pro Gly Leu Val Lys Lys Gln Arg Glu Tyr Thr Glu Gly Ser 580 585 590			1776
tct tgg aat tgg aag ggt caa tct ggc gtt ctg tac aca gtt tct agt Ser Trp Asn Trp Lys Gly Gln Ser Gly Val Leu Tyr Thr Val Ser Ser 595 600 605			1824
tgg gct gaa ata cat tcc ttc gtt ttg gga tgg tag Trp Ala Glu Ile His Ser Phe Val Leu Gly Trp 610 615			1860

<210> 14
 <211> 619
 <212> PRT
 <213> *Jatropha curcas*

<400> 14

ES 2 806 732 T3

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

ES 2 806 732 T3

Lys Leu Ser Val Pro Leu Thr Gly Glu Glu Leu Asp Ile Ile Glu Lys
 225 230 235 240
 Leu Tyr His Gln Ala Met Lys Leu Glu Ile Glu Phe Phe Asn Ala Gln
 245 250 255
 Pro Leu Asp Gln Pro Thr Val Val Pro Leu Thr Lys Glu His Asn Pro
 260 265 270
 Leu Glu Asp Arg Leu Val Ile Phe Ser Asp Phe Asp Leu Thr Cys Thr
 275 280 285
 Val Val Asp Ser Ser Ala Ile Leu Ala Glu Ile Ala Ile Leu Thr Ala
 290 295 300
 Ser Lys Ser Asp Gln Ser Gln Ser Asp Asn Gln Asn Ala Arg Met Ser
 305 310 315 320
 Ser Thr Glu Leu Arg Asn Thr Trp Val Leu Leu Ser Gly Gln Tyr Thr
 325 330 335
 Glu Glu Tyr Glu Gln Cys Ile Glu Ser Ile Leu Pro Ser Glu Lys Met
 340 345 350
 Glu Phe Asn Phe Glu Ala Leu Cys Lys Ala Leu Glu Gln Leu Ser Asp
 355 360 365
 Phe Glu Arg Arg Ala Asn Ala Arg Val Ile Lys Ser Gly Val Leu Lys
 370 375 380
 Gly Leu Asn Leu Glu Asp Ile Lys Arg Ala Val Glu Phe Asn Phe Glu
 385 390 395 400
 Ala Leu Cys Lys Ala Leu Glu Gln Leu Ser Asp Phe Glu Arg Arg Ala
 405 410 415
 Asn Ala Arg Val Ile Lys Ser Gly Val Leu Lys Gly Leu Asn Leu Glu
 420 425 430
 Asp Ile Lys Arg Ala Gly Glu Arg Leu Ile Leu Gln Asp Gly Cys Thr
 435 440 445
 Ser Phe Phe Gln Lys Ile Ser Lys Asn Glu Asn Leu Asn Ala Asn Ile
 450 455 460
 His Phe Leu Ser Tyr Cys Trp Cys Ala Asp Leu Ile Arg Ser Ala Phe
 465 470 475 480

ES 2 806 732 T3

Ser Ser Gly Gly Leu Asp Val Leu Asn Ile His Ala Asn Glu Phe Asp
485 490 495

Phe Val Glu Ser Ile Ser Thr Gly Glu Ile Ile Met Lys Val Glu Thr
500 505 510

Pro Thr Asp Lys Ala Gln Ala Phe Asn Asn Ile Leu Met Asn Tyr Ser
515 520 525

Pro Asp Lys Lys Asn Leu Thr Val Tyr Ile Gly Asp Ser Val Gly Asp
530 535 540

Leu Leu Cys Leu Leu Ala Ala Asp Ile Gly Ile Val Ile Gly Ser Ser
545 550 555 560

Ser Ser Leu Arg Arg Val Gly Ser Gln Phe Gly Val Thr Phe Leu Pro
565 570 575

Leu Tyr Pro Gly Leu Val Lys Lys Gln Arg Glu Tyr Thr Glu Gly Ser
580 585 590

Ser Trp Asn Trp Lys Gly Gln Ser Gly Val Leu Tyr Thr Val Ser Ser
595 600 605

Trp Ala Glu Ile His Ser Phe Val Leu Gly Trp
610 615

<210> 15

<211> 1842

5 <212> ADN

<213> *Citrus sinensis*

<220>

<221> CDS

10 <222> (1)..(1842)

<223> gen de *Citrus sinensis* que codifica para TMP fosfatasa [XP_006484613.1]

<400> 15

atg cgc ttc ctt ttc aca aac cca atc aaa acc cca tta ctc tct tct 48
Met Arg Phe Leu Phe Thr Asn Pro Ile Lys Thr Pro Leu Leu Ser Ser
1 5 10 15

att ctt ttc cat tgt ccc aac tcg ccc cga ctc ggc ctt ctt gac tca 96
Ile Leu Phe His Cys Pro Asn Ser Pro Arg Leu Gly Leu Leu Asp Ser
20 25 30

gtc cga gtc aac tca cct tct tct ttg aca act caa aga tcg tca ctt 144
Val Arg Val Asn Ser Pro Ser Ser Leu Thr Thr Gln Arg Ser Ser Leu
35 40 45

15 tcg atg gcg gcg att ccc cca aaa tcg ccg agc cct gag gag gag gga 192
Ser Met Ala Ala Ile Pro Pro Lys Ser Pro Ser Pro Glu Glu Glu Gly

ES 2 806 732 T3

50	55	60	
ctc gcg agg agg ttg tgg atc aag ttt aag aga gaa tct gtg ttt gcc Leu Ala Arg Arg Leu Trp Ile Lys Phe Lys Arg Glu Ser Val Phe Ala 65 70 75 80			240
atg tac tcc ccg ttt acg gtt tgt ttg gct tct ggg aac cta aag ctt Met Tyr Ser Pro Phe Thr Val Cys Leu Ala Ser Gly Asn Leu Lys Leu 85 90 95			288
gaa acc ttc agg cat tac atc gcc caa gat ttt cat ttt ctc aaa gct Glu Thr Phe Arg His Tyr Ile Ala Gln Asp Phe His Phe Leu Lys Ala 100 105 110			336
ttc gcc caa gcg tat gaa ctg gcg gaa gaa tgt gct gat gat gat gat Phe Ala Gln Ala Tyr Glu Leu Ala Glu Glu Cys Ala Asp Asp Asp Asp 115 120 125			384
gca aag tta tct atc tct gaa ttg agg aag ggt gta ctt gag gag tta Ala Lys Leu Ser Ile Ser Glu Leu Arg Lys Gly Val Leu Glu Glu Leu 130 135 140			432
aaa atg cat gat tcc ttt gtg aag gag tgg ggt aca gat ctt gct aaa Lys Met His Asp Ser Phe Val Lys Glu Trp Gly Thr Asp Leu Ala Lys 145 150 155 160			480
atg gct act gtt aac tct gca act gta aag tat aca gag ttc ttg ttg Met Ala Thr Val Asn Ser Ala Thr Val Lys Tyr Thr Glu Phe Leu Leu 165 170 175			528
gca aca gct tcc ggg aag gtc gaa ggt gtt aaa ggt cct gga aaa ctt Ala Thr Ala Ser Gly Lys Val Glu Gly Val Lys Gly Pro Gly Lys Leu 180 185 190			576
gca acc cca ttt gag aaa act aaa gtt gcc gct tac aca ttg ggt gcc Ala Thr Pro Phe Glu Lys Thr Lys Val Ala Ala Tyr Thr Leu Gly Ala 195 200 205			624
atg tca cct tgt atg agg ctc tat gct ttc ctt gga aag gaa ttc cat Met Ser Pro Cys Met Arg Leu Tyr Ala Phe Leu Gly Lys Glu Phe His 210 215 220			672
ggc ctc cta aat gct aat gaa ggc aat cat cct tac aag aag tgg att Gly Leu Leu Asn Ala Asn Glu Gly Asn His Pro Tyr Lys Lys Trp Ile 225 230 235 240			720
gac aat tat tct tct gaa agt ttt cag gcc tca gct ctg caa aat gag Asp Asn Tyr Ser Ser Glu Ser Phe Gln Ala Ser Ala Leu Gln Asn Glu 245 250 255			768
gac ttg ctg gat aaa ctt agt gtc tct ttg aca ggc gaa gaa cta gac Asp Leu Leu Asp Lys Leu Ser Val Ser Leu Thr Gly Glu Glu Leu Asp 260 265 270			816
ata ata gaa aag ctc tat cac caa gcc atg aaa ctt gaa gta gag ttc Ile Ile Glu Lys Leu Tyr His Gln Ala Met Lys Leu Glu Val Glu Phe 275 280 285			864
ttc tgt gct cag cca ctt gct cag ccc act gta gtt cct ctg att aaa Phe Cys Ala Gln Pro Leu Ala Gln Pro Thr Val Val Pro Leu Ile Lys 290 295 300			912
ggg cat aat cct gca gga gac cgt cta att ata ttt tct gat ttc gat			960

ES 2 806 732 T3

Gly 305	His	Asn	Pro	Ala	Gly 310	Asp	Arg	Leu	Ile	Ile 315	Phe	Ser	Asp	Phe	Asp 320	
ttg	act	tg	acc	att	gtt	gat	tcc	tct	gcc	att	ttg	gca	gag	atc	gca	1008
Leu	Thr	Cys	Thr	Ile	Val	Asp	Ser	Ser	Ala	Ile	Leu	Ala	Glu	Ile	Ala	
				325					330					335		
ata	gtg	aca	gca	cca	aaa	tct	gac	cag	aat	caa	cct	gaa	aat	caa	ctt	1056
Ile	Val	Thr	Ala	Pro	Lys	Ser	Asp	Gln	Asn	Gln	Pro	Glu	Asn	Gln	Leu	
			340					345					350			
ggt	cgg	atg	tca	tca	ggt	gag	ctg	agg	aac	aca	tgg	ggt	ctt	ctt	tcc	1104
Gly	Arg	Met	Ser	Ser	Gly	Glu	Leu	Arg	Asn	Thr	Trp	Gly	Leu	Leu	Ser	
		355					360					365				
aaa	cag	tac	aca	gag	gag	tac	gaa	caa	tgc	att	gaa	agc	ttc	atg	ccc	1152
Lys	Gln	Tyr	Thr	Glu	Glu	Tyr	Glu	Gln	Cys	Ile	Glu	Ser	Phe	Met	Pro	
	370					375					380					
tct	gag	aaa	gtg	gag	aat	ttc	aac	tat	gaa	act	ttg	cat	aaa	gca	ctt	1200
Ser	Glu	Lys	Val	Glu	Asn	Phe	Asn	Tyr	Glu	Thr	Leu	His	Lys	Ala	Leu	
385					390					395					400	
gag	caa	ctc	tca	cac	ttt	gag	aag	agg	gca	aat	tct	aga	gtg	atc	gaa	1248
Glu	Gln	Leu	Ser	His	Phe	Glu	Lys	Arg	Ala	Asn	Ser	Arg	Val	Ile	Glu	
				405					410					415		
tct	gga	gtt	ctc	aag	ggt	ata	aat	ctt	gaa	gat	att	aaa	aaa	gct	ggt	1296
Ser	Gly	Val	Leu	Lys	Gly	Ile	Asn	Leu	Glu	Asp	Ile	Lys	Lys	Ala	Gly	
			420					425					430			
gaa	cgc	ctg	agt	ctt	caa	gat	ggt	tgc	act	acc	ttc	ttt	cag	aaa	gtt	1344
Glu	Arg	Leu	Ser	Leu	Gln	Asp	Gly	Cys	Thr	Thr	Phe	Phe	Gln	Lys	Val	
		435					440					445				
gta	aag	aat	gaa	aat	ttg	aat	gct	aat	gtc	cat	gtg	ctt	tca	tac	tgt	1392
Val	Lys	Asn	Glu	Asn	Leu	Asn	Ala	Asn	Val	His	Val	Leu	Ser	Tyr	Cys	
	450					455					460					
tgg	tgt	ggt	gat	ctc	atc	aga	gca	tct	ttt	tct	tca	gca	ggt	tta	aat	1440
Trp	Cys	Gly	Asp	Leu	Ile	Arg	Ala	Ser	Phe	Ser	Ser	Ala	Gly	Leu	Asn	
465					470					475					480	
gca	ctg	aat	gta	cat	gcg	aat	gag	ttc	tca	ttc	aaa	gaa	tct	att	tca	1488
Ala	Leu	Asn	Val	His	Ala	Asn	Glu	Phe	Ser	Phe	Lys	Glu	Ser	Ile	Ser	
				485					490					495		
acg	ggt	gaa	att	att	gag	aaa	gtg	gag	tcc	ccc	att	gac	aaa	gtt	caa	1536
Thr	Gly	Glu	Ile	Ile	Glu	Lys	Val	Glu	Ser	Pro	Ile	Asp	Lys	Val	Gln	
			500						505				510			
gct	ttc	aac	aat	act	tta	gag	aaa	tac	gga	act	gac	aga	aag	aac	ttg	1584
Ala	Phe	Asn	Asn	Thr	Leu	Glu	Lys	Tyr	Gly	Thr	Asp	Arg	Lys	Asn	Leu	
		515					520					525				
agt	gtt	tac	att	gga	gac	tct	gtg	ggt	gac	ttg	ctt	tgt	ctg	ctt	gag	1632
Ser	Val	Tyr	Ile	Gly	Asp	Ser	Val	Gly	Asp	Leu	Leu	Cys	Leu	Leu	Glu	
		530					535					540				
gct	gat	ata	ggc	att	gta	atc	ggg	tct	agc	tca	agc	tta	agg	aga	gtg	1680
Ala	Asp	Ile	Gly	Ile	Val	Ile	Gly	Ser	Ser	Ser	Ser	Leu	Arg	Arg	Val	
545					550					555					560	

ES 2 806 732 T3

gga tct caa ttt ggt gtt aca ttt atc ccg ttg tac cct ggc ttg gtt 1728
Gly Ser Gln Phe Gly Val Thr Phe Ile Pro Leu Tyr Pro Gly Leu Val
565 570 575

aag aaa cag aag gag tac act gaa gga agc tct tct aac tgg aag gag 1776
Lys Lys Gln Lys Glu Tyr Thr Glu Gly Ser Ser Ser Asn Trp Lys Glu
580 585 590

aaa tct ggc ata ctt tac aca gtc tca agt tgg gct gaa gta cat gcc 1824
Lys Ser Gly Ile Leu Tyr Thr Val Ser Ser Trp Ala Glu Val His Ala
595 600 605

ttt atc ttg ggg tgg tag 1842
Phe Ile Leu Gly Trp
610

<210> 16

<211> 613

<212> PRT

<213> *Citrus sinensis*

<400> 16

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

Ile Leu Phe His Cys Pro Asn Ser Pro Arg Leu Gly Leu Leu Asp Ser
20 25 30

Val Arg Val Asn Ser Pro Ser Ser Leu Thr Thr Gln Arg Ser Ser Leu
35 40 45

Ser Met Ala Ala Ile Pro Pro Lys Ser Pro Ser Pro Glu Glu Glu Gly
50 55 60

Leu Ala Arg Arg Leu Trp Ile Lys Phe Lys Arg Glu Ser Val Phe Ala
65 70 75 80

Met Tyr Ser Pro Phe Thr Val Cys Leu Ala Ser Gly Asn Leu Lys Leu
85 90 95

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

Phe Ala Gln Ala Tyr Glu Leu Ala Glu Glu Cys Ala Asp Asp Asp Asp
115 120 125

Ala Lys Leu Ser Ile Ser Glu Leu Arg Lys Gly Val Leu Glu Glu Leu
130 135 140

Lys Met His Asp Ser Phe Val Lys Glu Trp Gly Thr Asp Leu Ala Lys
145 150 155 160

5

10

ES 2 806 732 T3

Met Ala Thr Val Asn Ser Ala Thr Val Lys Tyr Thr Glu Phe Leu Leu
165 170 175

Ala Thr Ala Ser Gly Lys Val Glu Gly Val Lys Gly Pro Gly Lys Leu
180 185 190

Ala Thr Pro Phe Glu Lys Thr Lys Val Ala Ala Tyr Thr Leu Gly Ala
195 200 205

Met Ser Pro Cys Met Arg Leu Tyr Ala Phe Leu Gly Lys Glu Phe His
210 215 220

Gly Leu Leu Asn Ala Asn Glu Gly Asn His Pro Tyr Lys Lys Trp Ile
225 230 235 240

Asp Asn Tyr Ser Ser Glu Ser Phe Gln Ala Ser Ala Leu Gln Asn Glu
245 250 255

Asp Leu Leu Asp Lys Leu Ser Val Ser Leu Thr Gly Glu Glu Leu Asp
260 265 270

Ile Ile Glu Lys Leu Tyr His Gln Ala Met Lys Leu Glu Val Glu Phe
275 280 285

Phe Cys Ala Gln Pro Leu Ala Gln Pro Thr Val Val Pro Leu Ile Lys
290 295 300

Gly His Asn Pro Ala Gly Asp Arg Leu Ile Ile Phe Ser Asp Phe Asp
305 310 315 320

Leu Thr Cys Thr Ile Val Asp Ser Ser Ala Ile Leu Ala Glu Ile Ala
325 330 335

Ile Val Thr Ala Pro Lys Ser Asp Gln Asn Gln Pro Glu Asn Gln Leu
340 345 350

Gly Arg Met Ser Ser Gly Glu Leu Arg Asn Thr Trp Gly Leu Leu Ser
355 360 365

Lys Gln Tyr Thr Glu Glu Tyr Glu Gln Cys Ile Glu Ser Phe Met Pro
370 375 380

Ser Glu Lys Val Glu Asn Phe Asn Tyr Glu Thr Leu His Lys Ala Leu
385 390 395 400

Glu Gln Leu Ser His Phe Glu Lys Arg Ala Asn Ser Arg Val Ile Glu
405 410 415

ES 2 806 732 T3

Ser Gly Val Leu Lys Gly Ile Asn Leu Glu Asp Ile Lys Lys Ala Gly
420 425 430

Glu Arg Leu Ser Leu Gln Asp Gly Cys Thr Thr Phe Phe Gln Lys Val
435 440 445

Val Lys Asn Glu Asn Leu Asn Ala Asn Val His Val Leu Ser Tyr Cys
450 455 460

Trp Cys Gly Asp Leu Ile Arg Ala Ser Phe Ser Ser Ala Gly Leu Asn
465 470 475 480

Ala Leu Asn Val His Ala Asn Glu Phe Ser Phe Lys Glu Ser Ile Ser
485 490 495

Thr Gly Glu Ile Ile Glu Lys Val Glu Ser Pro Ile Asp Lys Val Gln
500 505 510

Ala Phe Asn Asn Thr Leu Glu Lys Tyr Gly Thr Asp Arg Lys Asn Leu
515 520 525

Ser Val Tyr Ile Gly Asp Ser Val Gly Asp Leu Leu Cys Leu Leu Glu
530 535 540

Ala Asp Ile Gly Ile Val Ile Gly Ser Ser Ser Ser Leu Arg Arg Val
545 550 555 560

Gly Ser Gln Phe Gly Val Thr Phe Ile Pro Leu Tyr Pro Gly Leu Val
565 570 575

Lys Lys Gln Lys Glu Tyr Thr Glu Gly Ser Ser Ser Asn Trp Lys Glu
580 585 590

Lys Ser Gly Ile Leu Tyr Thr Val Ser Ser Trp Ala Glu Val His Ala
595 600 605

Phe Ile Leu Gly Trp
610

<210> 17

<211> 1728

5 <212> ADN

<213> *Prunus persica*

<220>

<221> CDS

10 <222> (1)..(1728)

<223> gen de *Prunus persica* que codifica para TMP fosfatasa [XP_007199656.1]

<400> 17

ES 2 806 732 T3

atg gcg gca ttg gct cgt cat agc att gtt aga ctc aat cac gaa gga Met Ala Ala Leu Ala Arg His Ser Ile Val Arg Leu Asn His Glu Gly 1 5 10 15	48
ggc cta gcc aga cgg ctg tgg ttc aag ttc aga gac gac tct gtt ttc Gly Leu Ala Arg Arg Leu Trp Phe Lys Phe Arg Asp Asp Ser Val Phe 20 25 30	96
tct ctc tac act ccc ttc ttc gtt ggc tta gcc tct gct act ctg cac Ser Leu Tyr Thr Pro Phe Phe Val Gly Leu Ala Ser Ala Thr Leu His 35 40 45	144
tct gaa act acc ttt cgc cat ttc atc tct cag gac ctc cat ttt ctc Ser Glu Thr Thr Phe Arg His Phe Ile Ser Gln Asp Leu His Phe Leu 50 55 60	192
aaa gcc ttc gtt ctc gca tat gaa ttg gcg gaa gat tgt gct gat gat Lys Ala Phe Val Leu Ala Tyr Glu Leu Ala Asp Cys Ala Asp Asp 65 70 75 80	240
gag gac gac aag aat ggt tta cgc gat ttg aga aaa cgt gcc gtc ggc Glu Asp Asp Lys Asn Gly Leu Arg Asp Leu Arg Lys Arg Ala Val Gly 85 90 95	288
agg ctt caa atg cac gac aca ttt gtc cga gaa tgg ggt ttt gaa ttc Arg Leu Gln Met His Asp Thr Phe Val Arg Glu Trp Gly Phe Glu Phe 100 105 110	336
cca aat gag gac att tct aaa gac att gca aca acc aaa tac aca gat Pro Asn Glu Asp Ile Ser Lys Asp Ile Ala Thr Thr Lys Tyr Thr Asp 115 120 125	384
ttc ttg ctt gca aca gca tca ggg aaa att gaa gga gaa aga tcg gtt Phe Leu Leu Ala Thr Ala Ser Gly Lys Ile Glu Gly Glu Arg Ser Val 130 135 140	432
ctg gac aaa atc gca acc cct ttc gaa aag acc aag gtt gct gca tat Leu Asp Lys Ile Ala Thr Pro Phe Glu Lys Thr Lys Val Ala Ala Tyr 145 150 155 160	480
aca ctt gct gct ctg gct cct tgt atg aga ctc tat gcc ttc atc agt Thr Leu Ala Ala Leu Ala Pro Cys Met Arg Leu Tyr Ala Phe Ile Ser 165 170 175	528
act gag atc caa ggc att ata aat cct gat caa gat agc act cac att Thr Glu Ile Gln Gly Ile Ile Asn Pro Asp Gln Asp Ser Thr His Ile 180 185 190	576
tac aaa agc tgg ata gaa aat tat tcg tct caa gtt ttc gag gaa ata Tyr Lys Ser Trp Ile Glu Asn Tyr Ser Ser Gln Val Phe Glu Glu Ile 195 200 205	624
gcc ctg caa aat gaa gac atg cta gat aaa ctt agt gtt tct ttg act Ala Leu Gln Asn Glu Asp Met Leu Asp Lys Leu Ser Val Ser Leu Thr 210 215 220	672
ggt gag gag ctt gag att ata gag aag ctc tat cat caa gct atg aag Gly Glu Glu Leu Glu Ile Ile Glu Lys Leu Tyr His Gln Ala Met Lys 225 230 235 240	720

ES 2 806 732 T3

ctt caa gta gat ttt att gct gct caa cca att tct gat cag caa tct Leu Gln Val Asp Phe Ile Ala Ala Gln Pro Ile Ser Asp Gln Gln Ser 245 250 255	768
gta gtt cct ttg tct cgg gtg cat gac ttt agc aaa cgc cat ctt acg Val Val Pro Leu Ser Arg Val His Asp Phe Ser Lys Arg His Leu Thr 260 265 270	816
ata ctt tgt gac ttt gat ttg gca tgc act gct ttt gat tct gct gcc Ile Leu Cys Asp Phe Asp Leu Ala Cys Thr Ala Phe Asp Ser Ala Ala 275 280 285	864
ata ttg gct gag att gcg atc ata aca gca cca aag gct gat atg gat Ile Leu Ala Glu Ile Ala Ile Ile Thr Ala Pro Lys Ala Asp Met Asp 290 295 300	912
gga tct gat caa acc caa ctt gct cgg atg cca tca gca gac tta agg Gly Ser Asp Gln Thr Gln Leu Ala Arg Met Pro Ser Ala Asp Leu Arg 305 310 315 320	960
agc aca tgg gat gtt ctt tca acc caa tac act gaa caa ttt gaa caa Ser Thr Trp Asp Val Leu Ser Thr Gln Tyr Thr Glu Gln Phe Glu Gln 325 330 335	1008
tgt gta gaa agc att gtg gcc agt gag aga gtg gaa gaa ttc gat tat Cys Val Glu Ser Ile Val Ala Ser Glu Arg Val Glu Glu Phe Asp Tyr 340 345 350	1056
gaa cgt ctg tgt agc gcg ctt gaa caa ctt gcg gag ttt gag aga aag Glu Arg Leu Cys Ser Ala Leu Glu Gln Leu Ala Glu Phe Glu Arg Lys 355 360 365	1104
gca aat gaa agg gtg gtt cag tca gga gtg ttg aag ggt tta aat gcg Ala Asn Glu Arg Val Val Gln Ser Gly Val Leu Lys Gly Leu Asn Ala 370 375 380	1152
gag gat ata aaa agg gct gga cag agc ctc att ctg caa gat ggt tgc Glu Asp Ile Lys Arg Ala Gly Gln Ser Leu Ile Leu Gln Asp Gly Cys 385 390 395 400	1200
aga agc ttc ttt cag aag att gtg aaa aat aaa aat ctg aaa act gat Arg Ser Phe Phe Gln Lys Ile Val Lys Asn Lys Asn Leu Lys Thr Asp 405 410 415	1248
gtt cat gtg ctt tca tac tgc tgg tgt aat gac ctc att gta tca gct Val His Val Leu Ser Tyr Cys Trp Cys Asn Asp Leu Ile Val Ser Ala 420 425 430	1296
ttc tct tca gga gat ttg aat gtc ttg aat gta cat tca aat gag ttg Phe Ser Ser Gly Asp Leu Asn Val Leu Asn Val His Ser Asn Glu Leu 435 440 445	1344
gtt tat caa gaa tct gtc aca act ggt gaa att gta aag aag atg gag Val Tyr Gln Glu Ser Val Thr Thr Gly Glu Ile Val Lys Lys Met Glu 450 455 460	1392
tct ccc atg gaa aag ctt caa gtc ttc aac gac gtc cta atc gac cgc Ser Pro Met Glu Lys Leu Gln Val Phe Asn Asp Val Leu Ile Asp Arg 465 470 475 480	1440
agg ggc gaa ggc aat aaa cac ttg aca gtt tac att gga ggc tca gtg Arg Gly Glu Gly Asn Lys His Leu Thr Val Tyr Ile Gly Gly Ser Val 485 490 495	1488

ES 2 806 732 T3

ggt gac ttg ctt tgc ctg ctt gaa gca gat ata ggc att gta gtt ggt 1536
 Gly Asp Leu Leu Cys Leu Leu Glu Ala Asp Ile Gly Ile Val Val Gly
 500 505 510
 tca agt tca agc cta agg aga cta ggt gat cat ttt ggt gtt tcc ttt 1584
 Ser Ser Ser Ser Leu Arg Arg Leu Gly Asp His Phe Gly Val Ser Phe
 515 520 525
 gtc cca ttg ttc tct ggc ttg gtg aag agg cag aaa gaa ctt gct gat 1632
 Val Pro Leu Phe Ser Gly Leu Val Lys Arg Gln Lys Glu Leu Ala Asp
 530 535 540
 caa gat tgt gcc tct aat tgg tgg aaa cca ttg tct ggt gtt ctt tat 1680
 Gln Asp Cys Ala Ser Asn Trp Trp Lys Pro Leu Ser Gly Val Leu Tyr
 545 550 555 560
 acg gtg tct agt tgg gct gaa ata cag gca ttc att ttg ggt aca tag 1728
 Thr Val Ser Ser Trp Ala Glu Ile Gln Ala Phe Ile Leu Gly Thr
 565 570 575

<210> 18
 <211> 575
 <212> PRT
 <213> *Prunus persica*

<400> 18

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

5

10

ES 2 806 732 T3

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

ES 2 806 732 T3

```

385                390                395                400

Arg Ser Phe Phe Gln Lys Ile Val Lys Asn Lys Asn Leu Lys Thr Asp
      405                410                415

Val His Val Leu Ser Tyr Cys Trp Cys Asn Asp Leu Ile Val Ser Ala
      420                425                430

Phe Ser Ser Gly Asp Leu Asn Val Leu Asn Val His Ser Asn Glu Leu
      435                440                445

Val Tyr Gln Glu Ser Val Thr Thr Gly Glu Ile Val Lys Lys Met Glu
      450                455                460

Ser Pro Met Glu Lys Leu Gln Val Phe Asn Asp Val Leu Ile Asp Arg
465                470                475                480

Arg Gly Glu Gly Asn Lys His Leu Thr Val Tyr Ile Gly Gly Ser Val
      485                490                495

Gly Asp Leu Leu Cys Leu Leu Glu Ala Asp Ile Gly Ile Val Val Gly
      500                505                510

Ser Ser Ser Ser Leu Arg Arg Leu Gly Asp His Phe Gly Val Ser Phe
      515                520                525

Val Pro Leu Phe Ser Gly Leu Val Lys Arg Gln Lys Glu Leu Ala Asp
      530                535                540

Gln Asp Cys Ala Ser Asn Trp Trp Lys Pro Leu Ser Gly Val Leu Tyr
545                550                555                560

Thr Val Ser Ser Trp Ala Glu Ile Gln Ala Phe Ile Leu Gly Thr
      565                570                575

```

<210> 19

<211> 1944

5 <212> ADN

<213> *Phoenix dactylifera*

<220>

<221> CDS

10 <222> (1)..(1944)

<223> gen de *Phoenix_dactylifera* que codifica para TMP fosfatasa [XP_008796407]

<400> 19

```

atg cga ttc ctc tcc cct ctt ctc ccc ctc cgc cga aac cca aac cct
Met Arg Phe Leu Ser Pro Leu Leu Pro Leu Arg Arg Asn Pro Asn Pro
15 1                5                10                15

```

48

ES 2 806 732 T3

agc cct agg ttc ttc tcg ctc tcc cct ccc ata tcc ctc gcc tcc gcc Ser Pro Arg Phe Phe Ser Leu Ser Pro Pro Ile Ser Leu Ala Ser Ala 20 25 30	96
tgc ccc cga ttc ggt ttc ttg aat cga gat cgc ccc cgg cgc cgc ctt Cys Pro Arg Phe Gly Phe Leu Asn Arg Asp Arg Pro Arg Arg Arg Leu 35 40 45	144
cca aag ggg ttc cga tcg atc gcc gcg gcg aat cag cgg gcg tcg cct Pro Lys Gly Phe Arg Ser Ile Ala Ala Ala Asn Gln Arg Ala Ser Pro 50 55 60	192
cca aga ttg gtg ccg gag agg gcg gcc gcc acg agt tct tgg cct tct Pro Arg Leu Val Pro Glu Arg Ala Ala Ala Thr Ser Ser Trp Pro Ser 65 70 75 80	240
tca gcc gga cga gcc atg gca gtg gtg gcg acg gcg gtt gaa gaa ggc Ser Ala Gly Arg Ala Met Ala Val Val Ala Thr Ala Val Glu Glu Gly 85 90 95	288
tcc gcg gcg aag cgg ttc tgg atc agg tcc cgg aag gag gcg gtg ttc Ser Ala Ala Lys Arg Phe Trp Ile Arg Ser Arg Lys Glu Ala Val Phe 100 105 110	336
gcg gag tac acc ccg ttc gtg gtg tgc ctg gcg gcg ggg aga ctg gag Ala Glu Tyr Thr Pro Phe Val Val Cys Leu Ala Ala Gly Arg Leu Glu 115 120 125	384
atg gag gcc ttc cgc gac tac att gct cag gac gtg cac ttc ctc aat Met Glu Ala Phe Arg Asp Tyr Ile Ala Gln Asp Val His Phe Leu Asn 130 135 140	432
act ttt gcc caa gcg tat gag atg gcg gaa gag tgt gct gat gat gat Thr Phe Ala Gln Ala Tyr Glu Met Ala Glu Glu Cys Ala Asp Asp Asp 145 150 155 160	480
gat gcg aag gct gca ata act gat ctg agg aaa gct gtt ttg gag gaa Asp Ala Lys Ala Ala Ile Thr Asp Leu Arg Lys Ala Val Leu Glu Glu 165 170 175	528
ctg aaa atg cat agt tca ttt gtc caa gaa tgg gga ata gac ccc act Leu Lys Met His Ser Ser Phe Val Gln Glu Trp Gly Ile Asp Pro Thr 180 185 190	576
aaa gaa atc att cct ttc cct gca aca gta aag tac acc gac ttc ctg Lys Glu Ile Ile Pro Phe Pro Ala Thr Val Lys Tyr Thr Asp Phe Leu 195 200 205	624
ctt gct aca gct gca gga aaa gtt gaa gga ggg aaa gat cct ggg aaa Leu Ala Thr Ala Ala Gly Lys Val Glu Gly Gly Lys Asp Pro Gly Lys 210 215 220	672
att gtc act cct ttt gag aag aca aaa att gct gct tat act gta ggt Ile Val Thr Pro Phe Glu Lys Thr Lys Ile Ala Ala Tyr Thr Val Gly 225 230 235 240	720
gcc atg gct cct tgc atg agg ctt tat gca ttc ttg gga aaa gag ctc Ala Met Ala Pro Cys Met Arg Leu Tyr Ala Phe Leu Gly Lys Glu Leu 245 250 255	768
cag acg tgt ctg caa ctt gac gaa aat tgt cat ccc tac aaa aag tgg Gln Thr Cys Leu Gln Leu Asp Glu Asn Cys His Pro Tyr Lys Lys Trp 260 265 270	816

ES 2 806 732 T3

att gat aat tat tcc tct gaa agt ttt gag aca gct gct gtg caa ata Ile Asp Asn Tyr Ser Ser Glu Ser Phe Glu Thr Ala Ala Val Gln Ile 275 280 285	864
gaa gaa ttg ctt gac aaa ttg agt gtt tca ttg act ggg gag gag ctt Glu Glu Leu Leu Asp Lys Leu Ser Val Ser Leu Thr Gly Glu Glu Leu 290 295 300	912
gaa gac ata gaa aag ctt tac cgc caa gct atg aaa ctt gaa att gaa Glu Asp Ile Glu Lys Leu Tyr Arg Gln Ala Met Lys Leu Glu Ile Glu 305 310 315 320	960
ttt ttt ctt gct cag cca att gtc cga cca gct gta gtt cct ttg aca Phe Phe Leu Ala Gln Pro Ile Val Arg Pro Ala Val Val Pro Leu Thr 325 330 335	1008
aga ctg cat gat ccg gca aat tgc ctt gtc att ttt tct gat ttt gac Arg Leu His Asp Pro Ala Asn Cys Leu Val Ile Phe Ser Asp Phe Asp 340 345 350	1056
ttg aca tgc agt gta gtt gat tcc tct gcc att tta gca gag att gca Leu Thr Cys Ser Val Val Asp Ser Ser Ala Ile Leu Ala Glu Ile Ala 355 360 365	1104
ata tta agt gca cca aag act gat aag act ggg act gat aat tta gat Ile Leu Ser Ala Pro Lys Thr Asp Lys Thr Gly Thr Asp Asn Leu Asp 370 375 380	1152
gct cga agg tct tct tca gaa atg aga aac tca tgg gat gct ctt tct Ala Arg Arg Ser Ser Ser Glu Met Arg Asn Ser Trp Asp Ala Leu Ser 385 390 395 400	1200
aaa cag tat aca gaa gag tat gag cag tgc ata gaa agc tta ctt cca Lys Gln Tyr Thr Glu Glu Tyr Glu Gln Cys Ile Glu Ser Leu Leu Pro 405 410 415	1248
tta gaa gaa gct aaa aca ttt gat tat gaa ggc ctt tgc aaa agt ttg Leu Glu Glu Ala Lys Thr Phe Asp Tyr Glu Gly Leu Cys Lys Ser Leu 420 425 430	1296
ggc cag ctc tct gag ttt gag aaa cga gca aat tcc agg gtt att gag Gly Gln Leu Ser Glu Phe Glu Lys Arg Ala Asn Ser Arg Val Ile Glu 435 440 445	1344
tct ggg gtg cta aag gga atg aat cta gat gac ata aaa aga gct ggg Ser Gly Val Leu Lys Gly Met Asn Leu Asp Asp Ile Lys Arg Ala Gly 450 455 460	1392
gaa cgt ttg atc ctc caa gat ggt tgt ata gat ttt ttt cag aag gtt Glu Arg Leu Ile Leu Gln Asp Gly Cys Ile Asp Phe Phe Gln Lys Val 465 470 475 480	1440
gta aag gaa aag gaa aat cta aat tta gat ctc cat gta ctt tct tat Val Lys Glu Lys Glu Asn Leu Asn Leu Asp Leu His Val Leu Ser Tyr 485 490 495	1488
tgt tgg tgt gcg gat cta ata agg tca gct ttt tca tca gta ggt tgc Cys Trp Cys Ala Asp Leu Ile Arg Ser Ala Phe Ser Ser Val Gly Cys 500 505 510	1536
cta aat gat ttg aac ata cac tca aat gag ttc aac tat caa gaa tct Leu Asn Asp Leu Asn Ile His Ser Asn Glu Phe Asn Tyr Gln Glu Ser	1584

ES 2 806 732 T3

515	520	525	
att tca acg ggt gaa att gtt agg aag atg gaa tca ccc atg gac aag			1632
Ile Ser Thr Gly Glu Ile Val Arg Lys Met Glu Ser Pro Met Asp Lys			
530	535	540	
gtt gaa gca ttc aaa agt atc tta agc aac ctt gga agc aat gag aag			1680
Val Glu Ala Phe Lys Ser Ile Leu Ser Asn Leu Gly Ser Asn Glu Lys			
545	550	555	560
cgc tta tct gtg tac att gga gat tcg gtt ggt gac ttg ctt tgc ctg			1728
Arg Leu Ser Val Tyr Ile Gly Asp Ser Val Gly Asp Leu Leu Cys Leu			
	565	570	575
ttg gaa gca gat gtt ggt att gtg att gga tca agc act agc tta agg			1776
Leu Glu Ala Asp Val Gly Ile Val Ile Gly Ser Ser Thr Ser Leu Arg			
	580	585	590
aga atc ggg aag cag ttt ggt gtt tct ttc att cca ctc ttc cgt ggt			1824
Arg Ile Gly Lys Gln Phe Gly Val Ser Phe Ile Pro Leu Phe Arg Gly			
	595	600	605
ttg gta aac aag caa aga caa ctt aat gaa aaa gac tca tct atc tgg			1872
Leu Val Asn Lys Gln Arg Gln Leu Asn Glu Lys Asp Ser Ser Ile Trp			
	610	615	620
aag ggg ttg tct ggt gtt ctt tat aca gca tca agc tgg tca gaa ata			1920
Lys Gly Leu Ser Gly Val Leu Tyr Thr Ala Ser Ser Trp Ser Glu Ile			
	625	630	635
caa gct ttt att ttg ggg gca taa			1944
Gln Ala Phe Ile Leu Gly Ala			
	645		

<210> 20
 <211> 647
 <212> PRT
 <213> *Phoenix dactylifera*

<400> 20

Met Arg Phe Leu Ser Pro Leu Leu Pro Leu Arg Arg Asn Pro Asn Pro	
1	15
Ser Pro Arg Phe Phe Ser Leu Ser Pro Pro Ile Ser Leu Ala Ser Ala	
20	30
Cys Pro Arg Phe Gly Phe Leu Asn Arg Asp Arg Pro Arg Arg Arg Leu	
35	45
Pro Lys Gly Phe Arg Ser Ile Ala Ala Ala Asn Gln Arg Ala Ser Pro	
50	60
Pro Arg Leu Val Pro Glu Arg Ala Ala Ala Thr Ser Ser Trp Pro Ser	
65	80
Ser Ala Gly Arg Ala Met Ala Val Val Ala Thr Ala Val Glu Glu Gly	

ES 2 806 732 T3

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

ES 2 806 732 T3

Arg Leu His Asp Pro Ala Asn Cys Leu Val Ile Phe Ser Asp Phe Asp
 340 345 350
 Leu Thr Cys Ser Val Val Asp Ser Ser Ala Ile Leu Ala Glu Ile Ala
 355 360 365
 Ile Leu Ser Ala Pro Lys Thr Asp Lys Thr Gly Thr Asp Asn Leu Asp
 370 375 380
 Ala Arg Arg Ser Ser Ser Glu Met Arg Asn Ser Trp Asp Ala Leu Ser
 385 390 395 400
 Lys Gln Tyr Thr Glu Glu Tyr Glu Gln Cys Ile Glu Ser Leu Leu Pro
 405 410 415
 Leu Glu Glu Ala Lys Thr Phe Asp Tyr Glu Gly Leu Cys Lys Ser Leu
 420 425 430
 Gly Gln Leu Ser Glu Phe Glu Lys Arg Ala Asn Ser Arg Val Ile Glu
 435 440 445
 Ser Gly Val Leu Lys Gly Met Asn Leu Asp Asp Ile Lys Arg Ala Gly
 450 455 460
 Glu Arg Leu Ile Leu Gln Asp Gly Cys Ile Asp Phe Phe Gln Lys Val
 465 470 475 480
 Val Lys Glu Lys Glu Asn Leu Asn Leu Asp Leu His Val Leu Ser Tyr
 485 490 495
 Cys Trp Cys Ala Asp Leu Ile Arg Ser Ala Phe Ser Ser Val Gly Cys
 500 505 510
 Leu Asn Asp Leu Asn Ile His Ser Asn Glu Phe Asn Tyr Gln Glu Ser
 515 520 525
 Ile Ser Thr Gly Glu Ile Val Arg Lys Met Glu Ser Pro Met Asp Lys
 530 535 540
 Val Glu Ala Phe Lys Ser Ile Leu Ser Asn Leu Gly Ser Asn Glu Lys
 545 550 555 560
 Arg Leu Ser Val Tyr Ile Gly Asp Ser Val Gly Asp Leu Leu Cys Leu
 565 570 575
 Leu Glu Ala Asp Val Gly Ile Val Ile Gly Ser Ser Thr Ser Leu Arg
 580 585 590

ES 2 806 732 T3

Arg Ile Gly Lys Gln Phe Gly Val Ser Phe Ile Pro Leu Phe Arg Gly
595 600 605

Leu Val Asn Lys Gln Arg Gln Leu Asn Glu Lys Asp Ser Ser Ile Trp
610 615 620

Lys Gly Leu Ser Gly Val Leu Tyr Thr Ala Ser Ser Trp Ser Glu Ile
625 630 635 640

Gln Ala Phe Ile Leu Gly Ala
645

<210> 21

<211> 1860

<212> ADN

<213> *Zea mays*

<220>

<221> CDS

<222> (1)..(1860)

<223> gen de *Zea mays* que codifica para TMP fosfatasa [XP_008678418.1]

<400> 21

atg ctt gtt ctc cgc cgt ctc cgc ctc cgc ctc cca ctg cca cgc cct 48
Met Leu Val Leu Arg Arg Leu Arg Leu Arg Leu Pro Leu Pro Arg Pro
1 5 10 15

ctt ctc gtc tcc tcc ttc tcc tcc acc tcc ccc tcc tcc tca ccc tcg 96
Leu Leu Val Ser Ser Phe Ser Ser Thr Ser Pro Ser Ser Ser Pro Ser
20 25 30

acc tct agc tcc tcc tcc tgt tgg tcg tcg aca ggc gaa agt aga agg 144
Thr Ser Ser Ser Ser Cys Trp Ser Ser Thr Gly Glu Ser Arg Arg
35 40 45

gcc atg gcg tca tct cct tct ccc gat tcg gcc gcg gtc gtt gcc gag 192
Ala Met Ala Ser Ser Pro Ser Pro Asp Ser Ala Ala Val Val Ala Glu
50 55 60

ggc tcc gcg gct cgc cgc ttc tgg atc gct gcc tcc acg cgc gag gcc 240
Gly Ser Ala Ala Arg Arg Phe Trp Ile Ala Ala Ser Thr Arg Glu Ala
65 70 75 80

gcc ttc gcc gca tac acg ccc ttc ctc ctc tcc ctc gcc gcc ggc aat 288
Ala Phe Ala Ala Tyr Thr Pro Phe Leu Leu Ser Leu Ala Ala Gly Asn
85 90 95

ctg cgg ctc aac gtg ttt cgc cac tac atc gcg cag gac gcg cac ttc 336
Leu Arg Leu Asn Val Phe Arg His Tyr Ile Ala Gln Asp Ala His Phe
100 105 110

ctt cac gcc ttc gct cgc gcg tac gaa atg gcc gag gac tgc gct gat 384
Leu His Ala Phe Ala Arg Ala Tyr Glu Met Ala Glu Asp Cys Ala Asp
115 120 125

gat gac gac gac atg gcc acc ata gcc gcc ctc agg aag gcc atc ctc 432
Asp Asp Asp Asp Met Ala Thr Ile Ala Ala Leu Arg Lys Ala Ile Leu

ES 2 806 732 T3

130	135	140	
caa gag ctc aac ctc cac tcc tcc gtt ctg aag gag tgg gga gtt gat Gln Glu Leu Asn Leu His Ser Ser Val Leu Lys Glu Trp Gly Val Asp 145 150 155 160			480
cct acc aaa gag ata cct cca agt gca gct aca acc aaa tat act gat Pro Thr Lys Glu Ile Pro Pro Ser Ala Thr Thr Lys Tyr Thr Asp 165 170 175			528
ttc cta ctt gca act gcg gct gga aaa gtt gat ggc aca aaa ggt tct Phe Leu Leu Ala Thr Ala Ala Gly Lys Val Asp Gly Thr Lys Gly Ser 180 185 190			576
gac aaa atg gtt act cca ttt gag aag act aaa att gct gca tac act Asp Lys Met Val Thr Pro Phe Glu Lys Thr Lys Ile Ala Ala Tyr Thr 195 200 205			624
gtt ggg gcc atg act cca tgc atg agg ctt tat gca tat cta ggc aaa Val Gly Ala Met Thr Pro Cys Met Arg Leu Tyr Ala Tyr Leu Gly Lys 210 215 220			672
gaa ctc atg gtt ttc ctt aaa caa gat gaa aat cat cca tac aag aaa Glu Leu Met Val Phe Leu Lys Gln Asp Glu Asn His Pro Tyr Lys Lys 225 230 235 240			720
tgg att aac aca tat gca tcc agt gat ttt gag gac acc aca ctc caa Trp Ile Asn Thr Tyr Ala Ser Ser Asp Phe Glu Asp Thr Thr Leu Gln 245 250 255			768
ata gaa gaa ttg cta gac aaa cta agt gtc tca tta act ggt gag gaa Ile Glu Glu Leu Leu Asp Lys Leu Ser Val Ser Leu Thr Gly Glu Glu 260 265 270			816
ctt gag att att ggc aag ctc tac cag caa gct atg aaa ctg gaa gtg Leu Glu Ile Ile Gly Lys Leu Tyr Gln Gln Ala Met Lys Leu Glu Val 275 280 285			864
gag ttc ttt tct tct cag ctt ata gac caa cct gtt gta gct cca ctt Glu Phe Phe Ser Ser Gln Leu Ile Asp Gln Pro Val Val Ala Pro Leu 290 295 300			912
tca aga tac tgt gat cca aaa tat aaa ctc ttg atc ttt tct gat ttt Ser Arg Tyr Cys Asp Pro Lys Tyr Lys Leu Leu Ile Phe Ser Asp Phe 305 310 315 320			960
gat ttg acg tgc act att gtt gat tca tct gcc att ttg gcg gag att Asp Leu Thr Cys Thr Ile Val Asp Ser Ser Ala Ile Leu Ala Glu Ile 325 330 335			1008
gca att ttg tca ttc caa aag gca aat caa agt ggg att gat aat aac Ala Ile Leu Ser Phe Gln Lys Ala Asn Gln Ser Gly Ile Asp Asn Asn 340 345 350			1056
ctc gac cgt gca aaa tcg gga gac ctg aga agt tcg tgg aac atg ctc Leu Asp Arg Ala Lys Ser Gly Asp Leu Arg Ser Ser Trp Asn Met Leu 355 360 365			1104
tct aag caa tac atg gaa gag tat gag aaa tgc atg gaa aga cta ctt Ser Lys Gln Tyr Met Glu Glu Tyr Glu Lys Cys Met Glu Arg Leu Leu 370 375 380			1152
cct cca gaa gaa tcg aag tca cta gat tat gat aaa ctg tat aaa ggc			1200

ES 2 806 732 T3

Pro	Pro	Glu	Glu	Ser	Lys	Ser	Leu	Asp	Tyr	Asp	Lys	Leu	Tyr	Lys	Gly			
385					390					395					400			
ctg	gag	gtg	cta	gct	gag	ttt	gag	aag	ctt	gca	aat	tct	agg	gtt	gtc		1248	
Leu	Glu	Val	Leu	Ala	Glu	Phe	Glu	Lys	Leu	Ala	Asn	Ser	Arg	Val	Val			
				405					410					415				
gac	tct	ggg	gtg	ctg	agg	gga	atg	aat	ttg	gaa	gac	atc	agg	aaa	gct		1296	
Asp	Ser	Gly	Val	Leu	Arg	Gly	Met	Asn	Leu	Glu	Asp	Ile	Arg	Lys	Ala			
			420					425					430					
ggg	gag	cgt	ctt	att	ctt	caa	ggg	ggc	tgt	aaa	aat	ttc	ttt	cag	aag		1344	
Gly	Glu	Arg	Leu	Ile	Leu	Gln	Gly	Gly	Cys	Lys	Asn	Phe	Phe	Gln	Lys			
		435					440					445						
att	gta	aaa	aca	agg	gag	aac	ctc	aat	ttg	gat	gtc	cat	att	ctt	tcc		1392	
Ile	Val	Lys	Thr	Arg	Glu	Asn	Leu	Asn	Leu	Asp	Val	His	Ile	Leu	Ser			
		450				455					460							
tat	tgc	tgg	tgt	gca	gaa	ctt	ata	aga	tca	gcc	ttc	tca	tca	gcc	ggg		1440	
Tyr	Cys	Trp	Cys	Ala	Glu	Leu	Ile	Arg	Ser	Ala	Phe	Ser	Ser	Ala	Gly			
465					470					475					480			
tgt	cta	gat	ggg	ttg	aac	ata	cat	tca	aat	gag	ttt	gcc	ttt	gag	gat		1488	
Cys	Leu	Asp	Gly	Leu	Asn	Ile	His	Ser	Asn	Glu	Phe	Ala	Phe	Glu	Asp			
				485					490					495				
tct	gtt	tca	act	ggg	gag	atc	gac	aga	aag	atg	cag	tct	ccg	cta	gac		1536	
Ser	Val	Ser	Thr	Gly	Glu	Ile	Asp	Arg	Lys	Met	Gln	Ser	Pro	Leu	Asp			
			500					505					510					
aaa	gtt	gaa	aag	ttc	aag	agc	atc	aga	agt	gac	gtg	gac	agt	aca	gtg		1584	
Lys	Val	Glu	Lys	Phe	Lys	Ser	Ile	Arg	Ser	Asp	Val	Asp	Ser	Thr	Val			
		515					520					525						
cca	ttc	cta	tct	gtt	tat	att	gga	gac	tcg	gtt	gga	gat	ttg	ctc	tgc		1632	
Pro	Phe	Leu	Ser	Val	Tyr	Ile	Gly	Asp	Ser	Val	Gly	Asp	Leu	Leu	Cys			
		530				535					540							
tta	ttg	gag	gct	gat	att	ggg	ata	gtc	att	ggg	tca	acc	aca	agt	ttg		1680	
Leu	Leu	Glu	Ala	Asp	Ile	Gly	Ile	Val	Ile	Gly	Ser	Thr	Thr	Ser	Leu			
545					550					555					560			
cgt	agg	gtg	ggc	aaa	cag	ttt	ggg	gtt	tct	ttt	gtc	cca	ttg	ttc	cct		1728	
Arg	Arg	Val	Gly	Lys	Gln	Phe	Gly	Val	Ser	Phe	Val	Pro	Leu	Phe	Pro			
				565					570					575				
ggg	cta	gta	gag	aag	cag	agg	caa	ctg	gcg	gag	gaa	gat	gca	tcc	gta		1776	
Gly	Leu	Val	Glu	Lys	Gln	Arg	Gln	Leu	Ala	Glu	Glu	Asp	Ala	Ser	Val			
			580					585					590					
ttc	aag	gca	cgg	tct	gga	gtc	ctc	tat	acg	gtt	tct	agc	tgg	tca	gaa		1824	
Phe	Lys	Ala	Arg	Ser	Gly	Val	Leu	Tyr	Thr	Val	Ser	Ser	Trp	Ser	Glu			
		595					600					605						
ata	cac	gcc	ttc	gta	ctg	gga	agt	gat	ttc	agc	tga						1860	
Ile	His	Ala	Phe	Val	Leu	Gly	Ser	Asp	Phe	Ser								
		610				615												

<210> 22
 <211> 619
 5 <212> PRT
 <213> *Zea mays*
 <400> 22

ES 2 806 732 T3

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

ES 2 806 732 T3

Trp Ile Asn Thr Tyr Ala Ser Ser Asp Phe Glu Asp Thr Thr Leu Gln
 245 250 255
 Ile Glu Glu Leu Leu Asp Lys Leu Ser Val Ser Leu Thr Gly Glu Glu
 260 265 270
 Leu Glu Ile Ile Gly Lys Leu Tyr Gln Gln Ala Met Lys Leu Glu Val
 275 280 285
 Glu Phe Phe Ser Ser Gln Leu Ile Asp Gln Pro Val Val Ala Pro Leu
 290 295 300
 Ser Arg Tyr Cys Asp Pro Lys Tyr Lys Leu Leu Ile Phe Ser Asp Phe
 305 310 315 320
 Asp Leu Thr Cys Thr Ile Val Asp Ser Ser Ala Ile Leu Ala Glu Ile
 325 330 335
 Ala Ile Leu Ser Phe Gln Lys Ala Asn Gln Ser Gly Ile Asp Asn Asn
 340 345 350
 Leu Asp Arg Ala Lys Ser Gly Asp Leu Arg Ser Ser Trp Asn Met Leu
 355 360 365
 Ser Lys Gln Tyr Met Glu Glu Tyr Glu Lys Cys Met Glu Arg Leu Leu
 370 375 380
 Pro Pro Glu Glu Ser Lys Ser Leu Asp Tyr Asp Lys Leu Tyr Lys Gly
 385 390 395 400
 Leu Glu Val Leu Ala Glu Phe Glu Lys Leu Ala Asn Ser Arg Val Val
 405 410 415
 Asp Ser Gly Val Leu Arg Gly Met Asn Leu Glu Asp Ile Arg Lys Ala
 420 425 430
 Gly Glu Arg Leu Ile Leu Gln Gly Gly Cys Lys Asn Phe Phe Gln Lys
 435 440 445
 Ile Val Lys Thr Arg Glu Asn Leu Asn Leu Asp Val His Ile Leu Ser
 450 455 460
 Tyr Cys Trp Cys Ala Glu Leu Ile Arg Ser Ala Phe Ser Ser Ala Gly
 465 470 475 480
 Cys Leu Asp Gly Leu Asn Ile His Ser Asn Glu Phe Ala Phe Glu Asp
 485 490 495

ES 2 806 732 T3

Ser Val Ser Thr Gly Glu Ile Asp Arg Lys Met Gln Ser Pro Leu Asp
500 505 510

Lys Val Glu Lys Phe Lys Ser Ile Arg Ser Asp Val Asp Ser Thr Val
515 520 525

Pro Phe Leu Ser Val Tyr Ile Gly Asp Ser Val Gly Asp Leu Leu Cys
530 535 540

Leu Leu Glu Ala Asp Ile Gly Ile Val Ile Gly Ser Thr Thr Ser Leu
545 550 555 560

Arg Arg Val Gly Lys Gln Phe Gly Val Ser Phe Val Pro Leu Phe Pro
565 570 575

Gly Leu Val Glu Lys Gln Arg Gln Leu Ala Glu Glu Asp Ala Ser Val
580 585 590

Phe Lys Ala Arg Ser Gly Val Leu Tyr Thr Val Ser Ser Trp Ser Glu
595 600 605

Ile His Ala Phe Val Leu Gly Ser Asp Phe Ser
610 615

<210> 23

<211> 1830

5 <212> ADN

<213> *Oryza sativa*

<220>

<221> CDS

10 <222> (1)..(1830)

<223> gen de *Oryza sativa* que codifica para TMP fosfatasa [NP_001062539.1]

<400> 23

atg	cgc	ggc	ctc	ctc	cgc	cgt	gtc	tac	ctc	cgc	ctc	ccc	cct	ttc	cct	48
Met	Arg	Gly	Leu	Leu	Arg	Arg	Val	Tyr	Leu	Arg	Leu	Pro	Pro	Phe	Pro	
1			5						10					15		

cct	gcc	acc	tct	ctt	tat	tat	tgg	tca	aga	aca	aga	cct	gca	gct	gca	96
Pro	Ala	Thr	Ser	Leu	Tyr	Tyr	Trp	Ser	Arg	Thr	Arg	Pro	Ala	Ala	Ala	
			20					25					30			

ggg	ccc	aac	cac	ccc	atc	cct	agg	cgc	atg	tcg	acg	tcc	tct	act	gcc	144
Gly	Pro	Asn	His	Pro	Ile	Pro	Arg	Arg	Met	Ser	Thr	Ser	Ser	Thr	Ala	
		35					40					45				

gcg	gcg	gtc	gtt	gcc	gag	ggc	tcc	gcc	gct	cgc	cgc	ttc	tgg	atc	gcc	192
Ala	Ala	Val	Val	Ala	Glu	Gly	Ser	Ala	Ala	Arg	Arg	Phe	Trp	Ile	Ala	
		50				55					60					

15	gcc	gcc	tcg	agg	gag	gcc	gcc	ttc	gcc	gcc	tac	acg	ccc	ttc	ctc	gtc	240
----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

ES 2 806 732 T3

Ala 65	Ala	Ser	Arg	Glu	Ala 70	Ala	Phe	Ala	Ala	Tyr 75	Thr	Pro	Phe	Leu	Val 80	
tcc	ctc	gcc	gcc	ggg	gcc	ctc	cgc	ctg	gat	tcc	ttc	cgc	caa	tac	atc	288
Ser	Leu	Ala	Ala	Gly 85	Ala	Leu	Arg	Leu	Asp 90	Ser	Phe	Arg	Gln	Tyr 95	Ile	
gcc	cag	gat	gcc	tac	ttc	ctc	cac	gcc	ttc	gcc	cgc	gcc	tat	gag	atg	336
Ala	Gln	Asp	Ala	Tyr 100	Phe	Leu	His	Ala	Phe 105	Ala	Arg	Ala	Tyr 110	Glu	Met	
gcc	gag	gag	tgc	gcc	gat	gac	gac	gac	gac	aag	gcc	acc	atc	gtc	gtc	384
Ala	Glu	Glu	Cys	Ala	Asp	Asp	Asp	Asp	Asp	Lys	Ala	Thr	Ile	Val	Val	
			115				120					125				
ctc	agg	aag	gcc	atc	ctc	cgc	gag	ctc	aac	ctc	cac	gct	tcc	gtc	ctt	432
Leu	Arg	Lys	Ala	Ile	Leu	Arg	Glu	Leu	Asn	Leu	His	Ala	Ser	Val	Leu	
			130				135					140				
cag	gaa	tgg	gga	gtc	gat	ccc	aac	aaa	gaa	atc	cct	cca	atc	cca	gcc	480
Gln	Glu	Trp	Gly	Val	Asp	Pro	Asn	Lys	Glu	Ile	Pro	Pro	Ile	Pro	Ala	
						150					155				160	
aca	act	aag	tac	act	gat	ttc	tta	ctt	gca	act	tcc	act	gga	aag	gtt	528
Thr	Thr	Lys	Tyr	Thr	Asp	Phe	Leu	Leu	Ala	Thr	Ser	Thr	Gly	Lys	Val	
						165					170				175	
gat	ggt	ggg	aaa	ggt	tct	gat	aaa	atg	gtc	aca	cca	ttc	gag	aag	acg	576
Asp	Gly	Gly	Lys	Gly	Ser	Asp	Lys	Met	Val	Thr	Pro	Phe	Glu	Lys	Thr	
			180					185					190			
aaa	att	gct	gca	tac	act	gtt	ggg	gct	atg	acc	cca	tgc	atg	agg	ctt	624
Lys	Ile	Ala	Ala	Tyr	Thr	Val	Gly	Ala	Met	Thr	Pro	Cys	Met	Arg	Leu	
			195					200				205				
tat	gcg	tat	ctg	ggc	aaa	gaa	ctt	gca	gtt	ttc	ttg	aaa	cag	gat	gaa	672
Tyr	Ala	Tyr	Leu	Gly	Lys	Glu	Leu	Ala	Val	Phe	Leu	Lys	Gln	Asp	Glu	
			210				215				220					
aat	cac	cca	tac	aag	aaa	tgg	att	gag	act	tat	gca	tcc	agt	gat	ttt	720
Asn	His	Pro	Tyr	Lys	Lys	Trp	Ile	Glu	Thr	Tyr	Ala	Ser	Ser	Asp	Phe	
					230					235					240	
gag	aat	aac	gca	ctc	caa	ata	gaa	gag	ttg	ctt	gat	aaa	cta	agt	gtc	768
Glu	Asn	Asn	Ala	Leu	Gln	Ile	Glu	Glu	Leu	Leu	Asp	Lys	Leu	Ser	Val	
				245					250					255		
tct	cta	act	ggc	gag	gag	ctt	gag	att	att	ggg	aag	ctc	tac	cag	caa	816
Ser	Leu	Thr	Gly	Glu	Glu	Leu	Glu	Ile	Ile	Gly	Lys	Leu	Tyr	Gln	Gln	
				260				265					270			
gct	atg	agg	ctg	gaa	gtt	gag	ttc	ttc	tct	gct	cag	cca	gta	gac	caa	864
Ala	Met	Arg	Leu	Glu	Val	Glu	Phe	Phe	Ser	Ala	Gln	Pro	Val	Asp	Gln	
				275			280					285				
cct	gtt	gta	gct	cca	ctc	tca	aga	tat	tgt	ggt	ccg	aaa	gat	aag	ctc	912
Pro	Val	Val	Ala	Pro	Leu	Ser	Arg	Tyr	Cys	Gly	Pro	Lys	Asp	Lys	Leu	
				290			295				300					
ttg	ata	ttt	tgt	gat	ttt	gat	ttg	aca	tgc	act	gtt	gtt	gat	tca	tct	960
Leu	Ile	Phe	Cys	Asp	Phe	Asp	Leu	Thr	Cys	Thr	Val	Val	Asp	Ser	Ser	
				305		310				315					320	

ES 2 806 732 T3

gcc att ttg gcg gag att gca atc ttg tca cac caa aag gct agt caa Ala Ile Leu Ala Glu Ile Ala Ile Leu Ser His Gln Lys Ala Ser Gln 325 330 335	1008
ggg ggg gct gat agt tcc ctt gat cgt aca aaa tca gcg gac ttg aga Gly Gly Ala Asp Ser Ser Leu Asp Arg Thr Lys Ser Ala Asp Leu Arg 340 345 350	1056
aat tca tgg aac atg ctc tca aat caa tac atg gaa gag tat gag caa Asn Ser Trp Asn Met Leu Ser Asn Gln Tyr Met Glu Glu Tyr Glu Gln 355 360 365	1104
tgc ata gca agc ttg ctt cct cca gaa gaa gca agg tca cta gac tat Cys Ile Ala Ser Leu Leu Pro Pro Glu Glu Ala Arg Ser Leu Asp Tyr 370 375 380	1152
gat caa ctg tat aaa ggt ttg gag gtg cta tcg cag ttt gag aaa ctt Asp Gln Leu Tyr Lys Gly Leu Glu Val Leu Ser Gln Phe Glu Lys Leu 385 390 395 400	1200
gca aac tct agg gtg gtt gat tct ggt gtc ctg agg gga atg aat tta Ala Asn Ser Arg Val Val Asp Ser Gly Val Leu Arg Gly Met Asn Leu 405 410 415	1248
gat gac atc cga aaa gct gga gag agg ctt att ctg caa gat gga tgc Asp Asp Ile Arg Lys Ala Gly Glu Arg Leu Ile Leu Gln Asp Gly Cys 420 425 430	1296
aaa att ttt ttt caa aag att ggc aaa aca agg gag aac ctc aat tta Lys Ile Phe Phe Gln Lys Ile Gly Lys Thr Arg Glu Asn Leu Asn Leu 435 440 445	1344
gat gtc cat att ctt tcc tat tgc tgg tgc gca gat ctt ata agg tca Asp Val His Ile Leu Ser Tyr Cys Trp Cys Ala Asp Leu Ile Arg Ser 450 455 460	1392
gct ttt tca tca gtt ggt tgt cta gac ggg ctg aac ata cat tca aat Ala Phe Ser Ser Val Gly Cys Leu Asp Gly Leu Asn Ile His Ser Asn 465 470 475 480	1440
gag ttt gct ttt gag gga tct gtt tca act ggt cat att aac aga caa Glu Phe Ala Phe Glu Gly Ser Val Ser Thr Gly His Ile Asn Arg Gln 485 490 495	1488
atg gag tct cct ctg gac aaa gct gaa aag ttc aag agc atc aaa agc Met Glu Ser Pro Leu Asp Lys Ala Glu Lys Phe Lys Ser Ile Lys Ser 500 505 510	1536
gac gtg ggt agt aca ggg aca tta ttg tca gtc tat att ggg gac tcg Asp Val Gly Ser Thr Gly Thr Leu Leu Ser Val Tyr Ile Gly Asp Ser 515 520 525	1584
gtt gga gat ttg ctt tgc ttg ttg gag gca gat att ggt att gtt gtt Val Gly Asp Leu Leu Cys Leu Leu Glu Ala Asp Ile Gly Ile Val Val 530 535 540	1632
gga tca agc aca acc ttg cgg aga gtg ggc aaa cag ttt ggt gtt tca Gly Ser Ser Thr Thr Leu Arg Arg Val Gly Lys Gln Phe Gly Val Ser 545 550 555 560	1680
ttt gtt cct ctg ttc act ggg ttg gta gag aag cag agg cga ata gaa Phe Val Pro Leu Phe Thr Gly Leu Val Glu Lys Gln Arg Arg Ile Glu 565 570 575	1728
aag gaa gaa tca tcc atc ttc aag gca cgg tct gga att ctt tat acg Lys Glu Glu Ser Ser Ile Phe Lys Ala Arg Ser Gly Ile Leu Tyr Thr 580 585 590	1776
gtt tct agc tgg tcg gag gta cag gct ttc atc ctg gga aat gat ttc Val Ser Ser Trp Ser Glu Val Gln Ala Phe Ile Leu Gly Asn Asp Phe 595 600 605	1824
agc tga Ser	1830

<210> 24

ES 2 806 732 T3

<211> 609
 <212> PRT
 <213> *Oryza sativa*

5 <400> 24

Met Arg Gly Leu Leu Arg Arg Val Tyr Leu Arg Leu Pro Pro Phe Pro
 1 5 10 15

Pro Ala Thr Ser Leu Tyr Tyr Trp Ser Arg Thr Arg Pro Ala Ala Ala
 20 25 30

Gly Pro Asn His Pro Ile Pro Arg Arg Met Ser Thr Ser Ser Thr Ala
 35 40 45

Ala Ala Val Val Ala Glu Gly Ser Ala Ala Arg Arg Phe Trp Ile Ala
 50 55 60

Ala Ala Ser Arg Glu Ala Ala Phe Ala Ala Tyr Thr Pro Phe Leu Val
 65 70 75 80

Ser Leu Ala Ala Gly Ala Leu Arg Leu Asp Ser Phe Arg Gln Tyr Ile
 85 90 95

Ala Gln Asp Ala Tyr Phe Leu His Ala Phe Ala Arg Ala Tyr Glu Met
 100 105 110

Ala Glu Glu Cys Ala Asp Asp Asp Asp Asp Lys Ala Thr Ile Val Val
 115 120 125

Leu Arg Lys Ala Ile Leu Arg Glu Leu Asn Leu His Ala Ser Val Leu
 130 135 140

Gln Glu Trp Gly Val Asp Pro Asn Lys Glu Ile Pro Pro Ile Pro Ala
 145 150 155 160

Thr Thr Lys Tyr Thr Asp Phe Leu Leu Ala Thr Ser Thr Gly Lys Val
 165 170 175

ES 2 806 732 T3

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

ES 2 806 732 T3

420 425 430
 Lys Ile Phe Phe Gln Lys Ile Gly Lys Thr Arg Glu Asn Leu Asn Leu
 435 440 445
 Asp Val His Ile Leu Ser Tyr Cys Trp Cys Ala Asp Leu Ile Arg Ser
 450 455 460
 Ala Phe Ser Ser Val Gly Cys Leu Asp Gly Leu Asn Ile His Ser Asn
 465 470 475 480
 Glu Phe Ala Phe Glu Gly Ser Val Ser Thr Gly His Ile Asn Arg Gln
 485 490 495
 Met Glu Ser Pro Leu Asp Lys Ala Glu Lys Phe Lys Ser Ile Lys Ser
 500 505 510
 Asp Val Gly Ser Thr Gly Thr Leu Leu Ser Val Tyr Ile Gly Asp Ser
 515 520 525
 Val Gly Asp Leu Leu Cys Leu Leu Glu Ala Asp Ile Gly Ile Val Val
 530 535 540
 Gly Ser Ser Thr Thr Leu Arg Arg Val Gly Lys Gln Phe Gly Val Ser
 545 550 555 560
 Phe Val Pro Leu Phe Thr Gly Leu Val Glu Lys Gln Arg Arg Ile Glu
 565 570 575
 Lys Glu Glu Ser Ser Ile Phe Lys Ala Arg Ser Gly Ile Leu Tyr Thr
 580 585 590
 Val Ser Ser Trp Ser Glu Val Gln Ala Phe Ile Leu Gly Asn Asp Phe
 595 600 605

Ser

<210> 25

<211> 1683

5 <212> ADN

<213> *Picea sitchensis*

<220>

<221> CDS

10 <222> (1)..(1683)

<223> gen de *Picea_sitchensis* que codifica para TMP fosfatasa [ABR16455]

<400> 25

ES 2 806 732 T3

atg ggg gtc gcc gat gaa gca gga gtc gcc aga agg cta tgg aca aag Met Gly Val Ala Asp Glu Ala Gly Val Ala Arg Arg Leu Trp Thr Lys 1 5 10 15	48
ttc aag aaa gac acc gcg ctt gca cag tat aat tcc ttt gtt gtt gct Phe Lys Lys Asp Thr Ala Leu Ala Gln Tyr Asn Ser Phe Val Val Ala 20 25 30	96
ttg gcg gcc ggg acg ctc aac atg acg tct ttt cag cag tac atg gcg Leu Ala Ala Gly Thr Leu Asn Met Thr Ser Phe Gln Gln Tyr Met Ala 35 40 45	144
cag gat gct tat ttt ctc aaa gca ttt gct cag gcg tac aca atg gca Gln Asp Ala Tyr Phe Leu Lys Ala Phe Ala Gln Ala Tyr Thr Met Ala 50 55 60	192
gag gat tgc gca gat gat gac gac gac gac aaa gca tcg atc cgt gaa cta Glu Asp Cys Ala Asp Asp Asp Asp Lys Ala Ser Ile Arg Glu Leu 65 70 75 80	240
cga aaa gcc gct gag gaa gag ctc aat ctg cac aat tcc ttg gct gag Arg Lys Ala Ala Glu Glu Glu Leu Asn Leu His Asn Ser Leu Ala Glu 85 90 95	288
gac tgg gac gtt gaa ttt gca aaa gag tgc tct ccc aat atg gca aca Asp Trp Asp Val Glu Phe Ala Lys Glu Cys Ser Pro Asn Met Ala Thr 100 105 110	336
gtc aag tac aca gaa ttt tta ttg gca aca gct gct ggc aag gtg gaa Val Lys Tyr Thr Glu Phe Leu Leu Ala Thr Ala Ala Gly Lys Val Glu 115 120 125	384
gga ggg aag gga cca agc aga agt gtg act cct ttt gag aaa aca aaa Gly Gly Lys Gly Pro Ser Arg Ser Val Thr Pro Phe Glu Lys Thr Lys 130 135 140	432
ata gca gca tac aca gtg ggt gcc atg acc ccg tgc atg agg ctt tat Ile Ala Ala Tyr Thr Val Gly Ala Met Thr Pro Cys Met Arg Leu Tyr 145 150 155 160	480
gct ttc ttg ggc caa gaa att gtc aaa gcc ctg gaa cct gat tgc agt Ala Phe Leu Gly Gln Glu Ile Val Lys Ala Leu Glu Pro Asp Cys Ser 165 170 175	528
aat cat cca tat aag cag tgg att gaa aca tac tct tct gca aag ttt Asn His Pro Tyr Lys Gln Trp Ile Glu Thr Tyr Ser Ser Ala Lys Phe 180 185 190	576
gag gca tcg gca tta caa act gaa gag ttg ctt gac aaa ctg gct att Glu Ala Ser Ala Leu Gln Thr Glu Glu Leu Leu Asp Lys Leu Ala Ile 195 200 205	624
tcg cta act ggg gaa gag ctt gaa gtg ctg cgg agg ttg tat tat cat Ser Leu Thr Gly Glu Glu Leu Glu Val Leu Arg Arg Leu Tyr Tyr His 210 215 220	672
gcc tta aaa cta gaa ata gaa ttc ttt tcc gct cag cct ttc tct cag Ala Leu Lys Leu Glu Ile Glu Phe Phe Ser Ala Gln Pro Phe Ser Gln 225 230 235 240	720
aga aca tta gtt ccg atg ttg aaa ctg ggt gat tca gcc agc cgc cga Arg Thr Leu Val Pro Met Leu Lys Leu Gly Asp Ser Ala Ser Arg Arg 245 250 255	768

ES 2 806 732 T3

tat acc att gtc tca gat ttc gat ttg tct tgc act gtc ttg gat tct Tyr Thr Ile Val Ser Asp Phe Asp Leu Ser Cys Thr Val Leu Asp Ser 260 265 270	816
tca gca gta tta gca gaa att gca ata ttg act act ctc aaa act gag Ser Ala Val Leu Ala Glu Ile Ala Ile Leu Thr Thr Leu Lys Thr Glu 275 280 285	864
caa aat ggt gct gaa aac tta agt gat cac aag tca tca tcg gag ttg Gln Asn Gly Ala Glu Asn Leu Ser Asp His Lys Ser Ser Ser Glu Leu 290 295 300	912
aga aaa act tgg gat gca ctt tct agt caa tat tct gaa gaa tgt gaa Arg Lys Thr Trp Asp Ala Leu Ser Ser Gln Tyr Ser Glu Glu Cys Glu 305 310 315 320	960
gaa tgc tta agg aag act ctg cca cct gaa gaa gtg ggc tct ttt gat Glu Cys Leu Arg Lys Thr Leu Pro Pro Glu Glu Val Gly Ser Phe Asp 325 330 335	1008
tat gaa ggc cta cac caa tct ctt gag cat ctg tct cag ttt gaa atg Tyr Glu Gly Leu His Gln Ser Leu Glu His Leu Ser Gln Phe Glu Met 340 345 350	1056
gag gca aac tct aaa gtt gtc gag tca ggt gtc ctt gag ggc att aat Glu Ala Asn Ser Lys Val Val Glu Ser Gly Val Leu Glu Gly Ile Asn 355 360 365	1104
ata gat gac att aaa aag gca gga gag cgt ctt gca ttt cag gat gga Ile Asp Asp Ile Lys Lys Ala Gly Glu Arg Leu Ala Phe Gln Asp Gly 370 375 380	1152
tgc gca aac ttt ttt gaa caa atc cta acg aaa atg gac agc tta aat Cys Ala Asn Phe Phe Glu Gln Ile Leu Thr Lys Met Asp Ser Leu Asn 385 390 395 400	1200
gtg gat gtg cac ata att tct gtt tgt tgg agt gga gat atc atc agg Val Asp Val His Ile Ile Ser Val Cys Trp Ser Gly Asp Ile Ile Arg 405 410 415	1248
gct gct ttt tca tca agc ggt ttg gat ggt tta cag gtt cat tca aat Ala Ala Phe Ser Ser Ser Gly Leu Asp Gly Leu Gln Val His Ser Asn 420 425 430	1296
gaa ctc acc ttt gtg gaa tca gtc tct act ggt ggt att gat agg cgt Glu Leu Thr Phe Val Glu Ser Val Ser Thr Gly Gly Ile Asp Arg Arg 435 440 445	1344
gtt gag tcc cca gtt gac aag ttg aaa atc ttc aat aat att tgg agt Val Glu Ser Pro Val Asp Lys Leu Lys Ile Phe Asn Asn Ile Trp Ser 450 455 460	1392
tct tca aag gac cag gac acg gaa cat atc tct ata tac att ggg gac Ser Ser Lys Asp Gln Asp Thr Glu His Ile Ser Ile Tyr Ile Gly Asp 465 470 475 480	1440
ggt tta ggt gac ttg ctt tgt ctt ctt cag gca gat att gga ata gtg Gly Leu Gly Asp Leu Leu Cys Leu Leu Gln Ala Asp Ile Gly Ile Val 485 490 495	1488
att ggt aca agc tca acg cta aga agg gtt gga aaa cgt ttt gga gta Ile Gly Thr Ser Ser Thr Leu Arg Arg Val Gly Lys Arg Phe Gly Val 1536	

ES 2 806 732 T3

	500	505	510	
	tcc ttt gtt cct ttg ttt tct ggt ctt ctc aaa cag gag aga gca tat			1584
	Ser Phe Val Pro Leu Phe Ser Gly Leu Leu Lys Gln Glu Arg Ala Tyr			
	515	520	525	
	gta gaa ggt tct agt tgt tgg aca aaa caa agt ggt att ctt tat acc			1632
	Val Glu Gly Ser Ser Cys Trp Thr Lys Gln Ser Gly Ile Leu Tyr Thr			
	530	535	540	
	gtc tct agt tgg agt gaa ata cat gct ttt att ttg ggc tct tcc aat			1680
	Val Ser Ser Trp Ser Glu Ile His Ala Phe Ile Leu Gly Ser Ser Asn			
	545	550	555	560
	tga			1683
	<210> 26			
	<211> 560			
5	<212> PRT			
	<213> <i>Picea sitchensis</i>			
	<400> 26			
	Met Gly Val Ala Asp Glu Ala Gly Val Ala Arg Arg Leu Trp Thr Lys			
	1	5	10	15
	Phe Lys Lys Asp Thr Ala Leu Ala Gln Tyr Asn Ser Phe Val Val Ala			
	20	25	30	
	Leu Ala Ala Gly Thr Leu Asn Met Thr Ser Phe Gln Gln Tyr Met Ala			
	35	40	45	
	Gln Asp Ala Tyr Phe Leu Lys Ala Phe Ala Gln Ala Tyr Thr Met Ala			
	50	55	60	
	Glu Asp Cys Ala Asp Asp Asp Asp Lys Ala Ser Ile Arg Glu Leu			
	65	70	75	80
	Arg Lys Ala Ala Glu Glu Glu Leu Asn Leu His Asn Ser Leu Ala Glu			
	85	90	95	
	Asp Trp Asp Val Glu Phe Ala Lys Glu Cys Ser Pro Asn Met Ala Thr			
	100	105	110	
	Val Lys Tyr Thr Glu Phe Leu Leu Ala Thr Ala Ala Gly Lys Val Glu			
	115	120	125	
	Gly Gly Lys Gly Pro Ser Arg Ser Val Thr Pro Phe Glu Lys Thr Lys			
	130	135	140	
10	Ile Ala Ala Tyr Thr Val Gly Ala Met Thr Pro Cys Met Arg Leu Tyr			
	145	150	155	160

ES 2 806 732 T3

Ala Phe Leu Gly Gln Glu Ile Val Lys Ala Leu Glu Pro Asp Cys Ser
165 170 175

Asn His Pro Tyr Lys Gln Trp Ile Glu Thr Tyr Ser Ser Ala Lys Phe
180 185 190

Glu Ala Ser Ala Leu Gln Thr Glu Glu Leu Leu Asp Lys Leu Ala Ile
195 200 205

Ser Leu Thr Gly Glu Glu Leu Glu Val Leu Arg Arg Leu Tyr Tyr His
210 215 220

Ala Leu Lys Leu Glu Ile Glu Phe Phe Ser Ala Gln Pro Phe Ser Gln
225 230 235 240

Arg Thr Leu Val Pro Met Leu Lys Leu Gly Asp Ser Ala Ser Arg Arg
245 250 255

Tyr Thr Ile Val Ser Asp Phe Asp Leu Ser Cys Thr Val Leu Asp Ser
260 265 270

Ser Ala Val Leu Ala Glu Ile Ala Ile Leu Thr Thr Leu Lys Thr Glu
275 280 285

Gln Asn Gly Ala Glu Asn Leu Ser Asp His Lys Ser Ser Ser Glu Leu
290 295 300

Arg Lys Thr Trp Asp Ala Leu Ser Ser Gln Tyr Ser Glu Glu Cys Glu
305 310 315 320

Glu Cys Leu Arg Lys Thr Leu Pro Pro Glu Glu Val Gly Ser Phe Asp
325 330 335

Tyr Glu Gly Leu His Gln Ser Leu Glu His Leu Ser Gln Phe Glu Met
340 345 350

Glu Ala Asn Ser Lys Val Val Glu Ser Gly Val Leu Glu Gly Ile Asn
355 360 365

Ile Asp Asp Ile Lys Lys Ala Gly Glu Arg Leu Ala Phe Gln Asp Gly
370 375 380

Cys Ala Asn Phe Phe Glu Gln Ile Leu Thr Lys Met Asp Ser Leu Asn
385 390 395 400

Val Asp Val His Ile Ile Ser Val Cys Trp Ser Gly Asp Ile Ile Arg
405 410 415

ES 2 806 732 T3

Ala Ala Phe Ser Ser Ser Gly Leu Asp Gly Leu Gln Val His Ser Asn
420 425 430

Glu Leu Thr Phe Val Glu Ser Val Ser Thr Gly Gly Ile Asp Arg Arg
435 440 445

Val Glu Ser Pro Val Asp Lys Leu Lys Ile Phe Asn Asn Ile Trp Ser
450 455 460

Ser Ser Lys Asp Gln Asp Thr Glu His Ile Ser Ile Tyr Ile Gly Asp
465 470 475 480

Gly Leu Gly Asp Leu Leu Cys Leu Leu Gln Ala Asp Ile Gly Ile Val
485 490 495

Ile Gly Thr Ser Ser Thr Leu Arg Arg Val Gly Lys Arg Phe Gly Val
500 505 510

Ser Phe Val Pro Leu Phe Ser Gly Leu Leu Lys Gln Glu Arg Ala Tyr
515 520 525

Val Glu Gly Ser Ser Cys Trp Thr Lys Gln Ser Gly Ile Leu Tyr Thr
530 535 540

Val Ser Ser Trp Ser Glu Ile His Ala Phe Ile Leu Gly Ser Ser Asn
545 550 555 560

<210> 27

<211> 1641

5 <212> ADN

<213> *Physcomitrella patens*

<220>

<221> CDS

10 <222> (1)..(1641)

<223> gen de *Physcomitrella patens* que codifica para TMP fosfatasa [XP_001769831]

<400> 27

atg aat ttg agc acg caa gct aac aca ggg ttg gcg aag agc ttc tgg 48
Met Asn Leu Ser Thr Gln Ala Asn Thr Gly Leu Ala Lys Ser Phe Trp
1 5 10 15

gct agt tgt aag aga gag gct tat gca tca ctc tac cat ccg ttt gtg 96
Ala Ser Cys Lys Arg Glu Ala Tyr Ala Ser Leu Tyr His Pro Phe Val
20 25 30

gtt gcg tta gcg gct gcc acc ttg cca aaa caa act ttt caa cgt tac 144
Val Ala Leu Ala Ala Gly Thr Leu Pro Lys Gln Thr Phe Gln Arg Tyr
35 40 45

15

ES 2 806 732 T3

atg gca cag gat gcc tat ttc ttg gag gcg ttc aag aat gcg tat caa Met Ala Gln Asp Ala Tyr Phe Leu Glu Ala Phe Lys Asn Ala Tyr Gln 50 55 60	192
ctg gct atg gaa acc act aca gac gaa gag gca aag gcc atc att gag Leu Ala Met Glu Thr Thr Thr Asp Glu Glu Ala Lys Ala Ile Ile Glu 65 70 75 80	240
tcc ctt cag aga gat gtg cag gaa gag ctc aat ttg cac tcg tcg atc Ser Leu Gln Arg Asp Val Gln Glu Glu Leu Asn Leu His Ser Ser Ile 85 90 95	288
atg cag tct ttg gat gct acc gat cag aat tgc ttt gaa cca aac atg Met Gln Ser Leu Asp Ala Thr Asp Gln Asn Cys Phe Glu Pro Asn Met 100 105 110	336
gca aca aca gcg tat tgt gat ttt ctg cta gcc aca gct aca gga agt Ala Thr Thr Ala Tyr Cys Asp Phe Leu Leu Ala Thr Ala Thr Gly Ser 115 120 125	384
aac gaa gca caa aaa ttt gga agc aca agt gct caa atc ata acc gct Asn Glu Ala Gln Lys Phe Gly Ser Thr Ser Ala Gln Ile Ile Thr Ala 130 135 140	432
atg act cct tgc atg cgg cta tat gca ttt ttg ggg cag gag ctc aaa Met Thr Pro Cys Met Arg Leu Tyr Ala Phe Leu Gly Gln Glu Leu Lys 145 150 155 160	480
aaa cac gtt gat cat gtt gct gac cat cct tac cag gag tgg att gat Lys His Val Asp His Val Ala Asp His Pro Tyr Gln Glu Trp Ile Asp 165 170 175	528
act tac tct gct gca gag ttc gag gct gca gct tcg aag att gag cag Thr Tyr Ser Ala Ala Glu Phe Glu Ala Ala Ala Ser Lys Ile Glu Gln 180 185 190	576
ctg cta gac aag tta act gct act ttg act gga aag cat gaa ata gca Leu Leu Asp Lys Leu Thr Ala Thr Leu Thr Gly Lys His Glu Ile Ala 195 200 205	624
ttc tta gaa agt ctc tat ctt caa gcc atg aac ttg gag gtg gat ttc Phe Leu Glu Ser Leu Tyr Leu Gln Ala Met Asn Leu Glu Val Asp Phe 210 215 220	672
ttc ggt gct cag ctg tta ggg cct gtg ctc gta ccc ttc ctc aaa tgc Phe Gly Ala Gln Leu Leu Gly Pro Val Leu Val Pro Phe Leu Lys Cys 225 230 235 240	720
caa ccg gct cca gag agc tat ata tta ctt gcg tct gac ttt gat tcc Gln Pro Ala Pro Glu Ser Tyr Ile Leu Leu Ala Ser Asp Phe Asp Ser 245 250 255	768
acg tgc acg ata tct gat tca tgc ccc ata ttg gca gac ctg acc gtg Thr Cys Thr Ile Ser Asp Ser Cys Pro Ile Leu Ala Asp Leu Thr Val 260 265 270	816
caa act gcg cga aaa tct cac ggt ggt cgt tca gtt ggt gaa tca ggg Gln Thr Ala Arg Lys Ser His Gly Gly Arg Ser Val Gly Glu Ser Gly 275 280 285	864
gcc agc ttg ttg aaa aaa aga tgg gat gat ctc gtc atg cag tat atg Ala Ser Leu Leu Lys Lys Arg Trp Asp Asp Leu Val Met Gln Tyr Met 290 295 300	912

ES 2 806 732 T3

gac gag tat gag gac gtt ctg aag cga agc ctg gtg aaa aaa gat aat Asp Glu Tyr Glu Asp Val Leu Lys Arg Ser Leu Val Lys Lys Asp Asn 305 310 315 320	960
ggc agt gtt aat gcg ctc agt gca gag aat ctc caa gag ttt ctg aag Gly Ser Val Asn Ala Leu Ser Ala Glu Asn Leu Gln Glu Phe Leu Lys 325 330 335	1008
gaa atg tcc aac ttc gaa cag aag gcc aat gcg agg gtc gaa gag gct Glu Met Ser Asn Phe Glu Gln Lys Ala Asn Ala Arg Val Glu Glu Ala 340 345 350	1056
gca gtt cta aaa gcg tta tct ctg gct tcg att caa gaa gct gga aaa Ala Val Leu Lys Gly Leu Ser Leu Ala Ser Ile Gln Glu Ala Gly Lys 355 360 365	1104
tcc atg cct ctt cgt gag gcc tgt tct gac ttt ttt aag cgt ctg gaa Ser Met Pro Leu Arg Glu Gly Cys Ser Asp Phe Phe Lys Arg Leu Glu 370 375 380	1152
tca gga gag gtt ctt gtt gac aca tgt ata ttg tct gtg tgc tgg agc Ser Gly Glu Val Leu Val Asp Thr Cys Ile Leu Ser Val Cys Trp Ser 385 390 395 400	1200
aaa acc ttc atc gaa gct gtc ttg gaa aag gtt cgt att cca aac atc Lys Thr Phe Ile Glu Ala Val Leu Glu Lys Val Arg Ile Pro Asn Ile 405 410 415	1248
aat gcc aac gag ctc gtt ttc gaa gga cgc att tcc acc ggt gct att Asn Ala Asn Glu Leu Val Phe Glu Gly Arg Ile Ser Thr Gly Ala Ile 420 425 430	1296
atc aaa aac gtc gaa acg gct ctt gac aag caa aga cac ttc gtt cag Ile Lys Asn Val Glu Thr Ala Leu Asp Lys Gln Arg His Phe Val Gln 435 440 445	1344
ttg ctg gat aat cta aaa cca act caa gac gtg ctg tcc att tat gtt Leu Leu Asp Asn Leu Lys Pro Thr Gln Asp Val Leu Ser Ile Tyr Val 450 455 460	1392
ggt gat agt ctg act gat ctt ctc tgc cta atc aga gca gac ctg ggt Gly Asp Ser Leu Thr Asp Leu Leu Cys Leu Ile Arg Ala Asp Leu Gly 465 470 475 480	1440
ata gtt ctc ggt gac agc agc gct ctg aag cag gtg tat ggg cca aaa Ile Val Leu Gly Asp Ser Ser Ala Leu Lys Gln Val Tyr Gly Pro Lys 485 490 495	1488
atg gcc ccc ctc ttc atg aaa gcc ata ctc ttg gag cag gca aac atg Met Ala Pro Leu Phe Met Lys Ala Ile Leu Leu Glu Gln Ala Asn Met 500 505 510	1536
cga ggc agg cag caa ccc aca ggt tac gtc ttc act gtc tcc agt tgg Arg Gly Arg Gln Gln Pro Thr Gly Tyr Val Phe Thr Val Ser Ser Trp 515 520 525	1584
tat gag gtg gaa gcc ttt ctg ttg ggt cct gct aga aac aga cct ttg Tyr Glu Val Glu Ala Phe Leu Leu Gly Pro Ala Arg Asn Arg Pro Leu 530 535 540	1632
tac atc tag Tyr Ile 545	1641

<210> 28
 5 <211> 546
 <212> PRT
 <213> *Physcomitrella patens*

<400> 28

ES 2 806 732 T3

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

ES 2 806 732 T3

210					215					220					
Phe 225	Gly	Ala	Gln	Leu	Leu 230	Gly	Pro	Val	Leu	Val 235	Pro	Phe	Leu	Lys	Cys 240
Gln	Pro	Ala	Pro	Glu 245	Ser	Tyr	Ile	Leu	Leu 250	Ala	Ser	Asp	Phe	Asp 255	Ser
Thr	Cys	Thr	Ile 260	Ser	Asp	Ser	Cys	Pro 265	Ile	Leu	Ala	Asp	Leu 270	Thr	Val
Gln	Thr	Ala 275	Arg	Lys	Ser	His	Gly 280	Gly	Arg	Ser	Val	Gly 285	Glu	Ser	Gly
Ala	Ser 290	Leu	Leu	Lys	Lys	Arg 295	Trp	Asp	Asp	Leu	Val 300	Met	Gln	Tyr	Met
Asp 305	Glu	Tyr	Glu	Asp	Val 310	Leu	Lys	Arg	Ser	Leu 315	Val	Lys	Lys	Asp	Asn 320
Gly	Ser	Val	Asn 325	Ala	Leu	Ser	Ala	Glu	Asn 330	Leu	Gln	Glu	Phe	Leu 335	Lys
Glu	Met	Ser	Asn 340	Phe	Glu	Gln	Lys	Ala 345	Asn	Ala	Arg	Val	Glu 350	Glu	Ala
Ala	Val	Leu 355	Lys	Gly	Leu	Ser	Leu 360	Ala	Ser	Ile	Gln	Glu 365	Ala	Gly	Lys
Ser	Met 370	Pro	Leu	Arg	Glu	Gly 375	Cys	Ser	Asp	Phe	Phe 380	Lys	Arg	Leu	Glu
Ser 385	Gly	Glu	Val	Leu	Val 390	Asp	Thr	Cys	Ile	Leu 395	Ser	Val	Cys	Trp	Ser 400
Lys	Thr	Phe	Ile	Glu 405	Ala	Val	Leu	Glu	Lys 410	Val	Arg	Ile	Pro	Asn 415	Ile
Asn	Ala	Asn	Glu 420	Leu	Val	Phe	Glu	Gly 425	Arg	Ile	Ser	Thr	Gly 430	Ala	Ile
Ile	Lys 435	Asn	Val	Glu	Thr	Ala	Leu 440	Asp	Lys	Gln	Arg	His 445	Phe	Val	Gln
Leu	Leu 450	Asp	Asn	Leu	Lys	Pro 455	Thr	Gln	Asp	Val	Leu 460	Ser	Ile	Tyr	Val

ES 2 806 732 T3

Gly Asp Ser Leu Thr Asp Leu Leu Cys Leu Ile Arg Ala Asp Leu Gly
465 470 475 480

Ile Val Leu Gly Asp Ser Ser Ala Leu Lys Gln Val Tyr Gly Pro Lys
485 490 495

Met Ala Pro Leu Phe Met Lys Ala Ile Leu Leu Glu Gln Ala Asn Met
500 505 510

Arg Gly Arg Gln Gln Pro Thr Gly Tyr Val Phe Thr Val Ser Ser Trp
515 520 525

Tyr Glu Val Glu Ala Phe Leu Leu Gly Pro Ala Arg Asn Arg Pro Leu
530 535 540

Tyr Ile
545

<210> 29

<211> 1593

5 <212> ADN

<213> *Selaginella moellendorffii*

<220>

<221> CDS

10 <222> (1)..(1593)

<223> gen de *Selaginella moellendorffii* que codifica para TMP fosfatasa [XP_002990363]

<400> 29

atg tcg tgt ttg ctt aga aat gta gtg gcc aga gga ttg agg agc ttg 48
Met Ser Cys Leu Leu Arg Asn Val Val Ala Arg Gly Leu Arg Ser Leu
1 5 10 15

gca agc gcc cag gcg atg gag cca tcc att tca aag cgc ttg tgg cag 96
Ala Ser Ala Gln Ala Met Glu Pro Ser Ile Ser Lys Arg Leu Trp Gln
20 25 30

caa tcc aag cgc gag gca atg gta tgt ctg tat cat cca ttt gtg gtg 144
Gln Ser Lys Arg Glu Ala Met Val Cys Leu Tyr His Pro Phe Val Val
35 40 45

tcc atc gct gct ggg acg ctg gat ctt cac agc ttc cag cga ttc ata 192
Ser Ile Ala Ala Gly Thr Leu Asp Leu His Ser Phe Gln Arg Phe Ile
50 55 60

gcg cag gat tcc ttc ttc ctg acg gca ttc gcg aaa gcc tat ggt ttg 240
Ala Gln Asp Ser Phe Phe Leu Thr Ala Phe Ala Lys Ala Tyr Gly Leu
65 70 75 80

gcc ata gag cgc agc gat gat cga gaa gtt aaa tct gag att tgc aag 288
Ala Ile Glu Arg Ser Asp Asp Arg Glu Val Lys Ser Glu Ile Cys Lys
85 90 95

ctc caa cag gct gtg tac gag gaa ctt gag ctc cat tct tcc ctc atg 336
15 Leu Gln Gln Ala Val Tyr Glu Glu Leu Glu Leu His Ser Ser Leu Met

ES 2 806 732 T3

100	105	110	
aag gct tgg aac ttc gat cat aca cca cca tcg cca gca act tgt gct Lys Ala Trp Asn Phe Asp His Thr Pro Pro Ser Pro Ala Thr Cys Ala 115 120 125			384
tac aca gat ttt ctc ctc gca gtg gct gct ggg aag aaa att gaa tgc Tyr Thr Asp Phe Leu Leu Ala Val Ala Ala Gly Lys Lys Ile Glu Cys 130 135 140			432
gag aaa act aag gtg ccg atg ctc gct ctg gca gca atg gct ccg tgc Glu Lys Thr Lys Val Met Leu Ala Leu Ala Met Ala Pro Cys 145 150 155 160			480
atg cgt ctc tac gct ttc cta ggc caa gag acg aga gtt ttc tct cga Met Arg Leu Tyr Phe Leu Gly Gln Thr Arg Val Phe Ser Arg 165 170 175			528
gaa aat cat cca tat cgc gac tgg att tcg act tac tcg tcg cct ggt Glu Asn His Pro Tyr Arg Asp Trp Ile Ser Thr Tyr Ser Ser Pro Gly 180 185 190			576
ttc gag act gct gct act cga ctc gag cag ctt ctc gat agc ctc tcg Phe Glu Thr Ala Ala Thr Arg Leu Glu Gln Leu Leu Asp Ser Leu Ser 195 200 205			624
gaa gct caa gag act acg gca gcg gaa ttt cag agt atg caa agt ttg Glu Ala Gln Glu Thr Thr Ala Ala Glu Phe Gln Ser Met Gln Ser Leu 210 215 220			672
tat cac cgt gcc ata gcg tac gag gtg agc ttc ttc gat gcc cag gaa Tyr His Arg Ala Ile Ala Tyr Glu Val Ser Phe Phe Asp Ala Gln Glu 225 230 235 240			720
gtg cgt ggc agc aac gct ttt gtc ccg ctg cta gag agt gta gca ctc Val Arg Gly Ser Asn Ala Phe Val Pro Leu Leu Glu Ser Val Ala Leu 245 250 255			768
aag gat cgc aac ttc gtc ctc atc tct gat ttt gat tct act tgc acc Lys Asp Arg Asn Phe Val Leu Ile Ser Asp Phe Asp Ser Thr Cys Thr 260 265 270			816
gtc tct gat tca tcc cca gtt cta gcg gag ctg gct atg gcg gtc gat Val Ser Asp Ser Ser Pro Val Leu Ala Glu Leu Ala Met Ala Val Asp 275 280 285			864
cca aat gta agg agg aaa tgg agc agc ctc tcg gac gag tat ttc agg Pro Asn Val Arg Arg Lys Trp Ser Ser Leu Ser Asp Glu Tyr Phe Arg 290 295 300			912
gac tac tcc aaa ctc ctg gaa gaa gtt gtt ctt cgt gag tac gac tac Asp Tyr Ser Lys Leu Leu Glu Glu Val Val Leu Arg Glu Tyr Asp Tyr 305 310 315 320			960
gat gcg atc aaa gag gct ctc caa gtt ctt tcc gag ttt gag aag caa Asp Ala Ile Lys Glu Ala Leu Gln Val Leu Ser Glu Phe Glu Lys Gln 325 330 335			1008
ggg aac gcg aaa atc gac gcc tcc cgc gtt ttg caa ggc att aag atc Gly Asn Ala Lys Ile Asp Ala Ser Arg Val Leu Gln Gly Ile Lys Ile 340 345 350			1056
gat gat atc aag caa gcc gga caa aac atg gca ctt caa gct ggc tgt			1104

ES 2 806 732 T3

Asp	Asp	Ile	Lys	Gln	Ala	Gly	Gln	Asn	Met	Ala	Leu	Gln	Ala	Gly	Cys		
		355					360					365					
gcc	agt	gtg	ctt	tgc	agg	cta	agt	tcc	aaa	atc	tct	tgt	caa	atc	ctc	1152	
Ala	Ser	Val	Leu	Cys	Arg	Leu	Ser	Ser	Lys	Ile	Ser	Cys	Gln	Ile	Leu		
		370				375					380						
tcg	gtt	tgc	tgg	agc	cgg	acc	ttc	atc	gaa	gca	gct	ttc	tcc	aaa	gag	1200	
Ser	Val	Cys	Trp	Ser	Arg	Thr	Phe	Ile	Glu	Ala	Ala	Phe	Ser	Lys	Glu		
					390				395						400		
aat	atc	acc	aat	gtt	cct	gtc	cat	tcc	aac	gaa	ctc	gaa	aac	gat	ggg	1248	
Asn	Ile	Thr	Asn	Val	Pro	Val	His	Ser	Asn	Glu	Leu	Glu	Asn	Asp	Gly		
				405					410					415			
aac	ttt	aca	acc	ggg	agc	ttg	atc	aga	cgc	gtc	gag	aca	ccg	att	gac	1296	
Asn	Phe	Thr	Thr	Gly	Ser	Leu	Ile	Arg	Arg	Val	Glu	Thr	Pro	Ile	Asp		
			420					425					430				
aag	gaa	gag	acg	atg	ttt	cgt	gag	att	cta	cac	gct	ccg	gac	gac	aag	1344	
Lys	Glu	Glu	Thr	Met	Phe	Arg	Glu	Ile	Leu	His	Ala	Pro	Asp	Asp	Lys		
		435					440					445					
ttt	gtg	att	ttc	att	gga	gac	agc	ctc	acg	gat	ctg	cta	gcc	ttg	ctc	1392	
Phe	Val	Ile	Phe	Ile	Gly	Asp	Ser	Leu	Thr	Asp	Leu	Leu	Ala	Leu	Leu		
		450				455					460						
cga	gct	gac	att	gga	att	gtt	cta	gga	acg	agc	tcc	agc	ctc	gat	cga	1440	
Arg	Ala	Asp	Ile	Gly	Ile	Val	Leu	Gly	Thr	Ser	Ser	Ser	Leu	Asp	Arg		
					470				475						480		
gcc	tcc	aaa	gcc	ttt	gga	gtg	aag	atc	gtg	cca	ctc	ttt	tcc	ggc	ctc	1488	
Ala	Ser	Lys	Ala	Phe	Gly	Val	Lys	Ile	Val	Pro	Leu	Phe	Ser	Gly	Leu		
				485					490					495			
gtc	cag	cgg	cag	caa	agc	tct	cga	tca	gcg	tgg	aga	aaa	gag	gaa	gga	1536	
Val	Gln	Arg	Gln	Gln	Ser	Ser	Arg	Ser	Ala	Trp	Arg	Lys	Glu	Glu	Gly		
			500					505					510				
gtt	ttg	tat	cga	gct	tct	gga	tgg	ctg	gag	ata	gaa	gcg	ttt	cta	gct	1584	
Val	Leu	Tyr	Arg	Ala	Ser	Gly	Trp	Leu	Glu	Ile	Glu	Ala	Phe	Leu	Ala		
		515				520						525					
ggt	aat	tag														1593	
Gly	Asn																
		530															

<210> 30

<211> 530

5 <212> PRT

<213> *Selaginella moellendorffii*

<400> 30

Met	Ser	Cys	Leu	Leu	Arg	Asn	Val	Val	Ala	Arg	Gly	Leu	Arg	Ser	Leu
1				5					10					15	

Ala	Ser	Ala	Gln	Ala	Met	Glu	Pro	Ser	Ile	Ser	Lys	Arg	Leu	Trp	Gln
			20					25					30		

10

ES 2 806 732 T3

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

ES 2 806 732 T3

Pro Asn Val Arg Arg Lys Trp Ser Ser Leu Ser Asp Glu Tyr Phe Arg
290 295 300

Asp Tyr Ser Lys Leu Leu Glu Glu Val Val Leu Arg Glu Tyr Asp Tyr
305 310 315 320

Asp Ala Ile Lys Glu Ala Leu Gln Val Leu Ser Glu Phe Glu Lys Gln
325 330 335

Gly Asn Ala Lys Ile Asp Ala Ser Arg Val Leu Gln Gly Ile Lys Ile
340 345 350

Asp Asp Ile Lys Gln Ala Gly Gln Asn Met Ala Leu Gln Ala Gly Cys
355 360 365

Ala Ser Val Leu Cys Arg Leu Ser Ser Lys Ile Ser Cys Gln Ile Leu
370 375 380

Ser Val Cys Trp Ser Arg Thr Phe Ile Glu Ala Ala Phe Ser Lys Glu
385 390 395 400

Asn Ile Thr Asn Val Pro Val His Ser Asn Glu Leu Glu Asn Asp Gly
405 410 415

Asn Phe Thr Thr Gly Ser Leu Ile Arg Arg Val Glu Thr Pro Ile Asp
420 425 430

Lys Glu Glu Thr Met Phe Arg Glu Ile Leu His Ala Pro Asp Asp Lys
435 440 445

Phe Val Ile Phe Ile Gly Asp Ser Leu Thr Asp Leu Leu Ala Leu Leu
450 455 460

Arg Ala Asp Ile Gly Ile Val Leu Gly Thr Ser Ser Ser Leu Asp Arg
465 470 475 480

Ala Ser Lys Ala Phe Gly Val Lys Ile Val Pro Leu Phe Ser Gly Leu
485 490 495

Val Gln Arg Gln Gln Ser Ser Arg Ser Ala Trp Arg Lys Glu Glu Gly
500 505 510

Val Leu Tyr Arg Ala Ser Gly Trp Leu Glu Ile Glu Ala Phe Leu Ala
515 520 525

Gly Asn
530

<210> 31

<211> 648

5 <212> ADN

<213> *Anaerotruncus colihominis*

<220>

<221> CDS

10 <222> (1)..(648)

<223> gen de *Anaerotruncus colihominis* que codifica para TMP fosfatasa [WP_006874980]

ES 2 806 732 T3

<400> 31

atg atc aaa ggc gcg att ttt gat atg gac ggt acg ctg att gat tcc	48
Met Ile Lys Gly Ala Ile Phe Asp Met Asp Gly Thr Leu Ile Asp Ser	
1 5 10 15	
atg cct cta tgg gag gac tgc gga cgg gcc ttt tta tcc gcg cgc ggc	96
Met Pro Leu Trp Glu Asp Cys Gly Arg Ala Phe Leu Ser Ala Arg Gly	
20 25 30	
att act gcg cgt gac gat ctg ggc gaa acg ctc aaa tcc ctg tcg atg	144
Ile Thr Ala Arg Asp Asp Leu Gly Glu Thr Leu Lys Ser Leu Ser Met	
35 40 45	
gag caa acg gct aat tat ttg cgg gac gca tac ggt att tcc gag aca	192
Glu Gln Thr Ala Asn Tyr Leu Arg Asp Ala Tyr Gly Ile Ser Glu Thr	
50 55 60	
acc tct gaa atc att gag atg atc aat gga atg gtt act gac gca tat	240
Thr Ser Glu Ile Ile Glu Met Ile Asn Gly Met Val Thr Asp Ala Tyr	
65 70 75 80	
cag cgc acc atc ccg ctt aaa cgt gac att gcc gcg ttt ctc gag cgc	288
Gln Arg Thr Ile Pro Leu Lys Arg Asp Ile Ala Ala Phe Leu Glu Arg	
85 90 95	
ctc agg cag gcg gat gtg cgc atg tgt gtc gca acg gca acg gac cgt	336
Leu Arg Gln Ala Asp Val Arg Met Cys Val Ala Thr Ala Thr Asp Arg	
100 105 110	
cca ctg gtg gag gcg gcg ctt gga cgc ctt gac ctc ctg ccc ttt ttt	384
Pro Leu Val Glu Ala Ala Leu Gly Arg Leu Asp Leu Leu Pro Phe Phe	
115 120 125	
gaa cgg att ttc acc tgt tcg gag gtg ggg gcc ggc aag gac cgc ccc	432
Glu Arg Ile Phe Thr Cys Ser Glu Val Gly Ala Gly Lys Asp Arg Pro	
130 135 140	
gat atc ttt gag cag gcg tgc gcc gcg ctt ggc acg ccg cgc ggc gaa	480
Asp Ile Phe Glu Gln Ala Cys Ala Ala Leu Gly Thr Pro Arg Gly Glu	
145 150 155 160	
acc gtc atc ttt gag gat gct ctt tat gcg att gaa aca gct cgg cgc	528
Thr Val Ile Phe Glu Asp Ala Leu Tyr Ala Ile Glu Thr Ala Arg Arg	
165 170 175	
gcc ggg ttc cgc gtt gtc gca atc gcg gac gac gcc tcc gcc ggc gac	576
Ala Gly Phe Arg Val Val Ala Ile Ala Asp Asp Ala Ser Ala Gly Asp	
180 185 190	
gag gcg cgc ata gcc gca ctg tct gag caa tat ata cat aac tat gag	624
Glu Ala Arg Ile Ala Ala Leu Ser Glu Gln Tyr Ile His Asn Tyr Glu	
195 200 205	
gaa tgc gag gta aac agt tta tga	648
Glu Cys Glu Val Asn Ser Leu	
210 215	

5

<210> 32

<211> 215

<212> PRT

<213> *Anaerotruncus colihominis*

10

<400> 32

ES 2 806 732 T3

Met Ile Lys Gly Ala Ile Phe Asp Met Asp Gly Thr Leu Ile Asp Ser
1 5 10 15

Met Pro Leu Trp Glu Asp Cys Gly Arg Ala Phe Leu Ser Ala Arg Gly
20 25 30

Ile Thr Ala Arg Asp Asp Leu Gly Glu Thr Leu Lys Ser Leu Ser Met
35 40 45

Glu Gln Thr Ala Asn Tyr Leu Arg Asp Ala Tyr Gly Ile Ser Glu Thr
50 55 60

Thr Ser Glu Ile Ile Glu Met Ile Asn Gly Met Val Thr Asp Ala Tyr
65 70 75 80

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

Leu Arg Gln Ala Asp Val Arg Met Cys Val Ala Thr Ala Thr Asp Arg
100 105 110

Pro Leu Val Glu Ala Ala Leu Gly Arg Leu Asp Leu Leu Pro Phe Phe
115 120 125

Glu Arg Ile Phe Thr Cys Ser Glu Val Gly Ala Gly Lys Asp Arg Pro
130 135 140

Asp Ile Phe Glu Gln Ala Cys Ala Ala Leu Gly Thr Pro Arg Gly Glu
145 150 155 160

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

Ala Gly Phe Arg Val Val Ala Ile Ala Asp Asp Ala Ser Ala Gly Asp
180 185 190
Glu Ala Arg Ile Ala Ala Leu Ser Glu Gln Tyr Ile His Asn Tyr Glu
195 200 205

Glu Cys Glu Val Asn Ser Leu
210 215

<210> 33
<211> 666
<212> ADN
<213> *Eubacterium ventriosum*

<220>
<221> CDS
<222> (1)..(666)
<223> gen de *Eubacterium ventriosum* que codifica para TMP fosfatasa [WP_005362972]

<400> 33

ES 2 806 732 T3

atg tca aca gga ttt ata ttt gat gta gat gga aca ata cta gac tca	48
Met Ser Thr Gly Phe Ile Phe Asp Val Asp Gly Thr Ile Leu Asp Ser	
1 5 10 15	
atg gga ata tgg atg aac gta gga gaa cta tat cta aaa gat atg gga	96
Met Gly Ile Trp Met Asn Val Gly Glu Leu Tyr Leu Lys Asp Met Gly	
20 25 30	
ata aag gcg gaa cca aat ctt gga gaa att cta ttc gaa atg aca atg	144
Ile Lys Ala Glu Pro Asn Leu Gly Glu Ile Leu Phe Glu Met Thr Met	
35 40 45	
aat gaa ggt gca gaa tac ata caa aaa aag tat aat cta aac ctt aca	192
Asn Glu Gly Ala Glu Tyr Ile Gln Lys Lys Tyr Asn Leu Asn Leu Thr	
50 55 60	
aca gaa gaa ata tgc acc gga ata aac aac cgt gta tac aaa ttc tac	240
Thr Glu Glu Ile Cys Thr Gly Ile Asn Asn Arg Val Tyr Lys Phe Tyr	
65 70 75 80	
gaa aaa gaa gca atg cca aaa cca aaa gtt atc gac ttt ata gaa caa	288
Glu Lys Glu Ala Met Pro Lys Pro Lys Val Ile Asp Phe Ile Glu Gln	
85 90 95	
gcc tac gag aac aaa atc cca atg aca ata gca acg tca aca gac aga	336
Ala Tyr Glu Asn Lys Ile Pro Met Thr Ile Ala Thr Ser Thr Asp Arg	
100 105 110	
cca atg ata gaa gca gct ttc aaa aga ctg cac ata gac aaa tat ttt	384
Pro Met Ile Glu Ala Ala Phe Lys Arg Leu His Ile Asp Lys Tyr Phe	
115 120 125	
aaa aaa ata ttt acc acg aca gag gtt ggg tat gga aaa gac aaa ccg	432
Lys Lys Ile Phe Thr Thr Thr Glu Val Gly Tyr Gly Lys Asp Lys Pro	
130 135 140	
gac atc ttc ata aaa gca atg gaa gaa atg gga aca aca cca aag caa	480
Asp Ile Phe Ile Lys Ala Met Glu Glu Met Gly Thr Thr Pro Lys Gln	
145 150 155 160	
aca tgg cta ttt gaa gat gga gca tac tca ata gaa aca gcc aaa caa	528
Thr Trp Leu Phe Glu Asp Gly Ala Tyr Ser Ile Glu Thr Ala Lys Gln	
165 170 175	
cta ggc ata aaa aca ata gga atc tac gat cct gca agc gaa aaa gac	576
Leu Gly Ile Lys Thr Ile Gly Ile Tyr Asp Pro Ala Ser Glu Lys Asp	
180 185 190	
cag gaa aaa ata aga aac cta aca aac atc tac ata aaa aat tgg aca	624
Gln Glu Lys Ile Arg Asn Leu Thr Asn Ile Tyr Ile Lys Asn Trp Thr	
195 200 205	
gaa cac aaa acc cta ctt aac caa ata caa aac aac aag tag	666
Glu His Lys Thr Leu Leu Asn Gln Ile Gln Asn Asn Lys	
210 215 220	

<210> 34

5 <211> 221

<212> PRT

<213> *Eubacterium ventriosum*

<400> 34

10

ES 2 806 732 T3

Met Ser Thr Gly Phe Ile Phe Asp Val Asp Gly Thr Ile Leu Asp Ser
1 5 10 15

Met Gly Ile Trp Met Asn Val Gly Glu Leu Tyr Leu Lys Asp Met Gly
20 25 30

Ile Lys Ala Glu Pro Asn Leu Gly Glu Ile Leu Phe Glu Met Thr Met
35 40 45

Asn Glu Gly Ala Glu Tyr Ile Gln Lys Lys Tyr Asn Leu Asn Leu Thr
50 55 60

Thr Glu Glu Ile Cys Thr Gly Ile Asn Asn Arg Val Tyr Lys Phe Tyr
65 70 75 80

Glu Lys Glu Ala Met Pro Lys Pro Lys Val Ile Asp Phe Ile Glu Gln
85 90 95

Ala Tyr Glu Asn Lys Ile Pro Met Thr Ile Ala Thr Ser Thr Asp Arg
100 105 110

Pro Met Ile Glu Ala Ala Phe Lys Arg Leu His Ile Asp Lys Tyr Phe
115 120 125

Lys Lys Ile Phe Thr Thr Thr Glu Val Gly Tyr Gly Lys Asp Lys Pro
130 135 140

Asp Ile Phe Ile Lys Ala Met Glu Glu Met Gly Thr Thr Pro Lys Gln
145 150 155 160
Thr Trp Leu Phe Glu Asp Gly Ala Tyr Ser Ile Glu Thr Ala Lys Gln
165 170 175

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

Gln Glu Lys Ile Arg Asn Leu Thr Asn Ile Tyr Ile Lys Asn Trp Thr
195 200 205

Glu His Lys Thr Leu Leu Asn Gln Ile Gln Asn Asn Lys
210 215 220

<210> 35

5 <211> 1482

<212> ADN

<213> *Coprococcus eutactus*

<220>

10 <221> CDS

<222> (1)..(1482)

<223> gen de *Coprococcus eutactus* ATCC 27759 que codifica para TMP fosfatasa [EDP27707]

<400> 35

15

ES 2 806 732 T3

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

ES 2 806 732 T3

130	135	140	
ggc gtt gga aag gct gcc tct cca aag ata tat gag ctt gcg gcc gaa Gly Val Gly Lys Ala Ala Ser Pro Lys Ile Tyr Glu Leu Ala Ala Glu 145 150 155 160			480
ttt atg ggg acg aaa gtc ggc gag tca ttt gtg ttc gag gat gcc tat Phe Met Gly Thr Lys Val Gly Glu Ser Phe Val Phe Glu Asp Ala Tyr 165 170 175			528
cat gcg gcc gag aca gct cag aat gcg gga ttt aca gtt gta gga ctc His Ala Ala Glu Thr Ala Gln Asn Ala Gly Phe Thr Val Val Gly Leu 180 185 190			576
tat gac gag tca agc cgt gac atg caa gca gaa ctt aag gtt cac tgc Tyr Asp Glu Ser Ser Arg Asp Met Gln Ala Glu Leu Lys Val His Cys 195 200 205			624
aat tat tac tat ttg gga ttt gcc gag ctt ata gat gag ctg ctg cct Asn Tyr Tyr Leu Gly Phe Ala Glu Leu Ile Asp Glu Leu Leu Pro 210 215 220			672
gac aga agc cag ctt gca ccg gtt ctt acc atc gcg ggc agt gat tca Asp Arg Ser Gln Leu Ala Pro Val Leu Thr Ile Ala Gly Ser Asp Ser 225 230 235 240			720
tcg gga ggt gcg gga ata cag gca gat ctt aag acc atg cag gca aat Ser Gly Gly Ala Gly Ile Gln Ala Asp Leu Lys Thr Met Gln Ala Asn 245 250 255			768
gga gtg ttt ggc atg agc gca gta act gcc ttg acg gcg cag aat acc Gly Val Phe Gly Met Ser Ala Val Thr Ala Leu Thr Ala Gln Asn Thr 260 265 270			816
aca ggt gtg aca tcc atc atg aat gtg aca cct gac ata ctt gca gat Thr Gly Val Thr Ser Ile Met Asn Val Thr Pro Asp Ile Leu Ala Asp 275 280 285			864
cag ata gat gca gta ttt aca gat ata aga cca cag gcg gtc aag ata Gln Ile Asp Ala Val Phe Thr Asp Ile Arg Pro Gln Ala Val Lys Ile 290 295 300			912
ggt atg gtg tct gtg cca gaa ctt ata aat gtg atc gca gac aag ctt Gly Met Val Ser Val Pro Glu Leu Ile Asn Val Ile Ala Asp Lys Leu 305 310 315 320			960
gaa ttt tac agg gcg gag aat gtg gtg ctt gat cct gtg atg gtt gcg Glu Phe Tyr Arg Ala Glu Asn Val Val Leu Asp Pro Val Met Val Ala 325 330 335			1008
aca agc ggt gct aaa ctc ata agc gat gat gct gtg gac gtt ttg aca Thr Ser Gly Ala Lys Leu Ile Ser Asp Asp Ala Val Asp Val Leu Thr 340 345 350			1056
gga agg ctg ttc cca ctt gca aag ctg atc acc cca aat att cca gag Gly Arg Leu Phe Pro Leu Ala Lys Leu Ile Thr Pro Asn Ile Pro Glu 355 360 365			1104
aca gag gcc ctc aca ggt atg agt atc cgg tct aag gaa gat atg gaa Thr Glu Ala Leu Thr Gly Met Ser Ile Arg Ser Lys Glu Asp Met Glu 370 375 380			1152
agt gca gca agg aaa ata tat gaa aaa tat ggc tgc tca gtt ctt gtg			1200

ES 2 806 732 T3

Ser	Ala	Ala	Arg	Lys	Ile	Tyr	Glu	Lys	Tyr	Gly	Cys	Ser	Val	Leu	Val	
385					390					395					400	
aag	ggc	gga	cat	agc	ata	aac	gat	gcg	aat	gat	atg	ctg	ttt	gat	gga	1248
Lys	Gly	Gly	His	Ser	Ile	Asn	Asp	Ala	Asn	Asp	Met	Leu	Phe	Asp	Gly	
			405					410					415			
gag	aat	gta	tca	tggt	ttt	tca	ggt	gag	aga	ata	gaa	aat	ccg	aat	acc	1296
Glu	Asn	Val	Ser	Trp	Phe	Ser	Gly	Glu	Arg	Ile	Glu	Asn	Pro	Asn	Thr	
			420					425					430			
cat	gga	acg	ggg	tgt	aca	ctc	tca	agt	gca	ata	gcc	tcc	aac	ctt	gca	1344
His	Gly	Thr	Gly	Cys	Thr	Leu	Ser	Ser	Ala	Ile	Ala	Ser	Asn	Leu	Ala	
		435					440					445				
aag	gga	tat	gat	ata	gaa	act	tct	gtg	cag	aga	gca	aaa	gcg	tac	atc	1392
Lys	Gly	Tyr	Asp	Ile	Glu	Thr	Ser	Val	Gln	Arg	Ala	Lys	Ala	Tyr	Ile	
	450					455					460					
tca	gga	gcc	ctg	gct	gcg	atg	ctt	gat	cta	gga	aga	gga	agc	ggc	ccg	1440
Ser	Gly	Ala	Leu	Ala	Ala	Met	Leu	Asp	Leu	Gly	Arg	Gly	Ser	Gly	Pro	
465					470					475					480	
tta	aac	cat	ggc	ttt	gat	ata	gac	agc	aga	ttc	atg	ata	taa			1482
Leu	Asn	His	Gly	Phe	Asp	Ile	Asp	Ser	Arg	Phe	Met	Ile				
			485						490							

<210> 36

<211> 493

5 <212> PRT

<213> *Coprococcus eutactus*

<400> 36

Met	Lys	Lys	Ile	Val	Ile	Ser	Asp	Ile	Lys	Gly	Ala	Ile	Phe	Asp	Met	
1				5					10					15		
Asp	Gly	Val	Leu	Leu	Asp	Ser	Met	Pro	Met	Trp	Asp	His	Ala	Gly	Glu	
		20						25					30			
Met	Tyr	Leu	Ala	Gly	Gln	Gly	Ile	Glu	Ala	Glu	Pro	Asp	Leu	Glu	Lys	
		35					40					45				
Val	Leu	Phe	Thr	Met	Thr	Met	Gln	Lys	Gly	Ala	Glu	Tyr	Ile	Arg	Asp	
	50					55					60					
His	Tyr	Gly	Leu	Lys	Leu	Thr	Ala	Asp	Glu	Ile	Ile	Asp	Gly	Ile	Asn	
65					70					75					80	
Glu	Thr	Val	Arg	Asp	Phe	Tyr	Ala	Asn	Lys	Val	Val	Pro	Lys	Asn	Gly	
			85						90					95		
Val	Leu	Lys	Phe	Leu	Arg	Leu	Leu	Lys	Ser	His	Asn	Ile	Pro	Val	Thr	
			100					105						110		

10

ES 2 806 732 T3

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

ES 2 806 732 T3

Thr Glu Ala Leu Thr Gly Met Ser Ile Arg Ser Lys Glu Asp Met Glu
370 375 380

Ser Ala Ala Arg Lys Ile Tyr Glu Lys Tyr Gly Cys Ser Val Leu Val
385 390 395 400

Lys Gly Gly His Ser Ile Asn Asp Ala Asn Asp Met Leu Phe Asp Gly
405 410 415

Glu Asn Val Ser Trp Phe Ser Gly Glu Arg Ile Glu Asn Pro Asn Thr
420 425 430

His Gly Thr Gly Cys Thr Leu Ser Ser Ala Ile Ala Ser Asn Leu Ala
435 440 445

Lys Gly Tyr Asp Ile Glu Thr Ser Val Gln Arg Ala Lys Ala Tyr Ile
450 455 460

Ser Gly Ala Leu Ala Ala Met Leu Asp Leu Gly Arg Gly Ser Gly Pro
465 470 475 480

Leu Asn His Gly Phe Asp Ile Asp Ser Arg Phe Met Ile
485 490

<210> 37

<211> 663

5 <212> ADN

<213> *Ruminococcus bromii*

<220>

<221> CDS

10 <222> (1)..(663)

<223> gen de *Ruminococcus bromii* L2-63 que codifica para TMP fosfatasa [CBL14666]

<400> 37

atg att aaa tct gca ata ttt gat gtt gac ggc aca ctt ctc gat tca 48
Met Ile Lys Ser Ala Ile Phe Asp Val Asp Gly Thr Leu Leu Asp Ser
1 5 10 15

atg aag ata tgg gat gat gca gga gag cgt tac ctc tcg tct gtc ggc 96
Met Lys Ile Trp Asp Asp Ala Gly Glu Arg Tyr Leu Ser Ser Val Gly
20 25 30

aaa aca gcc gaa aac gga ctt tcc gaa aag ctc tgt gat atg agt ctg 144
Lys Thr Ala Glu Asn Gly Leu Ser Glu Lys Leu Cys Asp Met Ser Leu
35 40 45

acg gag ggt gcg gag tat atg aaa aag cag tat gct ctt tcc ttt tca 192
Thr Glu Gly Ala Glu Tyr Met Lys Lys Gln Tyr Ala Leu Ser Phe Ser
50 55 60

15 act gat gaa ata gtt tcg ggt gtg ctg aaa atc att gaa gat ttt tac 240

ES 2 806 732 T3

Thr	Asp	Glu	Ile	Val	Ser	Gly	Val	Leu	Lys	Ile	Ile	Glu	Asp	Phe	Tyr	
65					70					75					80	
ttt	tat	gag	gtc	ggg	tta	aaa	aac	gat	gca	aaa	gaa	att	ttg	cag	ttt	288
Phe	Tyr	Glu	Val	Gly	Leu	Lys	Asn	Asp	Ala	Lys	Glu	Ile	Leu	Gln	Phe	
			85						90					95		
ttg	gaa	tcg	aac	aat	atc	aaa	atg	att	att	gca	aca	tca	agc	gac	aaa	336
Leu	Glu	Ser	Asn	Asn	Ile	Lys	Met	Ile	Ile	Ala	Thr	Ser	Ser	Asp	Lys	
			100					105					110			
acg	cat	att	aaa	aag	gca	ttt	gaa	agg	ctc	ggg	att	cta	aaa	tat	ttt	384
Thr	His	Ile	Lys	Lys	Ala	Phe	Glu	Arg	Leu	Gly	Ile	Leu	Lys	Tyr	Phe	
		115					120					125				
acg	gat	att	gtg	acc	tgt	tca	caa	gtc	gga	aaa	ggc	aaa	aca	agc	ccc	432
Thr	Asp	Ile	Val	Thr	Cys	Ser	Gln	Val	Gly	Lys	Gly	Lys	Thr	Ser	Pro	
	130					135					140					
gac	att	tac	ctt	gtc	tgt	gca	gat	aaa	ctc	gga	aca	gct	ccg	agt	gaa	480
Asp	Ile	Tyr	Leu	Val	Cys	Ala	Asp	Lys	Leu	Gly	Thr	Ala	Pro	Ser	Glu	
145					150					155					160	
acg	ctt	gta	ttc	gag	gac	gct	gtt	ttt	gcc	gca	gaa	act	gct	cac	aag	528
Thr	Leu	Val	Phe	Glu	Asp	Ala	Val	Phe	Ala	Ala	Glu	Thr	Ala	His	Lys	
			165						170					175		
gca	ggg	ttc	aaa	acg	gtg	gga	gtg	tat	gac	gaa	ttg	agc	agg	aat	aat	576
Ala	Gly	Phe	Lys	Thr	Val	Gly	Val	Tyr	Asp	Glu	Leu	Ser	Arg	Asn	Asn	
			180					185					190			
aaa	aac	aga	ata	aaa	gcc	gtt	tgc	gat	tac	tac	gca	gac	agc	ttt	gaa	624
Lys	Asn	Arg	Ile	Lys	Ala	Val	Cys	Asp	Tyr	Tyr	Ala	Asp	Ser	Phe	Glu	
		195					200					205				
aaa	gcg	gca	gat	tgg	ggg	cac	cac	ctt	ttg	tcg	ctg	taa				663
Lys	Ala	Ala	Asp	Trp	Gly	His	His	Leu	Leu	Ser	Leu					
	210					215					220					

<210> 38

<211> 220

5 <212> PRT

<213> *Ruminococcus bromii*

<400> 38

Met	Ile	Lys	Ser	Ala	Ile	Phe	Asp	Val	Asp	Gly	Thr	Leu	Leu	Asp	Ser
1			5						10					15	

Met	Lys	Ile	Trp	Asp	Asp	Ala	Gly	Glu	Arg	Tyr	Leu	Ser	Ser	Val	Gly
			20					25					30		

Lys	Thr	Ala	Glu	Asn	Gly	Leu	Ser	Glu	Lys	Leu	Cys	Asp	Met	Ser	Leu
		35					40					45			

Thr	Glu	Gly	Ala	Glu	Tyr	Met	Lys	Lys	Gln	Tyr	Ala	Leu	Ser	Phe	Ser
10	50					55					60				

ES 2 806 732 T3

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

Phe Tyr Glu Val Gly Leu Lys Asn Asp Ala Lys Glu Ile Leu Gln Phe
85 90 95

Leu Glu Ser Asn Asn Ile Lys Met Ile Ile Ala Thr Ser Ser Asp Lys
100 105 110

Thr His Ile Lys Lys Ala Phe Glu Arg Leu Gly Ile Leu Lys Tyr Phe
115 120 125

Thr Asp Ile Val Thr Cys Ser Gln Val Gly Lys Gly Lys Thr Ser Pro
130 135 140

Asp Ile Tyr Leu Val Cys Ala Asp Lys Leu Gly Thr Ala Pro Ser Glu
145 150 155 160

Thr Leu Val Phe Glu Asp Ala Val Phe Ala Ala Glu Thr Ala His Lys
165 170 175

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

Lys Asn Arg Ile Lys Ala Val Cys Asp Tyr Tyr Ala Asp Ser Phe Glu
195 200 205

Lys Ala Ala Asp Trp Gly His His Leu Leu Ser Leu
210 215 220

<210> 39

<211> 1434

5 <212> ADN

<213> *Dorea longicatena*

<220>

<221> CDS

10 <222> (1)..(1434)

<223> gen de *Dorea longicatena* DSM13814 que codifica para TMP fosfatasa [EDM62146]

<400> 39

atg ata aaa gga gca ata ttt gat gta gac gga acc ctt ctg gat tcc 48
Met Ile Lys Gly Ala Ile Phe Asp Val Asp Gly Thr Leu Leu Asp Ser
1 5 10 15

atg gag atc tgg gaa gac gta gga gtc cgt tat ctg aac agt atc ggt 96
Met Glu Ile Trp Glu Asp Val Gly Val Arg Tyr Leu Asn Ser Ile Gly
20 25 30

15 ata gag gca gag ccg gat ctt ggg acg gtg tta ttt aca atg agc atc 144
Ile Glu Ala Glu Pro Asp Leu Gly Thr Val Leu Phe Thr Met Ser Ile

ES 2 806 732 T3

35	40	45	
cag gaa ggt gca gca tat gta aaa gaa cat tat cat ctg tcc cag gag Gln Glu Gly Ala Ala Tyr Val Lys Glu His Tyr His Leu Ser Gln Glu 50 55 60			192
ccg gaa gaa att gtg cag gga gtt ctg gac atc atc agc aat tat tat Pro Glu Glu Ile Val Gln Gly Val Leu Asp Ile Ile Ser Asn Tyr Tyr 65 70 75 80			240
aag aaa acc gca cta tta aag agt gga gtg aag gaa ctt ctg gaa aag Lys Lys Thr Ala Leu Lys Ser Gly Val Lys Glu Leu Leu Glu Lys 85 90 95			288
ctt gat aag cat aat atc cca atg acg gtt gca tca tcc aat aat aaa Leu Asp Lys His Asn Ile Pro Met Thr Val Ala Ser Ser Asn Asn Lys 100 105 110			336
aaa gag ata gag atg gca ttt gag cgt ctg gga att gca aaa tat ttt Lys Glu Ile Glu Met Ala Phe Glu Arg Leu Gly Ile Ala Lys Tyr Phe 115 120 125			384
gac cgg atc ttt acc tgt gaa gag gtc ggt gcg gga aag acg aag ccg Asp Arg Ile Phe Thr Cys Glu Glu Val Gly Ala Gly Lys Thr Lys Pro 130 135 140			432
gat att tat ctg cgg gca gca gaa tat ctc gga acc cgt ccg gag gag Asp Ile Tyr Leu Arg Ala Ala Glu Tyr Leu Gly Thr Arg Pro Glu Glu 145 150 155 160			480
acg gtt gta ttc gaa gat gtc att cat gca atc cgt act gca aag cag Thr Val Val Phe Glu Asp Val Ile His Ala Ile Arg Thr Ala Lys Gln 165 170 175			528
gca ggg ttc cag gtt gta gga atc tat gat gaa gca agt aag gat gac Ala Gly Phe Gln Val Val Gly Ile Tyr Asp Glu Ala Ser Lys Asp Asp 180 185 190			576
cag gaa gag gtt cag aga gaa gta gac tgg tat tgt aga gag tgg gca Gln Glu Glu Val Gln Arg Glu Val Asp Trp Tyr Cys Arg Glu Trp Ala 195 200 205			624
gaa ctt atg aaa aaa aag aca gca att aca atc gcc gga agt gat tca Glu Leu Met Lys Lys Lys Thr Ala Ile Thr Ile Ala Gly Ser Asp Ser 210 215 220			672
agt gga ggt gca gga att cag gca gac atc aag acg atg cag gca aac Ser Gly Gly Ala Gly Ile Gln Ala Asp Ile Lys Thr Met Gln Ala Asn 225 230 235 240			720
gga gtc tac gca atg agt gca atc acc gca ctg aca gcc cag aat aca Gly Val Tyr Ala Met Ser Ala Ile Thr Ala Leu Thr Ala Gln Asn Thr 245 250 255			768
acc gga gta acc gga atc atg gaa gta tct ccg gaa ttt cta gaa caa Thr Gly Val Thr Gly Ile Met Glu Val Ser Pro Glu Phe Leu Glu Gln 260 265 270			816
cag ttg gac gca gtt atc aca gac atc cgt ccg gat gca gtg aaa atc Gln Leu Asp Ala Val Ile Thr Asp Ile Arg Pro Asp Ala Val Lys Ile 275 280 285			864
ggt atg gtg tca tca gaa gag tta ata aaa atg ata tca aag aaa cta			912

ES 2 806 732 T3

Gly	Met	Val	Ser	Ser	Glu	Glu	Leu	Ile	Lys	Met	Ile	Ser	Lys	Lys	Leu	
290						295					300					
aaa	gag	tac	cat	ctg	gag	aat	atc	gta	gtt	gat	cca	gtg	atg	gta	gca	960
Lys	Glu	Tyr	His	Leu	Glu	Asn	Ile	Val	Val	Asp	Pro	Val	Met	Val	Ala	
305					310					315					320	
aca	agc	gga	tcc	aga	ctg	atc	agt	gaa	acc	gcg	att	gat	aca	tta	aaa	1008
Thr	Ser	Gly	Ser	Arg	Leu	Ile	Ser	Glu	Thr	Ala	Ile	Asp	Thr	Leu	Lys	
				325				330						335		
aca	cag	ctg	ctg	cca	atg	gca	act	gtg	atc	aca	ccg	aat	atc	cca	gag	1056
Thr	Gln	Leu	Leu	Pro	Met	Ala	Thr	Val	Ile	Thr	Pro	Asn	Ile	Pro	Glu	
				340				345						350		
gca	gaa	gtt	ctt	gca	gaa	atg	gag	att	aga	tca	gaa	gat	gat	atg	gtg	1104
Ala	Glu	Val	Leu	Ala	Glu	Met	Glu	Ile	Arg	Ser	Glu	Asp	Asp	Met	Val	
		355					360					365				
gaa	gca	gca	aag	aag	att	cat	gaa	atg	tat	cac	tgt	gca	gtc	tta	tgc	1152
Glu	Ala	Ala	Lys	Lys	Ile	His	Glu	Met	Tyr	His	Cys	Ala	Val	Leu	Cys	
						375					380					
aaa	ggc	gga	cac	agc	ctg	aat	gat	gcg	aat	gat	ctc	cta	tac	cag	gat	1200
Lys	Gly	Gly	His	Ser	Leu	Asn	Asp	Ala	Asn	Asp	Leu	Leu	Tyr	Gln	Asp	
385					390					395					400	
gga	gaa	aca	aca	tgg	ttc	cac	gga	aaa	aga	atc	aac	aac	ccg	aac	act	1248
Gly	Glu	Thr	Thr	Trp	Phe	His	Gly	Lys	Arg	Ile	Asn	Asn	Pro	Asn	Thr	
				405					410					415		
cac	gga	acc	ggc	tgt	acc	tta	tcc	agc	gca	atc	gca	tcc	aat	ctg	gca	1296
His	Gly	Thr		Gly	Cys	Thr	Leu	Ser	Ser	Ala	Ile	Ala	Ser	Asn	Leu	Ala
				420				425						430		
aaa	gga	tat	tct	ctg	gaa	gaa	tct	att	cac	cgc	gcg	aaa	gag	tat	atc	1344
Lys	Gly	Tyr	Ser	Leu	Glu	Glu	Ser	Ile	His	Arg	Ala	Lys	Glu	Tyr	Ile	
				435				440					445			
agc	ggg	gcg	ttg	gaa	gcc	atg	tta	gat	ctg	gga	aaa	gga	agc	gga	ccg	1392
Ser	Gly	Ala	Leu	Glu	Ala	Met	Leu	Asp	Leu	Gly	Lys	Gly	Ser	Gly	Pro	
				450			455				460					
atg	gat	cat	ggg	ttt	gag	atg	cgg	ggg	aga	ttt	tct	att	taa			1434
Met	Asp	His	Gly	Phe	Glu	Met	Arg	Gly	Arg	Phe	Ser	Ile				
465					470					475						

<210> 40
 <211> 477
 5 <212> PRT
 <213> *Dorea longicatena*

<400> 40

Met	Ile	Lys	Gly	Ala	Ile	Phe	Asp	Val	Asp	Gly	Thr	Leu	Leu	Asp	Ser
1				5					10					15	

Met	Glu	Ile	Trp	Glu	Asp	Val	Gly	Val	Arg	Tyr	Leu	Asn	Ser	Ile	Gly
			20					25					30		

10

ES 2 806 732 T3

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

ES 2 806 732 T3

Gly Met Val Ser Ser Glu Glu Leu Ile Lys Met Ile Ser Lys Lys Leu
290 295 300

Lys Glu Tyr His Leu Glu Asn Ile Val Val Asp Pro Val Met Val Ala
305 310 315 320

Thr Ser Gly Ser Arg Leu Ile Ser Glu Thr Ala Ile Asp Thr Leu Lys
325 330 335

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

Ala Glu Val Leu Ala Glu Met Glu Ile Arg Ser Glu Asp Asp Met Val
355 360 365

Glu Ala Ala Lys Lys Ile His Glu Met Tyr His Cys Ala Val Leu Cys
370 375 380

Lys Gly Gly His Ser Leu Asn Asp Ala Asn Asp Leu Leu Tyr Gln Asp
385 390 395 400

Gly Glu Thr Thr Trp Phe His Gly Lys Arg Ile Asn Asn Pro Asn Thr
405 410 415

His Gly Thr Gly Cys Thr Leu Ser Ser Ala Ile Ala Ser Asn Leu Ala
420 425 430

Lys Gly Tyr Ser Leu Glu Glu Ser Ile His Arg Ala Lys Glu Tyr Ile
435 440 445

Ser Gly Ala Leu Glu Ala Met Leu Asp Leu Gly Lys Gly Ser Gly Pro
450 455 460

Met Asp His Gly Phe Glu Met Arg Gly Arg Phe Ser Ile
465 470 475

<210> 41

<211> 1305

5 <212> ADN

<213> *Lachnospiraceae bacterium*

<220>

<221> CDS

10 <222> (1)..(1305)

<223> gen de *Lachnospiraceae bacterium_3_1_57FAA_CT1* que codifica para TMP fosfatasa [EPC05128]

<400> 41

15 atg aaa tgt gac aga aag aca atg ctt ctt tat gcg gtg acc gat cgg

48

ES 2 806 732 T3

Met 1	Lys	Cys	Asp	Arg 5	Lys	Thr	Met	Leu	Leu 10	Tyr	Ala	Val	Thr	Asp 15	Arg	
gcc Ala	tgg Trp	aca Thr	gga Gly	gaa Glu	aag Lys	aca Thr	ctg Leu	ctt Leu	atg Met	cag Gln	gtc Val	gag Glu	gaa Glu	gcg Ala	ctg Leu	96
		20					25					30				
gca Ala	gga Gly	ggt Gly	gtg Val	acc Thr	tgt Cys	gtc Val	cag Gln	ctt Leu	cgt Arg	gaa Glu	aag Lys	gat Asp	atg Met	cca Pro	aag Lys	144
		35					40					45				
gag Glu	cag Gln	ttc Phe	ctg Leu	gaa Glu	gaa Glu	gcg Ala	gag Glu	agt Ser	ata Ile	aaa Lys	aga Arg	ctt Leu	tgt Cys	cat His	aaa Lys	192
	50					55					60					
tat Tyr	ggg Gly	atc Ile	cct Pro	ttt Phe	ata Ile	att Ile	gac Asp	gat Asp	gat Asp	gtg Val	gag Glu	ctg Leu	gcc Ala	gta Val	cgc Arg	240
65					70					75					80	
tgc Cys	ggc Gly	gcg Ala	gac Asp	ggg Val	gtg His	cat Val	gtg Gly	gga Gln	cag His	cat Asp	gat Met	atg Glu	gag Ala	gca Gly	ggc Gly	288
				85					90					95		
gcg Ala	gtc Val	cgc Arg	cgg Lys	aaa Ile	atc Gly	gga Asp	gac Gly	ggc Met	atg Leu	ctg Leu	ctg Gly	ggc Val	gta Ser	tca Val	gtc Val	336
			100				105						110			
cag Gln	act Thr	gtg Glu	gaa Gln	cag Ala	gca Val	gtg Val	gaa Ala	gcc Glu	gaa Lys	aaa Lys	aag Lys	gga Gly	gcg Ala	gat Asp	tac Tyr	384
		115					120					125				
ctt Leu	ggt Gly	gtg Val	ggc Gly	gct Ala	gtg Val	ttt Phe	tcc Ser	act Thr	tcc Ser	acg Thr	aaa Lys	acg Thr	gac Asp	gca Ala	cag Gln	432
	130					135					140					
gag Glu	gtt Val	tcc Ser	ctg Leu	gat Asp	acc Thr	ctc Leu	cgg Arg	gaa Glu	atc Ile	tgc Cys	cgg Arg	gcg Ala	gtg Val	tcc Ser	gta Val	480
145					150					155					160	
ccc Pro	gtc Val	tgt Cys	gca Ala	atc Ile	gga Gly	ggg Gly	ata Ile	cac His	aaa Lys	gga Gly	aat Asn	atg Met	cat His	ttg Leu	ctg Leu	528
			165						170				175			
cag Gln	gat Asp	acg Thr	gga Gly	atc Ile	gat Asp	ggg Gly	gtg Val	gct Ala	ttg Leu	gtg Val	tgc Ser	gcc Ala	atc Ile	ttt Phe	tcc Ser	576
		180					185						190			
agt Ser	ccc Pro	tgc Cys	ata Ile	cag Gln	aag Lys	gaa Glu	tgc Cys	agg Arg	gag Glu	ctg Leu	cgg Arg	gtc Val	ctg Leu	gca Ala	gag Glu	624
		195					200					205				
aga Arg	ctg Leu	aaa Lys	agg Arg	aaa Lys	ggg Gly	gct Ala	att Ile	ttt Phe	gat Asp	gcg Ala	gac Asp	gga Gly	acc Thr	ctg Leu	ctg Leu	672
	210					215					220					
gat Asp	tcc Ser	atg Met	tcc Ser	gtt Val	tgg Trp	gat Asp	act Thr	ctg Leu	ggt Gly	gaa Glu	aaa Lys	tat Tyr	ctg Leu	cgg Arg	aaa Lys	720
225					230					235					240	
aag Lys	ggt Gly	att Ile	gtt Val	ccg Pro	gaa Glu	aag Lys	aac Asn	atc Ile	agg Arg	gaa Glu	aca Thr	ata Ile	aaa Lys	aat Asn	atg Met	768
				245					250					255		

ES 2 806 732 T3

```

agt ctt cct cag gct gcg gtc tat ttt cag act gct tat ggg att gcg      816
Ser Leu Pro Gln Ala Ala Val Tyr Phe Gln Thr Ala Tyr Gly Ile Ala
      260      265      270

gat gca gaa gac aag att ata gag gat att aat gga ata gcg gcg tcc      864
Asp Ala Glu Asp Lys Ile Ile Glu Asp Ile Asn Gly Ile Ala Ala Ser
      275      280      285

ttt tac atc aat gag gtg aag ctg aag gaa ggc gtg aaa acg gtt ctg      912
Phe Tyr Ile Asn Glu Val Lys Leu Lys Glu Gly Val Lys Thr Val Leu
      290      295      300

gac aag ctg aag cag aaa aac gta aag atg tgt gtg gcg acg gct acg      960
Asp Lys Leu Lys Gln Lys Asn Val Lys Met Cys Val Ala Thr Ala Thr
305      310      315      320

gac aaa ggg ctg att gaa aag gca ctt gag aga aac gga atc aga gat      1008
Asp Lys Gly Leu Ile Glu Lys Ala Leu Glu Arg Asn Gly Ile Arg Asp
      325      330      335

tat ttt gag gct gtc ctc acc tgc acg gat gtg ggc gcg gga aag gat      1056
Tyr Phe Glu Ala Val Leu Thr Cys Thr Asp Val Gly Ala Gly Lys Asp
      340      345      350

gag ccg gtt atc ttc cgt aag gcc ggg cag ctt ctc gga aca gca aaa      1104
Glu Pro Val Ile Phe Arg Lys Ala Gly Gln Leu Leu Gly Thr Ala Lys
      355      360      365

gag gat acc att gta att gaa gat gcc ttg tat gct gtt aag aca gcg      1152
Glu Asp Thr Ile Val Ile Glu Asp Ala Leu Tyr Ala Val Lys Thr Ala
      370      375      380

aaa gag gac ggt ttc ctg gtg gcg gct gtt tat gat ccg tca gca gaa      1200
Lys Glu Asp Gly Phe Leu Val Ala Ala Val Tyr Asp Pro Ser Ala Glu
385      390      395      400

aag gag gaa ccg gag atc cgg gag atc tct gac ttc tat ttc cgg tca      1248
Lys Glu Glu Pro Glu Ile Arg Glu Ile Ser Asp Phe Tyr Phe Arg Ser
      405      410      415

ttt aat gaa atg gag agt tat ctg aat gaa aaa agt tct tac gat agc      1296
Phe Asn Glu Met Glu Ser Tyr Leu Asn Glu Lys Ser Ser Tyr Asp Ser
      420      425      430

ggg ctc tga      1305
Gly Leu

```

<210> 42

<211> 434

5 <212> PRT

<213> Bacteria *Lachnospiraceae*

<400> 42

```

Met Lys Cys Asp Arg Lys Thr Met Leu Leu Tyr Ala Val Thr Asp Arg
1      5      10      15

```

```

10 Ala Trp Thr Gly Glu Lys Thr Leu Leu Met Gln Val Glu Glu Ala Leu
      20      25      30

```

ES 2 806 732 T3

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

ES 2 806 732 T3

Phe Tyr Ile Asn Glu Val Lys Leu Lys Glu Gly Val Lys Thr Val Leu
290 295 300

Asp Lys Leu Lys Gln Lys Asn Val Lys Met Cys Val Ala Thr Ala Thr
305 310 315 320

Asp Lys Gly Leu Ile Glu Lys Ala Leu Glu Arg Asn Gly Ile Arg Asp
325 330 335

Tyr Phe Glu Ala Val Leu Thr Cys Thr Asp Val Gly Ala Gly Lys Asp
340 345 350

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

Glu Asp Thr Ile Val Ile Glu Asp Ala Leu Tyr Ala Val Lys Thr Ala
370 375 380

Lys Glu Asp Gly Phe Leu Val Ala Ala Val Tyr Asp Pro Ser Ala Glu
385 390 395 400

Lys Glu Glu Pro Glu Ile Arg Glu Ile Ser Asp Phe Tyr Phe Arg Ser
405 410 415

Phe Asn Glu Met Glu Ser Tyr Leu Asn Glu Lys Ser Ser Tyr Asp Ser
420 425 430

Gly Leu

<210> 43

<211> 1305

5 <212> ADN

<213> *Fusicatenibacter*

<220>

<221> CDS

10 <222> (1)..(1305)

<223> gen de *Fusicatenibacter* que codifica para TMP fosfatasa [CUQ30753]

<400> 43

atg aaa tgt aac aga aag aca atg ctt ctt tat gcg gtg acc gac cgg 48
Met Lys Cys Asn Arg Lys Thr Met Leu Leu Tyr Ala Val Thr Asp Arg
1 5 10 15

gcc tgg aca gga gaa aag aca ctg ctt acg cag gtc gag gaa gcg ctg 96
Ala Trp Thr Gly Glu Lys Thr Leu Leu Thr Gln Val Glu Glu Ala Leu
20 25 30

15 gca gga ggt gta acc tgt gtc cag ctt cgt gaa aag gat atg cca aag 144

ES 2 806 732 T3

Ala	Gly	Gly	Val	Thr	Cys	Val	Gln	Leu	Arg	Glu	Lys	Asp	Met	Pro	Lys	
	35						40					45				
gag	cag	ttc	ctg	gaa	gaa	gcg	gag	agt	ata	aaa	aga	ctt	tgc	cat	aaa	192
Glu	Gln	Phe	Leu	Glu	Glu	Ala	Glu	Ser	Ile	Lys	Arg	Leu	Cys	His	Lys	
	50					55				60						
tat	ggt	gtc	cct	ttt	ata	att	gac	gat	gat	gtg	gag	ctg	gcc	gta	cgc	240
Tyr	Gly	Val	Pro	Phe	Ile	Ile	Asp	Asp	Asp	Val	Glu	Leu	Ala	Val	Arg	
65				70						75					80	
tgc	ggt	gcg	gac	ggg	gta	cat	gtg	gga	cag	cat	gat	atg	gag	gca	ggc	288
Cys	Gly	Ala	Asp	Gly	Val	His	Val	Gly	Gln	His	Asp	Met	Glu	Ala	Gly	
			85						90					95		
gcg	gtc	cgc	cgg	aaa	atc	gga	gac	ggc	atg	ctg	ctg	ggc	gta	tca	gtc	336
Ala	Val	Arg		Lys	Ile	Gly	Asp	Gly	Met	Leu	Leu	Gly	Val	Ser	Val	
			100					105					110			
cag	act	gtg	gaa	cag	gca	gtg	gaa	gcc	gag	aaa	aag	gga	gcg	gat	tac	384
Gln	Thr	Val	Glu	Gln	Ala	Val	Glu	Ala	Glu	Lys	Lys	Gly	Ala	Asp	Tyr	
		115					120					125				
ctt	ggt	gtg	ggc	gct	gtg	ttt	tcc	act	tcc	acg	aaa	acg	gac	gca	cag	432
Leu	Gly	Val	Gly	Ala	Val		Ser	Thr	Ser	Thr	Lys	Thr	Asp	Ala	Gln	
	130					135					140					
gag	gtt	tcc	ctg	gat	acc	ctc	cgg	gaa	atc	tgc	cgg	gcg	gtg	tcc	gta	480
Glu	Val	Ser	Leu	Asp	Thr	Leu	Arg	Glu	Ile	Cys	Arg	Ala	Val	Ser	Val	
145					150					155					160	
ccc	gtc	tgt	gca	atc	gga	ggg	ata	cac	aaa	gga	aat	atg	cat	ttg	ctg	528
Pro	Val	Cys	Ala	Ile	Gly	Gly	Ile	His	Lys	Gly	Asn	Met	His	Leu	Leu	
			165						170					175		
cag	gat	acg	gga	atc	gat	ggg	gtg	gct	ttg	gtg	tgc	gcc	atc	ttt	tcc	576
Gln	Asp	Thr	Gly	Ile	Asp	Gly	Val	Ala	Leu	Val	Ser	Ala	Ile	Phe	Ser	
		180					185						190			
agt	ccc	tgc	ata	cag	aag	gaa	tgc	agg	gag	ctg	cgg	gcc	ctg	gca	gag	624
Ser	Pro	Cys	Ile	Gln	Lys	Glu	Cys	Arg	Glu	Leu	Arg	Ala	Leu	Ala	Glu	
		195				200						205				
agg	ctg	aaa	agg	aaa	ggg	gct	att	ttt	gat	gcg	gac	gga	acc	ctg	ctg	672
Arg	Leu	Lys	Arg	Lys	Gly	Ala	Ile	Phe	Asp	Ala	Asp	Gly	Thr	Leu	Leu	
	210				215						220					
gat	tcc	atg	tct	gtt	tgg	gat	acc	ctg	ggt	gaa	aaa	tat	ctg	cgg	aaa	720
Asp	Ser	Met	Ser	Val	Trp	Asp	Thr	Leu	Gly	Glu	Lys	Tyr	Leu	Arg	Lys	
225					230					235					240	
aag	ggt	att	gtt	ccg	gaa	aag	aac	atc	agg	gaa	aca	ata	aaa	aat	atg	768
Lys	Gly	Ile	Val	Pro	Glu	Lys	Asn	Ile	Arg	Glu	Thr	Ile	Lys	Asn	Met	
			245						250					255		
agt	ctt	cct	cag	gcc	gcg	gtc	tat	ttt	caa	act	gct	tat	ggg	atc	acg	816
Ser	Leu	Pro	Gln	Ala	Ala	Val	Tyr	Phe	Gln	Thr	Ala	Tyr	Gly	Ile	Thr	
		260					265						270			
gat	gca	gaa	gac	aag	att	ata	gag	gat	att	aat	gga	ata	gcg	gcg	tcc	864
Asp	Ala	Glu	Asp	Lys	Ile	Ile	Glu	Asp	Ile	Asn	Gly	Ile	Ala	Ala	Ser	
		275					280					285				

ES 2 806 732 T3

ttt tac atc aat gag gtg aag ctg aag gaa ggc gtg aaa acg gtt ctg 912
 Phe Tyr Ile Asn Glu Val Lys Leu Lys Glu Gly Val Lys Thr Val Leu
 290 295 300

 gac aag ctg aag cag aaa aac gta aag atg tgt gtg gcg acg gct acg 960
 Asp Lys Leu Lys Gln Lys Asn Val Lys Met Cys Val Ala Thr Ala Thr
 305 310 315 320

 gac aag ggg ctg att gaa aag gca ctt gag aga aac gga atc aga gat 1008
 Asp Lys Gly Leu Ile Glu Lys Ala Leu Glu Arg Asn Gly Ile Arg Asp
 325 330 335

 tat ttt gag gct gtc ctc acc tgc acg gat gtg ggc gcg gga aag gat 1056
 Tyr Phe Glu Ala Val Leu Thr Cys Thr Asp Val Gly Ala Gly Lys Asp
 340 345 350

 gag ccg gtt atc ttc cgt aag gcc ggg cag ctt ctc gga acc gca aaa 1104
 Glu Pro Val Ile Phe Arg Lys Ala Gly Gln Leu Leu Gly Thr Ala Lys
 355 360 365

 gag gat acc att gta att gaa gat gcc ttg tat gct gtt aag aca gcg 1152
 Glu Asp Thr Ile Val Ile Glu Asp Ala Leu Tyr Ala Val Lys Thr Ala
 370 375 380

 aaa gag gac ggt ttc ctg gtg gcg gct gtt tat gat ccg tca gca gaa 1200
 Lys Glu Asp Gly Phe Leu Val Ala Ala Val Tyr Asp Pro Ser Ala Glu
 385 390 395 400

 aag gag gaa ccg gag atc cgg gag atc tct gac ttc tat ttc cgg tca 1248
 Lys Glu Glu Pro Glu Ile Arg Glu Ile Ser Asp Phe Tyr Phe Arg Ser
 405 410 415

 ttt aat gaa atg gag agt tat ctg aat gaa aaa agt tct tac gat agc 1296
 Phe Asn Glu Met Glu Ser Tyr Leu Asn Glu Lys Ser Ser Tyr Asp Ser
 420 425 430

 ggg ctc tga 1305
 Gly Leu

<210> 44
 <211> 434
 <212> PRT
 <213> *Fusicatenibacter*

<400> 44

Met Lys Cys Asn Arg Lys Thr Met Leu Leu Tyr Ala Val Thr Asp Arg
 1 5 10 15

Ala Trp Thr Gly Glu Lys Thr Leu Leu Thr Gln Val Glu Glu Ala Leu
 20 25 30

Ala Gly Gly Val Thr Cys Val Gln Leu Arg Glu Lys Asp Met Pro Lys
 35 40 45

Glu Gln Phe Leu Glu Glu Ala Glu Ser Ile Lys Arg Leu Cys His Lys
 50 55 60

ES 2 806 732 T3

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

ES 2 806 732 T3

Asp Lys Gly Leu Ile Glu Lys Ala Leu Glu Arg Asn Gly Ile Arg Asp
325 330 335

Tyr Phe Glu Ala Val Leu Thr Cys Thr Asp Val Gly Ala Gly Lys Asp
340 345 350

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

Glu Asp Thr Ile Val Ile Glu Asp Ala Leu Tyr Ala Val Lys Thr Ala
370 375 380

Lys Glu Asp Gly Phe Leu Val Ala Ala Val Tyr Asp Pro Ser Ala Glu
385 390 395 400

Lys Glu Glu Pro Glu Ile Arg Glu Ile Ser Asp Phe Tyr Phe Arg Ser
405 410 415

Phe Asn Glu Met Glu Ser Tyr Leu Asn Glu Lys Ser Ser Tyr Asp Ser
420 425 430

Gly Leu

<210> 45

<211> 1296

5 <212> ADN

<213> Especie *Clostridium*

<220>

<221> CDS

10 <222> (1)..(1296)

<223> gen de *Clostridium* sp KLE1755 que codifica para TMP fosfatasa [ERI68966]:

<400> 45

atg aaa tgt gac aga agc atg ctg ctc ctc tat gcc gta acc gac cgt 48
Met Lys Cys Asp Arg Ser Met Leu Leu Leu Tyr Ala Val Thr Asp Arg
1 5 10 15

gcc tgg acg ggt aaa aaa aca ctg ctg cag cag gtg gag gaa gcc ctg 96
Ala Trp Thr Gly Lys Lys Thr Leu Leu Gln Gln Val Glu Glu Ala Leu
20 25 30

gca ggc ggc gcc acc tgc atc cag ctt cgg gaa aag gag ctg ccg gag 144
Ala Gly Gly Ala Thr Cys Ile Gln Leu Arg Glu Lys Glu Leu Pro Glu
35 40 45

gaa gaa ttc cgg cag gaa gcc ctg gct gtg aaa gaa ctt tgc cgc aga 192
Glu Glu Phe Arg Gln Glu Ala Leu Ala Val Lys Glu Leu Cys Arg Arg
50 55 60

15 tac cat gtc cct ttc ctc att aac gac aac gta gag ctg gct gtc agc 240

ES 2 806 732 T3

Tyr 65	His	Val	Pro	Phe	Leu 70	Ile	Asn	Asp	Asn	Val 75	Glu	Leu	Ala	Val	Ser 80		
tgc	ggc	gcg	gac	ggc	gtc	cat	gtg	ggc	cag	cac	gac	atg	tct	gcg	gcg	288	
Cys	Gly	Ala	Asp	Gly	Val	His	Val	Gly	Gln	His	Asp	Met	Ser	Ala	Ala		
				85					90					95			
gat	gtg	cgc	cgc	aga	atc	ggc	ccc	ggc	aaa	ata	ctg	gga	gta	tcc	gcg	336	
Asp	Val	Arg	Arg	Arg	Ile	Gly	Pro	Gly	Lys	Ile	Leu	Gly	Val	Ser	Ala		
			100					105					110				
cag	acg	gtg	gag	cag	gcc	cgc	cag	gcg	gaa	gaa	gac	ggc	gca	gat	tat	384	
Gln	Thr	Val	Glu	Gln	Ala	Arg	Gln	Ala	Glu	Glu	Asp	Gly	Ala	Asp	Tyr		
		115					120					125					
ctg	ggc	gtg	ggc	gct	gtt	ttt	tcc	acc	tcc	acc	aaa	tcc	gac	gca	gac	432	
Leu	Gly	Val	Gly	Ala	Val		Phe	Ser	Thr	Ser	Thr	Lys	Ser	Asp	Ala	Asp	
	130						135					140					
gcg	gta	tcc	cat	gag	acc	ctg	caa	aag	atc	tgc	gcc	gca	gta	tcc	atc	480	
Ala	Val	Ser	His	Glu	Thr	Leu	Gln	Lys	Ile	Cys	Ala	Ala	Val	Ser	Ile		
	145				150					155					160		
ccc	gtc	tgc	gcc	ata	ggc	ggc	atc	cat	aaa	gaa	aat	ctg	cat	ttg	ctc	528	
Pro	Val	Cys	Ala	Ile	Gly	Gly	Ile	His	Lys	Glu	Asn	Leu	His	Leu	Leu		
			165						170					175			
aaa	ggc	aca	ggc	atc	gcc	ggc	gtg	gcc	ctt	gtt	tcc	gcc	atc	ttc	gca	576	
Lys	Gly	Thr		Ile	Ala	Gly	Val	Ala	Leu	Val	Ser	Ala	Ile	Phe	Ala		
		180						185					190				
agc	ccg	gat	atc	cgt	aag	tcc	tgc	gaa	gac	ctg	aaa	aaa	ctg	gcc	ctg	624	
Ser	Pro	Asp	Ile	Arg	Lys	Ser	Cys	Glu	Asp	Leu	Lys	Lys	Leu	Ala	Leu		
		195					200					205					
cag	ata	aac	gcg	cag	gac	aca	ctg	gaa	gca	ctg	ctg	cat	aca	aac	atc	672	
Gln	Ile	Asn	Ala	Gln	Asp	Thr	Leu	Glu	Ala	Leu	Leu	His	Thr	Asn	Ile		
	210					215					220						
cgc	gga	gcc	atc	ttt	gac	gcg	gac	ggc	acc	ctt	tta	gac	tcc	atg	ggc	720	
Arg	Gly	Ala	Ile	Phe	Asp	Ala	Asp	Gly	Thr	Leu	Leu	Asp	Ser	Met	Gly		
	225				230				235					240			
atc	tgg	gat	act	ctg	ggg	gaa	gat	tac	ctg	cgt	aca	aaa	ggg	aaa	atc	768	
Ile	Trp	Asp	Thr	Leu	Gly	Glu	Asp	Tyr	Leu	Arg	Thr	Lys	Gly	Lys	Ile		
			245					250					255				
ccc	cgg	gaa	aac	ctg	cgt	gaa	acc	ttc	cgc	gac	atg	agc	ctt	ctc	cag	816	
Pro	Arg	Glu	Asn	Leu	Arg	Glu	Thr	Phe	Arg	Asp	Met	Ser	Leu	Leu	Gln		
			260					265					270				
gcc	gcc	tgc	tat	tac	cgg	gaa	aat	tac	gcc	ctt	acg	gaa	agc	cct	gaa	864	
Ala	Ala	Cys	Tyr	Tyr	Arg	Glu	Asn	Tyr	Ala	Leu	Thr	Glu	Ser	Pro	Glu		
		275					280					285					
aaa	ata	gtg	gaa	gag	ctt	aac	gcc	atg	atc	gcc	tcc	ttc	tat	gaa	aaa	912	
Lys	Ile	Val	Glu	Glu	Leu	Asn	Ala	Met	Ile	Ala	Ser	Phe	Tyr	Glu	Lys		
	290					295					300						
gaa	gcc	ccc	ctg	aag	gaa	gga	gcc	gcc	gcc	ttc	ctg	gaa	gcg	ctt	tgc	960	
Glu	Ala	Pro	Leu	Lys	Glu	Gly	Ala	Ala	Ala	Phe	Leu	Glu	Ala	Leu	Cys		
	305				310				315						320		

ES 2 806 732 T3

caa aga aac ata aaa atg tgc att gca aca gcc acc gat cac agc ctt 1008
Gln Arg Asn Ile Lys Met Cys Ile Ala Thr Ala Thr Asp His Ser Leu
325 330 335

atc cgg gcc gcc ctg aag cga tgc gga gtg ctg cat tac ttt act ttt 1056
Ile Arg Ala Ala Leu Lys Arg Cys Gly Val Leu His Tyr Phe Thr Phe
340 345 350

ata ctt acc tgc gga caa gca gga gcg gga aaa gac acc ccc gcc att 1104
Ile Leu Thr Cys Gly Gln Ala Gly Ala Gly Lys Asp Thr Pro Ala Ile
355 360 365

tat gaa gaa gcc ctg gcc ctg ctt gga acc gga aaa aaa gaa acc ttc 1152
Tyr Glu Glu Ala Leu Ala Leu Leu Gly Thr Gly Lys Lys Glu Thr Phe
370 375 380

gtt ttt gaa gat gcc ctg tac gcc ctg aaa acg gcg aaa aca gcc ggc 1200
Val Phe Glu Asp Ala Leu Tyr Ala Leu Lys Thr Ala Lys Thr Ala Gly
385 390 395 400

ttt cct aca gtc ggt gta aaa gac ccc tcc tcc gcc gga cag gaa ggg 1248
Phe Pro Thr Val Gly Val Lys Asp Pro Ser Ser Ala Gly Gln Glu Gly
405 410 415

gag att ata aaa caa gcc gat tac tat ctt tat acc ttc acg aaa tga 1296
Glu Ile Ile Lys Gln Ala Asp Tyr Tyr Leu Tyr Thr Phe Thr Lys
420 425 430

<210> 46

<211> 431

5 <212> PRT

<213> Especie *Clostridium*

<400> 46

Met Lys Cys Asp Arg Ser Met Leu Leu Leu Tyr Ala Val Thr Asp Arg
1 5 10 15

Ala Trp Thr Gly Lys Lys Thr Leu Leu Gln Gln Val Glu Glu Ala Leu
20 25 30

Ala Gly Gly Ala Thr Cys Ile Gln Leu Arg Glu Lys Glu Leu Pro Glu
35 40 45

Glu Glu Phe Arg Gln Glu Ala Leu Ala Val Lys Glu Leu Cys Arg Arg
50 55 60

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

Cys Gly Ala Asp Gly Val His Val Gly Gln His Asp Met Ser Ala Ala
85 90 95

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

10

ES 2 806 732 T3

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

ES 2 806 732 T3

Tyr Glu Glu Ala Leu Ala Leu Leu Gly Thr Gly Lys Lys Glu Thr Phe
370 375 380

Val Phe Glu Asp Ala Leu Tyr Ala Leu Lys Thr Ala Lys Thr Ala Gly
385 390 395 400

Phe Pro Thr Val Gly Val Lys Asp Pro Ser Ser Ala Gly Gln Glu Gly
405 410 415

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

<210> 47

<211> 1452

5 <212> ADN

<213> *Eubacterium hallii*

<220>

<221> CDS

10 <222> (1)..(1452)

<223> gen de *Eubacterium hallii* que codifica para TMP fosfatasa [EEG35494]

<400> 47

atg ata aaa gga gca atc ttt gat att gat gga act tta ctt gat tcc 48
Met Ile Lys Gly Ala Ile Phe Asp Ile Asp Gly Thr Leu Leu Asp Ser
1 5 10 15

atg ccc atc tgg gaa aat gca gga gcg aga tat ctt gct act ctt ggc 96
Met Pro Ile Trp Glu Asn Ala Gly Ala Arg Tyr Leu Ala Thr Leu Gly
20 25 30

att aag gca aag cca gat tta aaa gaa cgg ctg gat gct tta tct ttg 144
Ile Lys Ala Lys Pro Asp Leu Lys Glu Arg Leu Asp Ala Leu Ser Leu
35 40 45

cca gaa gga gcc atc tat atg caa aaa gag tat ggc ctt tcg gta tca 192
Pro Glu Gly Ala Ile Tyr Met Gln Lys Glu Tyr Gly Leu Ser Val Ser
50 55 60

gca gaa gac att tta gaa gga gtc aat cag gtt gta aaa gat ttt tac 240
Ala Glu Asp Ile Leu Glu Gly Val Asn Gln Val Val Lys Asp Phe Tyr
65 70 75 80

tat aaa gaa gcg gtc atg aag ccg gga gcc tat gcc tta gta aaa cgt 288
Tyr Lys Glu Ala Val Met Lys Pro Gly Ala Tyr Ala Leu Val Lys Arg
85 90 95

ctg aaa gaa aat ggt gtg aag tta att ata gcc aca gcg aca gat aag 336
Leu Lys Glu Asn Gly Val Lys Leu Ile Ile Ala Thr Ala Thr Asp Lys
100 105 110

gag atg gca aag gcg gcg ctt att cgt aac ggc ata tgg cag gac ttt 384
Glu Met Ala Lys Ala Ala Leu Ile Arg Asn Gly Ile Trp Gln Asp Phe
115 120 125

15 acg gga atg att acc tgc gag gaa gcc gga gcc gga aag aca agc ccg 432

ES 2 806 732 T3

Thr	Gly	Met	Ile	Thr	Cys	Glu	Glu	Ala	Gly	Ala	Gly	Lys	Thr	Ser	Pro	
130						135					140					
aag	gta	ttt	gag	ctt	gca	agg	caa	aag	cta	ggc	act	aaa	aaa	gag	gaa	480
Lys	Val	Phe	Glu	Leu	Ala	Arg	Gln	Lys	Leu	Gly	Thr	Lys	Lys	Glu	Glu	
145					150					155					160	
aca	tgg	gta	ttt	gaa	gat	tct	tta	tat	gcg	gtg	aaa	act	gct	act	gaa	528
Thr	Trp	Val	Phe	Glu	Asp	Ser	Leu	Tyr	Ala	Val	Lys	Thr	Ala	Thr	Glu	
				165					170					175		
gct	gga	ttt	cca	gta	tgc	agt	atc	tac	gat	acc	tac	agt	gtg	gga	aat	576
Ala	Gly	Phe	Pro	Val	Cys	Ser	Ile	Tyr	Asp	Thr	Tyr	Ser	Val	Gly	Asn	
			180					185					190			
gcg	aaa	gaa	atc	cag	aaa	ctt	tct	aat	att	tat	gtg	aga	gat	ttt	tcg	624
Ala	Lys	Glu	Ile	Gln	Lys	Leu	Ser	Asn	Ile	Tyr	Val	Arg	Asp	Phe	Ser	
		195				200						205				
gag	ata	ggg	gat	tat	tct	ttt	tca	aat	atg	aaa	aca	gtt	ctt	aca	att	672
Glu	Ile	Gly	Asp	Tyr	Ser	Phe	Ser	Asn	Met	Lys	Thr	Val	Leu	Thr	Ile	
	210					215						220				
gca	ggc	agt	gat	tcg	agc	gga	gga	gca	ggg	att	caa	gcg	gat	atc	aag	720
Ala	Gly	Ser	Asp	Ser	Ser	Gly	Gly	Ala	Gly	Ile	Gln	Ala	Asp	Ile	Lys	
225					230					235					240	
act	tta	act	gtt	cat	aaa	gta	tat	gcc	atg	act	tgt	atc	acc	gca	ctt	768
Thr	Leu	Thr	Val	His	Lys	Val	Tyr	Ala	Met	Thr	Cys	Ile	Thr	Ala	Leu	
				245					250					255		
acc	gca	caa	aat	aca	gtc	gga	att	acc	ggg	att	atg	cca	gta	cca	gca	816
Thr	Ala	Gln	Asn	Thr	Val	Gly	Ile	Thr	Gly	Ile	Met	Pro	Val	Pro	Ala	
			260					265					270			
gaa	ttt	ttt	aaa	aaa	cag	atg	gaa	agc	att	ttc	aca	gat	ata	aag	cca	864
Glu	Phe	Phe	Lys	Lys	Gln	Met	Glu	Ser	Ile	Phe	Thr	Asp	Ile	Lys	Pro	
		275				280						285				
gat	gcg	gtg	aaa	att	gga	atg	att	gct	tca	aag	gaa	cag	gca	gag	att	912
Asp	Ala	Val	Lys	Ile	Gly	Met	Ile	Ala	Ser	Lys	Glu	Gln	Ala	Glu	Ile	
	290					295					300					
atc	gca	gaa	tac	ctg	gaa	aaa	tat	tct	atc	aaa	aat	gta	gtg	gca	gac	960
Ile	Ala	Glu	Tyr	Leu	Glu	Lys	Tyr	Ser	Ile	Lys	Asn	Val	Val	Ala	Asp	
305				310						315					320	
ccg	gtg	atg	att	tcg	aca	agc	ggg	acg	gtt	tta	gta	gaa	gaa	aca	acg	1008
Pro	Val	Met	Ile	Ser	Thr	Ser	Gly	Thr	Val	Leu	Val	Glu	Glu	Thr	Thr	
				325					330					335		
aga	aag	ata	tta	tat	gag	aaa	tta	tat	cca	aaa	gtt	tcc	ctg	cta	acc	1056
Arg	Lys	Ile	Leu	Tyr	Glu	Lys	Leu	Tyr	Pro	Lys	Val	Ser	Leu	Leu	Thr	
			340					345					350			
ccg	aac	att	cca	gaa	acc	gaa	ttt	tta	tcc	ggg	ata	aaa	att	acc	gat	1104
Pro	Asn	Ile	Pro	Glu	Thr	Glu	Phe	Leu	Ser	Gly	Ile	Lys	Ile	Thr	Asp	
		355					360					365				
aaa	aaa	aca	agg	gaa	gaa	gca	gca	aaa	gtc	att	gca	gac	agg	tgg	aat	1152
Lys	Lys	Thr	Arg	Glu	Glu	Ala	Ala	Lys	Val	Ile	Ala	Asp	Arg	Trp	Asn	
	370					375					380					

ES 2 806 732 T3

tgt gcg gtc tta agt aag ggc ggt cac agc gaa gaa aat gcg gac gat 1200
Cys Ala Val Leu Ser Lys Gly Gly His Ser Glu Glu Asn Ala Asp Asp
385 390 395 400

ttg ctt tat gag agt ttt ttg cag gaa gaa aaa aaa gaa aaa gcc gtt 1248
Leu Leu Tyr Glu Ser Phe Leu Gln Glu Glu Lys Lys Glu Lys Ala Val
405 410 415

tgg ttt cca gaa gaa aga att gat aat cca aac aca cac gga acc ggc 1296
Trp Phe Pro Glu Glu Arg Ile Asp Asn Pro Asn Thr His Gly Thr Gly
420 425 430

tgt aca ctt tca agt gcg gta gcg gca aat ctg gca aag gga ttt cct 1344
Cys Thr Leu Ser Ser Ala Val Ala Ala Asn Leu Ala Lys Gly Phe Pro
435 440 445

gta gaa gaa tcc gta aaa aag gca aaa gta tac atc agc gga gca att 1392
Val Glu Glu Ser Val Lys Lys Ala Lys Val Tyr Ile Ser Gly Ala Ile
450 455 460

aga gca atg ctg aat ctt gga cag gga aat ggc ccg cta aat cat atg 1440
Arg Ala Met Leu Asn Leu Gly Gln Gly Asn Gly Pro Leu Asn His Met
465 470 475 480

tgg gat ttg taa 1452
Trp Asp Leu

<210> 48

<211> 483

5 <212> PRT

<213> *Eubacterium hallii*

<400> 48

Met Ile Lys Gly Ala Ile Phe Asp Ile Asp Gly Thr Leu Leu Asp Ser
1 5 10 15

Met Pro Ile Trp Glu Asn Ala Gly Ala Arg Tyr Leu Ala Thr Leu Gly
20 25 30

Ile Lys Ala Lys Pro Asp Leu Lys Glu Arg Leu Asp Ala Leu Ser Leu
35 40 45

Pro Glu Gly Ala Ile Tyr Met Gln Lys Glu Tyr Gly Leu Ser Val Ser
50 55 60

Ala Glu Asp Ile Leu Glu Gly Val Asn Gln Val Val Lys Asp Phe Tyr
65 70 75 80

Tyr Lys Glu Ala Val Met Lys Pro Gly Ala Tyr Ala Leu Val Lys Arg
85 90 95

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

10

ES 2 806 732 T3

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

ES 2 806 732 T3

Lys Lys Thr Arg Glu Glu Ala Ala Lys Val Ile Ala Asp Arg Trp Asn
370 375 380

Cys Ala Val Leu Ser Lys Gly Gly His Ser Glu Glu Asn Ala Asp Asp
385 390 395 400

Leu Leu Tyr Glu Ser Phe Leu Gln Glu Glu Lys Lys Glu Lys Ala Val
405 410 415

Trp Phe Pro Glu Glu Arg Ile Asp Asn Pro Asn Thr His Gly Thr Gly
420 425 430

Cys Thr Leu Ser Ser Ala Val Ala Ala Asn Leu Ala Lys Gly Phe Pro
435 440 445

Val Glu Glu Ser Val Lys Lys Ala Lys Val Tyr Ile Ser Gly Ala Ile
450 455 460

Arg Ala Met Leu Asn Leu Gly Gln Gly Asn Gly Pro Leu Asn His Met
465 470 475 480

Trp Asp Leu

<210> 49

<211> 1362

5 <212> ADN

<213> Especie *Eubacterium*

<220>

<221> CDS

10 <222> (1)..(1362)

<223> gen de *Eubacterium* sp. CAG:252 que codifica para TMP fosfatasa [CDB67556]

<400> 49

atg aaa aat aaa ttt ttc aca cgc gag att tgt gtc tgc gtg cac ttg 48
Met Lys Asn Lys Phe Phe Thr Arg Glu Ile Cys Val Cys Val His Leu
1 5 10 15

aca caa act cgt tat gcg caa aaa acg tgc gca gaa atg agg aat agt 96
Thr Gln Thr Arg Tyr Ala Gln Lys Thr Cys Ala Glu Met Arg Asn Ser
20 25 30

gtg aag gtt aaa gct gag gat atg cag cta tac gct gtt aca gat aca 144
Val Lys Val Lys Ala Glu Asp Met Gln Leu Tyr Ala Val Thr Asp Thr
35 40 45

cag tgg ctt aat gga cgt gac ttt ctt gaa gta ata gaa agc gtt ctt 192
Gln Trp Leu Asn Gly Arg Asp Phe Leu Glu Val Ile Glu Ser Val Leu
50 55 60

15 gca aat gga gct aca ttt tta cag tta agg gaa aaa aat gcc aca cat 240

ES 2 806 732 T3

Ala 65	Asn	Gly	Ala	Thr	Phe 70	Leu	Gln	Leu	Arg	Glu 75	Lys	Asn	Ala	Thr	His 80	
gag Glu	gaa Glu	ata Ile	gtg Val	gca Ala	aag Lys	gcg Ala	aag Lys	gca Ala	ata Ile	aag Lys	cca Pro	ata Ile	gct Ala	aag Lys	aag Lys	288
tac Tyr	gga Gly	gtg Val	cct Pro	ttt Phe	gtc Val	ata Ile	gat Asp	gat Asp	gac Asp	ata Ile	tat Tyr	gca Ala	gct Ala	aaa Lys	gag Glu	336
gca Ala	gac Asp	gtg Val	gat Asp	ggt Gly	gtc Val	cac His	ata Ile	ggg Gly	cag Gln	aat Asn	gat Asp	gca Ala	agc Ser	tat Tyr	gag Glu	384
aag Lys	gca Ala	aga Arg	gaa Glu	gtt Val	ctt Leu	gga Gly	gaa Glu	ggc Gly	aag Lys	ata Ile	ata Ile	gga Gly	atg Met	acg Thr	gtc Val	432
aag Lys	aca Thr	agg Arg	cag Gln	cag Gln	gca Ala	gaa Glu	aat Asn	gcc Ala	ata Ile	aga Arg	ctt Leu	ggc Gly	gct Ala	gac Asp	tat Tyr	480
gtt Val	gga Gly	atg Met	ggg Gly	gca Ala	gtg Val	ttt Phe	cat His	aca Thr	agc Ser	act Thr	aaa Lys	aaa Lys	gat Asp	gca Ala	aag Lys	528
gat Asp	atg Met	agc Ser	agg Arg	gaa Glu	aca Thr	ctt Leu	tta Leu	gag Glu	ctt Leu	gca Ala	ggg Gly	atg Met	atg Met	gag Glu	gat Asp	576
att Ile	ccg Pro	gtg Val	gtc Val	gcc Ala	att Ile	ggc Gly	ggc Gly	ata Ile	agc Ser	tat Tyr	gat Asp	aac Asn	tgc Cys	gat Asp	tac Tyr	624
tta Leu	aag Lys	gac Asp	aca Thr	ggt Gly	gtt Val	gat Asp	gga Gly	ata Ile	gca Ala	gtt Val	gtt Val	tca Ser	gcc Ala	ata Ile	ttt Phe	672
gca Ala	agt Ser	gat Asp	gac Asp	tgt Cys	gcg Ala	ctt Leu	gcc Ala	aca Thr	aga Arg	aag Lys	ctt Leu	ttt Phe	gta Val	aag Lys	aca Thr	720
agg Arg	gaa Glu	ttg Leu	ttt Phe	gga Gly	aag Lys	aaa Lys	aga Arg	aac Asn	ata Ile	ata Ile	atg Met	gat Asp	atg Met	gat Asp	ggt Gly	768
acg Thr	ctt Leu	gca Ala	gac Asp	tct Ser	atg Met	cct Pro	ttc Phe	tgg Trp	aaa Lys	aac Asn	agc Ser	gca Ala	aga Arg	gag Glu	tat Tyr	816
gcg Ala	ata Ile	tta Leu	cgt Arg	gga Gly	gca Ala	gat Asp	att Ile	ccg Pro	gat Asp	aat Asn	ttc Phe	gat Asp	gag Glu	ata Ile	act Thr	864
ggc Gly	gtt Val	atg Met	gac Asp	ctt Leu	aat Asn	gat Asp	tat Tyr	gct Ala	gag Glu	tat Tyr	gtt Val	aaa Lys	aat Asn	gtt Val	ctt Leu	912
ggc Gly	ata Ile	gat Asp	act Thr	aat Asn	ctt Leu	gag Glu	cag Gln	ata Ile	aca Thr	gaa Glu	gcg Ala	gct Ala	gtc Val	gag Glu	att Ile	960

ES 2 806 732 T3

atg aat aaa cat tac gaa aaa gat ata cct gca aag gac ggt atg aca	1008
Met Asn Lys His Tyr Glu Lys Asp Ile Pro Ala Lys Asp Gly Met Thr	
325 330 335	
gag ctt gtc acg aga gaa tat aag gcc gga agc aga ctt gtt gtg ttt	1056
Glu Leu Val Thr Arg Glu Tyr Lys Ala Gly Ser Arg Leu Val Val Phe	
340 345 350	
acg gct tca gat aga aga agt gtt gaa att ctt ctt tca cac ctt gga	1104
Thr Ala Ser Asp Arg Arg Ser Val Glu Ile Leu Leu Ser His Leu Gly	
355 360 365	
ata aga gaa tgt ttt tat gat ata tat aca gtc tat gat gta gga ctt	1152
Ile Arg Glu Cys Phe Tyr Asp Ile Tyr Thr Val Tyr Asp Val Gly Leu	
370 375 380	
aag aag agt gat aag aac agc tat ctt aag gtg gca gag ctt gca ggc	1200
Lys Lys Ser Asp Lys Asn Ser Tyr Leu Lys Val Ala Glu Leu Ala Gly	
385 390 395 400	
atg aaa gat aca tca cag gta tgg gta tat gag gat ata tta aga ggt	1248
Met Lys Asp Thr Ser Gln Val Trp Val Tyr Glu Asp Ile Leu Arg Gly	
405 410 415	
gta aag gca gcg aaa gag gcc gga ctt aat gtg tgt gca gtg tat gat	1296
Val Lys Ala Ala Lys Glu Ala Gly Leu Asn Val Cys Ala Val Tyr Asp	
420 425 430	
gag gac tcc gca ggc gac tgg gag gac ata aaa gag ctt gcg gat aag	1344
Glu Asp Ser Ala Gly Asp Trp Glu Asp Ile Lys Glu Leu Ala Asp Lys	
435 440 445	
acc ctt gaa ctt gtg taa	1362
Thr Leu Glu Leu Val	
450	

<210> 50

<211> 453

5 <212> PRT

<213> Especie *Eubacterium*

<400> 50

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

10

ES 2 806 732 T3

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

ES 2 806 732 T3

Glu Leu Val Thr Arg Glu Tyr Lys Ala Gly Ser Arg Leu Val Val Phe
340 345 350

Thr Ala Ser Asp Arg Arg Ser Val Glu Ile Leu Leu Ser His Leu Gly
355 360 365

Ile Arg Glu Cys Phe Tyr Asp Ile Tyr Thr Val Tyr Asp Val Gly Leu
370 375 380

Lys Lys Ser Asp Lys Asn Ser Tyr Leu Lys Val Ala Glu Leu Ala Gly
385 390 395 400

Met Lys Asp Thr Ser Gln Val Trp Val Tyr Glu Asp Ile Leu Arg Gly
405 410 415

Val Lys Ala Ala Lys Glu Ala Gly Leu Asn Val Cys Ala Val Tyr Asp
420 425 430

Glu Asp Ser Ala Gly Asp Trp Glu Asp Ile Lys Glu Leu Ala Asp Lys
435 440 445

Thr Leu Glu Leu Val
450

<210> 51

<211> 1260

5 <212> ADN

<213> *Lachnospiraceae pectinoschiza*

<220>

<221> CDS

10 <222> (1)..(1260)

<223> gen de *Lachnospiraceae pectinoschiza* que codifica para TMP fosfatasa [CUQ76318]

<400> 51

atg aaa gtt acc cgt gaa gat atg cag ctt tac gcc gtt aca gat acg 48
Met Lys Val Thr Arg Glu Asp Met Gln Leu Tyr Ala Val Thr Asp Thr
1 5 10 15

caa tgg ctt aat ggc agg gat ttc tat gaa gag att gag aaa gtc ctt 96
Gln Trp Leu Asn Gly Arg Asp Phe Tyr Glu Glu Ile Glu Lys Val Leu
20 25 30

gcg gca gga gct aca ttt ttg cag tta aga gaa aag gat tcg aca cac 144
Ala Ala Gly Ala Thr Phe Leu Gln Leu Arg Glu Lys Asp Ser Thr His
35 40 45

gag gag att gta aaa aaa gca ttg gca att aaa ccg ata gca aga aga 192
Glu Glu Ile Val Lys Lys Ala Leu Ala Ile Lys Pro Ile Ala Arg Arg
50 55 60

15

ES 2 806 732 T3

tat ggt gtg cca ttt gtt ata gat gat gat ata tac gcg gcg tta gag Tyr Gly Val Pro Phe Val Ile Asp Asp Asp Ile Tyr Ala Ala Leu Glu 65 70 75 80	240
gca gat gtt gac gga gtt cat ata gga caa agt gat gca agc tac gaa Ala Asp Val Asp Gly Val His Ile Gly Gln Ser Asp Ala Ser Tyr Glu 85 90 95	288
aca gca aga gag ctt cta gga cct gac aag ata ata gga atg aca gta Thr Ala Arg Glu Leu Leu Gly Pro Asp Lys Ile Ile Gly Met Thr Val 100 105 110	336
aag aca cca gag cag gcg gca aat gcg gca aga ctt ggt gct gat tat Lys Thr Pro Glu Gln Ala Ala Asn Ala Ala Arg Leu Gly Ala Asp Tyr 115 120 125	384
gtt gga atg gga gct gta ttt cat aca agc acg aag aaa gat gcc aaa Val Gly Met Gly Ala Val Phe His Thr Ser Thr Lys Lys Asp Ala Lys 130 135 140	432
gat tta agc agg gat aat ctt ctt aag ctt aca gct atg ctt gat atg Asp Leu Ser Arg Asp Asn Leu Leu Lys Leu Thr Ala Met Leu Asp Met 145 150 155 160	480
ccg ata gtt gca att ggc ggc att aat tat gac aac tgt gat tat tta Pro Ile Val Ala Ile Gly Gly Ile Asn Tyr Asp Asn Cys Asp Tyr Leu 165 170 175	528
aaa gat aca ggc gtg gac gga att gct gtt gta tcg gcg ata ttt gca Lys Asp Thr Gly Val Asp Gly Ile Ala Val Val Ser Ala Ile Phe Ala 180 185 190	576
agt gat gac tgc gcg gag gcg aca cga aag ctt tat aag aag aca aga Ser Asp Asp Cys Ala Glu Ala Thr Arg Lys Leu Tyr Lys Lys Thr Arg 195 200 205	624
aag ctg ttt aat tat aat aag aac ata ata ttt gat atg gac gga aca Lys Leu Phe Asn Tyr Asn Lys Asn Ile Ile Phe Asp Met Asp Gly Thr 210 215 220	672
ctt gtt gac tct atg ccg ttc tgg aag aat agt gca agg gaa tat gcc Leu Val Asp Ser Met Pro Phe Trp Lys Asn Ser Ala Arg Glu Tyr Ala 225 230 235 240	720
att tta aga ggt gct aag ctt cca aag aat ttt gat gag ata aca gga Ile Leu Arg Gly Ala Lys Leu Pro Lys Asn Phe Asp Glu Ile Thr Gly 245 250 255	768
gtt atg gac ctt tcg gaa tat gcg gct tat ctg caa aat gtt ctt ggg Val Met Asp Leu Ser Glu Tyr Ala Ala Tyr Leu Gln Asn Val Leu Gly 260 265 270	816
att gat aca tcg cta gaa cag ata aca gag gca gca gtt gat att atg Ile Asp Thr Ser Leu Glu Gln Ile Thr Glu Ala Ala Val Asp Ile Met 275 280 285	864
aat aag cat tat gca agt gat att cct gca aag aag gga atg ata aag Asn Lys His Tyr Ala Ser Asp Ile Pro Ala Lys Lys Gly Met Ile Lys 290 295 300	912
ctt ata aga aga gaa tat gag gct gga agc aag ctt gta ata ttc agt Leu Ile Arg Arg Glu Tyr Glu Ala Gly Ser Lys Leu Val Ile Phe Ser 305 310 315 320	960

ES 2 806 732 T3

gct tcc gat act tcc agt gtg gaa att ctt ctt aaa agg tta gaa ata 1008
 Ala Ser Asp Thr Ser Ser Val Glu Ile Leu Leu Lys Arg Leu Glu Ile
 325 330 335

tat gaa tgt ttt gag gga ata tac aca gta tat gat gtc ggc ata gga 1056
 Tyr Glu Cys Phe Glu Gly Ile Tyr Thr Val Tyr Asp Val Gly Ile Gly
 340 345 350

aag agt gat aag gaa agc tat aaa aag gtt gcc agg tca gca gga atg 1104
 Lys Ser Asp Lys Glu Ser Tyr Lys Lys Val Ala Arg Ser Ala Gly Met
 355 360 365

gat ata tct gat acg tgg gtg tat gag gat att cta aga ggc gtt cgg 1152
 Asp Ile Ser Asp Thr Trp Val Tyr Glu Asp Ile Leu Arg Gly Val Arg
 370 375 380

gcg gca cat aat gct gga ttg aaa gtg tgt gcg gta tat gat aaa gac 1200
 Ala Ala His Asn Ala Gly Leu Lys Val Cys Ala Val Tyr Asp Lys Asp
 385 390 395 400

tcg gca gat gac tgg gat gag ata tgc agt att gca gat aaa tgt ata 1248
 Ser Ala Asp Asp Trp Asp Glu Ile Cys Ser Ile Ala Asp Lys Cys Ile
 405 410 415

ata acc gga taa 1260
 Ile Thr Gly

<210> 52

<211> 419

5 <212> PRT

<213> *Lachnospiraceae pectinoschiza*

<400> 52

Met Lys Val Thr Arg Glu Asp Met Gln Leu Tyr Ala Val Thr Asp Thr
 1 5 10 15

Gln Trp Leu Asn Gly Arg Asp Phe Tyr Glu Glu Ile Glu Lys Val Leu
 20 25 30

Ala Ala Gly Ala Thr Phe Leu Gln Leu Arg Glu Lys Asp Ser Thr His
 35 40 45

Glu Glu Ile Val Lys Lys Ala Leu Ala Ile Lys Pro Ile Ala Arg Arg
 50 55 60

Tyr Gly Val Pro Phe Val Ile Asp Asp Asp Ile Tyr Ala Ala Leu Glu
 65 70 75 80

Ala Asp Val Asp Gly Val His Ile Gly Gln Ser Asp Ala Ser Tyr Glu
 85 90 95

10 Thr Ala Arg Glu Leu Leu Gly Pro Asp Lys Ile Ile Gly Met Thr Val
 100 105 110

ES 2 806 732 T3

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

ES 2 806 732 T3

355 360 365

Asp Ile Ser Asp Thr Trp Val Tyr Glu Asp Ile Leu Arg Gly Val Arg
370 375 380

Ala Ala His Asn Ala Gly Leu Lys Val Cys Ala Val Tyr Asp Lys Asp
385 390 395 400

Ser Ala Asp Asp Trp Asp Glu Ile Cys Ser Ile Ala Asp Lys Cys Ile
405 410 415

Ile Thr Gly

<210> 53

<211> 1296

5 <212> ADN

<213> Bacteria *Peptostreptococcaceae*

<220>

<221> CDS

10 <222> (1)..(1296)

<223> gen de bacteria *Peptostreptococcaceae* OBRC8 que codifica para TMP fosfatasa [WP_009530263]

<400> 53

atg aaa aat att gac tat aca atg tat tac gtc acc gat gaa gac ctt	48
Met Lys Asn Ile Asp Tyr Thr Met Tyr Tyr Val Thr Asp Glu Asp Leu	
1 5 10 15	
ttg agc agt aat cat acc ttg gaa aca tct gta caa gat gcc att tta	96
Leu Ser Ser Asn His Thr Leu Glu Thr Ser Val Gln Asp Ala Ile Leu	
20 25 30	
ggg ggc tgt aca atg ata cag ctt cga gaa aaa cat tca tcc act ctc	144
Gly Gly Cys Thr Met Ile Gln Leu Arg Glu Lys His Ser Ser Thr Leu	
35 40 45	
gat ttt tat aac aaa gcc ata aaa att aaa gcc att tgc gac aag tac	192
Asp Phe Tyr Asn Lys Ala Ile Lys Ile Lys Ala Ile Cys Asp Lys Tyr	
50 55 60	
aac ata cct ctt ata ata aat gac aga ata gat gta gct ctt gca ata	240
Asn Ile Pro Leu Ile Ile Asn Asp Arg Ile Asp Val Ala Leu Ala Ile	
65 70 75 80	
aac gca gac gga gta cat ctc gga caa gac gat atg cct ctt gat att	288
Asn Ala Asp Gly Val His Leu Gly Gln Asp Asp Met Pro Leu Asp Ile	
85 90 95	
gca aga aaa att atg gga gat ggc aaa att ata gga ata tca act gca	336
Ala Arg Lys Ile Met Gly Asp Gly Lys Ile Ile Gly Ile Ser Thr Ala	
100 105 110	
act tta gat gaa gct cta atc gct caa caa ggc ggt gca gat tat gta	384
Thr Leu Asp Glu Ala Leu Ile Ala Gln Gln Gly Gly Ala Asp Tyr Val	
115 120 125	

15

ES 2 806 732 T3

gga gta ggt gct atg tac agc aca aac aca aaa acc gat gcc aat ttg Gly Val Gly Ala Met Tyr Ser Thr Asn Thr Lys Thr Asp Ala Asn Leu 130 135 140	432
aca act ata aac gag ctt aca aaa ata aaa aac aat cta aaa ata cct Thr Thr Ile Asn Glu Leu Thr Lys Ile Lys Asn Asn Leu Lys Ile Pro 145 150 155 160	480
gta gtt gca atc ggc ggt ata aac ctt gac aca ata cct gct cta aaa Val Val Ala Ile Gly Gly Ile Asn Leu Asp Thr Ile Pro Ala Leu Lys 165 170 175	528
cct gca caa ata gac gga gtt gca ata gta tcc gct ata tct atg cag Pro Ala Gln Ile Asp Gly Val Ala Ile Val Ser Ala Ile Ser Met Gln 180 185 190	576
gaa gat acc gta tct gca aca aga aaa tta aaa aat act ttt ttg aaa Glu Asp Thr Val Ser Ala Thr Arg Lys Leu Lys Asn Thr Phe Leu Lys 195 200 205	624
caa tat caa act aaa ggc gta ata ttc gat att gac ggt act ctg ctt Gln Tyr Gln Thr Lys Gly Val Ile Phe Asp Ile Asp Gly Thr Leu Leu 210 215 220	672
gaa act atg aac ata tgg gac aat gta ctt cta aac ctt atg aat aca Glu Thr Met Asn Ile Trp Asp Asn Val Leu Leu Asn Leu Met Asn Thr 225 230 235 240	720
ctt aat atc agc tat acc gaa gat gaa ata caa aaa ata tgg aat atg Leu Asn Ile Ser Tyr Thr Glu Asp Glu Ile Gln Lys Ile Trp Asn Met 245 250 255	768
ggt ttt gca gag ctt gca cag ttc agc ata aaa aaa ttc aag ctt gat Gly Phe Ala Glu Leu Ala Gln Phe Ser Ile Lys Lys Phe Lys Leu Asp 260 265 270	816
atg agt gta aaa gaa ttt tgg caa ctt ata aaa aaa tta tca gtc gaa Met Ser Val Lys Glu Phe Trp Gln Leu Ile Lys Lys Leu Ser Val Glu 275 280 285	864
gag tat aaa aat agc aaa ata cac tta aaa aaa ggt gca aaa aaa ctg Glu Tyr Lys Asn Ser Lys Ile His Leu Lys Lys Gly Ala Lys Lys Leu 290 295 300	912
ctt gag tat ctc aaa gaa aaa ggt gta aaa tta gcc ata gca act gcc Leu Glu Tyr Leu Lys Glu Lys Gly Val Lys Leu Ala Ile Ala Thr Ala 305 310 315 320	960
ctt tgc aaa gaa cag tat gaa ata gtg ctt aca aag aca ggt atc ata Leu Cys Lys Glu Gln Tyr Glu Ile Val Leu Thr Lys Thr Gly Ile Ile 325 330 335	1008
gac tat ttt gac ata ata gca tca agc gta gat tta aaa atg gaa aaa Asp Tyr Phe Asp Ile Ile Ala Ser Val Asp Leu Lys Met Glu Lys 340 345 350	1056
tca gac aga caa ata ttt gac tat ata gca aaa aat cta caa gtt cca Ser Asp Arg Gln Ile Phe Asp Tyr Ile Ala Lys Asn Leu Gln Val Pro 355 360 365	1104
aac aaa aat ctt att ttc ttt gaa gac gac ata aac tcg tca aca ggt Asn Lys Asn Leu Ile Phe Phe Glu Asp Asp Ile Asn Ser Ser Thr Gly 1152	1152

ES 2 806 732 T3

	370		375			380	
	gcc aag ttg gca gga cta aaa ctg tgc att gta tca aac aag aaa tat						1200
	Ala Lys Leu Ala Gly Leu Lys Leu Cys Ile Val Ser Asn Lys Lys Tyr						
	385		390			395	400
	aac ggt aac agc aaa ttt gac gct ctc ata gat tat aaa ata gat gat						1248
	Asn Gly Asn Ser Lys Phe Asp Ala Leu Ile Asp Tyr Lys Ile Asp Asp						
		405		410		415	
	ttt gaa aat aaa ttg ata tat gat gaa ata ata gtg gag aaa aat tag						1296
	Phe Glu Asn Lys Leu Ile Tyr Asp Glu Ile Ile Val Glu Lys Asn						
		420		425		430	
	<210> 54						
	<211> 431						
5	<212> PRT						
	<213> Bacteria <i>Peptostreptococcaceae</i>						
	<400> 54						
	Met Lys Asn Ile Asp Tyr Thr Met Tyr Tyr Val Thr Asp Glu Asp Leu						
	1	5		10		15	
	Leu Ser Ser Asn His Thr Leu Glu Thr Ser Val Gln Asp Ala Ile Leu						
		20		25		30	
	Gly Gly Cys Thr Met Ile Gln Leu Arg Glu Lys His Ser Ser Thr Leu						
		35		40		45	
	Asp Phe Tyr Asn Lys Ala Ile Lys Ile Lys Ala Ile Cys Asp Lys Tyr						
		50		55		60	
	Asn Ile Pro Leu Ile Ile Asn Asp Arg Ile Asp Val Ala Leu Ala Ile						
	65	70		75		80	
	Asn Ala Asp Gly Val His Leu Gly Gln Asp Asp Met Pro Leu Asp Ile						
		85		90		95	
	Ala Arg Lys Ile Met Gly Asp Gly Lys Ile Ile Gly Ile Ser Thr Ala						
		100		105		110	
	Thr Leu Asp Glu Ala Leu Ile Ala Gln Gln Gly Gly Ala Asp Tyr Val						
		115		120		125	
	Gly Val Gly Ala Met Tyr Ser Thr Asn Thr Lys Thr Asp Ala Asn Leu						
		130		135		140	
	Thr Thr Ile Asn Glu Leu Thr Lys Ile Lys Asn Asn Leu Lys Ile Pro						
	145	150		155		160	
10	Val Val Ala Ile Gly Gly Ile Asn Leu Asp Thr Ile Pro Ala Leu Lys						

ES 2 806 732 T3

165	170	175
Pro Ala Gln Ile Asp Gly Val Ala Ile Val Ser Ala Ile Ser Met Gln		
180	185	190
Glu Asp Thr Val Ser Ala Thr Arg Lys Leu Lys Asn Thr Phe Leu Lys		
195	200	205
Gln Tyr Gln Thr Lys Gly Val Ile Phe Asp Ile Asp Gly Thr Leu Leu		
210	215	220
Glu Thr Met Asn Ile Trp Asp Asn Val Leu Leu Asn Leu Met Asn Thr		
225	230	235
Leu Asn Ile Ser Tyr Thr Glu Asp Glu Ile Gln Lys Ile Trp Asn Met		
245	250	255
Gly Phe Ala Glu Leu Ala Gln Phe Ser Ile Lys Lys Phe Lys Leu Asp		
260	265	270
Met Ser Val Lys Glu Phe Trp Gln Leu Ile Lys Lys Leu Ser Val Glu		
275	280	285
Glu Tyr Lys Asn Ser Lys Ile His Leu Lys Lys Gly Ala Lys Lys Leu		
290	295	300
Leu Glu Tyr Leu Lys Glu Lys Gly Val Lys Leu Ala Ile Ala Thr Ala		
305	310	315
Leu Cys Lys Glu Gln Tyr Glu Ile Val Leu Thr Lys Thr Gly Ile Ile		
325	330	335
Asp Tyr Phe Asp Ile Ile Ala Ser Ser Val Asp Leu Lys Met Glu Lys		
340	345	350
Ser Asp Arg Gln Ile Phe Asp Tyr Ile Ala Lys Asn Leu Gln Val Pro		
355	360	365
Asn Lys Asn Leu Ile Phe Phe Glu Asp Asp Ile Asn Ser Ser Thr Gly		
370	375	380
Ala Lys Leu Ala Gly Leu Lys Leu Cys Ile Val Ser Asn Lys Lys Tyr		
385	390	395
Asn Gly Asn Ser Lys Phe Asp Ala Leu Ile Asp Tyr Lys Ile Asp Asp		
405	410	415
Phe Glu Asn Lys Leu Ile Tyr Asp Glu Ile Ile Val Glu Lys Asn		
420	425	430

<210> 55
 5 <211> 1296
 <212> ADN
 <213> Bacteria *Peptostreptococcaceae*

<220>
 10 <221> CDS
 <222> (1)..(1296)

<223> gen de bacteria *Peptostreptococcaceae* CM2 que codifica para TMP fosfatasa [WP_009527854]

<400> 55

atg aaa aat att gac tat aca atg tat tac gtc acc gat gaa gac ctt	48
Met Lys Asn Ile Asp Tyr Thr Met Tyr Tyr Val Thr Asp Glu Asp Leu	
1 5 10 15	
ttg agc agt aat cac acc ttg gaa aca tct gtg caa gat gcc att tta	96
Leu Ser Ser Asn His Thr Leu Glu Thr Ser Val Gln Asp Ala Ile Leu	
20 25 30	
ggt ggc tgt aca atg ata cag ctt cga gaa aaa cat tca tcc act ctc	144
Gly Gly Cys Thr Met Ile Gln Leu Arg Glu Lys His Ser Ser Thr Leu	
35 40 45	
gat ttt tat aac aaa gcc gta aaa att aaa gct att tgc gac aag tac	192
Asp Phe Tyr Asn Lys Ala Val Lys Ile Lys Ala Ile Cys Asp Lys Tyr	
50 55 60	
aac ata cct ctt ata ata aat gac aga ata gac gta gct ctt gca ata	240
Asn Ile Pro Leu Ile Ile Asn Asp Arg Ile Asp Val Ala Leu Ala Ile	
65 70 75 80	
aat gca gac gga gta cat ctc gga caa gac gat atg cct ctt gat att	288
Asn Ala Asp Gly Val His Leu Gly Gln Asp Asp Met Pro Leu Asp Ile	
85 90 95	
gca aga aaa att atg gga gat ggc aaa att ata gga ata tca acc tca	336
Ala Arg Lys Ile Met Gly Asp Gly Lys Ile Ile Gly Ile Ser Thr Ser	
100 105 110	
act tta gat gaa gct cta atc gct caa caa ggc ggt gca gat tat gta	384
Thr Leu Asp Glu Ala Leu Ile Ala Gln Gln Gly Gly Ala Asp Tyr Val	
115 120 125	
ggt gta ggt gct atg tac agc aca aac aca aaa act gat gcc aat ttg	432
Gly Val Gly Ala Met Tyr Ser Thr Asn Thr Lys Thr Asp Ala Asn Leu	
130 135 140	
aca act ata gac gag ctt aca aaa ata aaa aac aat tta aaa ata cct	480
Thr Thr Ile Asp Glu Leu Thr Lys Ile Lys Asn Asn Leu Lys Ile Pro	
145 150 155 160	
gtt gtt gca atc ggc ggt ata aac ctt gac act ata ccc gct cta aaa	528
Val Val Ala Ile Gly Gly Ile Asn Leu Asp Thr Ile Pro Ala Leu Lys	
165 170 175	
cct gcg caa ata gac gga gtt gca ata gta tcc gct ata tct atg cag	576
Pro Ala Gln Ile Asp Gly Val Ala Ile Val Ser Ala Ile Ser Met Gln	

5

ES 2 806 732 T3

180	185	190	
gaa gat acc gta tct gca aca aga aaa tta aaa aat act ttt ttg aaa Glu Asp Thr Val Ser Ala Thr Arg Lys Leu Lys Asn Thr Phe Leu Lys 195 200 205			624
caa tat caa act aaa ggc gta ata ttc gat att gac ggt act ctg ctt Gln Tyr Gln Thr Lys Gly Val Ile Phe Asp Ile Asp Gly Thr Leu Leu 210 215 220			672
gaa act atg aac ata tgg gac aat gta ctt cta aat ctt atg aat acg Glu Thr Met Asn Ile Trp Asp Asn Val Leu Leu Asn Leu Met Asn Thr 225 230 235 240			720
ctt aat atc cgc tat acc gaa gat gaa ata caa aag ata tgg aat atg Leu Asn Ile Arg Tyr Thr Glu Asp Glu Ile Gln Lys Ile Trp Asn Met 245 250 255			768
ggt ttt gca gag ctt gca cag ttc agc ata aaa aaa ttc aag ctt gat Gly Phe Ala Glu Leu Ala Gln Phe Ser Ile Lys Lys Phe Lys Leu Asp 260 265 270			816
atg agt gta aaa gaa ttt tgg caa ctt ata aaa aaa tta tca gtc gaa Met Ser Val Lys Glu Phe Trp Gln Leu Ile Lys Lys Leu Ser Val Glu 275 280 285			864
gag tat aaa aat agc aaa ata cac tta aaa aaa ggt gca aaa aaa ctg Glu Tyr Lys Asn Ser Lys Ile His Leu Lys Lys Gly Ala Lys Lys Leu 290 295 300			912
ctt gag tat ctc aaa gaa aaa ggt gta aaa tta gcc ata gca act gcc Leu Glu Tyr Leu Lys Glu Lys Gly Val Lys Leu Ala Ile Ala Thr Ala 305 310 315 320			960
ctt tgc aaa gaa cag tat gaa ata gtg ctt aca aag aca ggt atc ata Leu Cys Lys Glu Gln Tyr Glu Ile Val Leu Thr Lys Thr Gly Ile Ile 325 330 335			1008
gac tat ttt gac ata ata gca tca agc gta gat tta aaa atg gaa aaa Asp Tyr Phe Asp Ile Ile Ala Ser Ser Val Asp Leu Lys Met Glu Lys 340 345 350			1056
tca gat aga caa ata ttt gac tat ata gca aaa aat cta caa gtt cca Ser Asp Arg Gln Ile Phe Asp Tyr Ile Ala Lys Asn Leu Gln Val Pro 355 360 365			1104
aac aaa aat ttt att ttc ttt gaa gac gac ata aac tcg tca aca ggt Asn Lys Asn Phe Ile Phe Phe Glu Asp Asp Ile Asn Ser Ser Thr Gly 370 375 380			1152
gca aaa cgt gca gga gta aaa ctg tgc att gta tca aac aag aaa tat Ala Lys Arg Ala Gly Val Lys Leu Cys Ile Val Ser Asn Lys Lys Tyr 385 390 395 400			1200
aat ggt aac agc aaa ttt gac gct ctc ata gat tat aaa ata gat gat Asn Gly Asn Ser Lys Phe Asp Ala Leu Ile Asp Tyr Lys Ile Asp Asp 405 410 415			1248
ttt gaa aat aaa ttg ata tat gat gaa ata ata gtg gag aaa aat tag Phe Glu Asn Lys Leu Ile Tyr Asp Glu Ile Ile Val Glu Lys Asn 420 425 430			1296

<210> 56

<211> 431

5 <212> PRT

<213> Bacteria *Peptostreptococcaceae*

<400> 56

ES 2 806 732 T3

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

ES 2 806 732 T3

Glu Thr Met Asn Ile Trp Asp Asn Val Leu Leu Asn Leu Met Asn Thr
225 230 235 240

Leu Asn Ile Arg Tyr Thr Glu Asp Glu Ile Gln Lys Ile Trp Asn Met
245 250 255

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

Met Ser Val Lys Glu Phe Trp Gln Leu Ile Lys Lys Leu Ser Val Glu
275 280 285

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

Leu Glu Tyr Leu Lys Glu Lys Gly Val Lys Leu Ala Ile Ala Thr Ala
305 310 315 320

Leu Cys Lys Glu Gln Tyr Glu Ile Val Leu Thr Lys Thr Gly Ile Ile
325 330 335

Asp Tyr Phe Asp Ile Ile Ala Ser Ser Val Asp Leu Lys Met Glu Lys
340 345 350

Ser Asp Arg Gln Ile Phe Asp Tyr Ile Ala Lys Asn Leu Gln Val Pro
355 360 365

Asn Lys Asn Phe Ile Phe Phe Glu Asp Asp Ile Asn Ser Ser Thr Gly
370 375 380

Ala Lys Arg Ala Gly Val Lys Leu Cys Ile Val Ser Asn Lys Lys Tyr
385 390 395 400

Asn Gly Asn Ser Lys Phe Asp Ala Leu Ile Asp Tyr Lys Ile Asp Asp
405 410 415

Phe Glu Asn Lys Leu Ile Tyr Asp Glu Ile Ile Val Glu Lys Asn
420 425 430

<210> 57

<211> 1365

5 <212> ADN

<213> Especie *Atopobium*

<220>

<221> CDS

10 <222> (1)..(1365)

<223> gen de *Atopobium* sp. ICM42b que codifica para TMP fosfatasa [WP_035427744]

<400> 57

ES 2 806 732 T3

atg cag gtg acc ggt gca att ttt gat tgc gat gga act ctt gtt gat Met Gln Val Thr Gly Ala Ile Phe Asp Cys Asp Gly Thr Leu Val Asp 1 5 10 15	48
tca atg cgc gtt tgg cat aac gtt ttt ggc gct gtt ctt cct aaa tat Ser Met Arg Val Trp His Asn Val Phe Gly Ala Val Leu Pro Lys Tyr 20 25 30	96
ggc aag act att gat tgc gat att ttt gac cgc gta gag gct gtt tcc Gly Lys Thr Ile Asp Ser Asp Ile Phe Asp Arg Val Glu Ala Val Ser 35 40 45	144
ctc att ggt gga tgt cag att tgc gtt gat gaa ctg gat ttg cct att Leu Ile Gly Gly Cys Gln Ile Cys Val Asp Glu Leu Asp Leu Pro Ile 50 55 60	192
aca gcg gaa gct ttg tat gaa gag ttc tgc gcg tac gta att gat cag Thr Ala Glu Ala Leu Tyr Glu Glu Phe Cys Ala Tyr Val Ile Asp Gln 65 70 75 80	240
tac caa cat cat gtt tca atc att ccc ggt gca aag gag ttc tta cag Tyr Gln His His Val Ser Ile Ile Pro Gly Ala Lys Glu Phe Leu Gln 85 90 95	288
gag ctc tac gat gca ggt att cct atg gcc gtt gct tgc tca act ccc Glu Leu Tyr Asp Ala Gly Ile Pro Met Ala Val Ala Ser Ser Thr Pro 100 105 110	336
gtg cga gaa gtt cgt gca gct ctg gca gct caa ggt att gag cac ctc Val Arg Glu Val Arg Ala Ala Leu Ala Ala Gln Gly Ile Glu His Leu 115 120 125	384
ttc aaa aca gtg gtc tca aca gaa gat gtg ggg gga gtg gac aag gtt Phe Lys Thr Val Val Ser Thr Glu Asp Val Gly Gly Val Asp Lys Val 130 135 140	432
gag cct gat gtt tat ctt gag gct ctt cgc cgt ctt ggc acc gat aag Glu Pro Asp Val Tyr Leu Glu Ala Leu Arg Arg Leu Gly Thr Asp Lys 145 150 155 160	480
gca act acc tgg gtc ttc gag gat gcc ccg ttt ggc gca cag aca gca Ala Thr Thr Trp Val Phe Glu Asp Ala Pro Phe Gly Ala Gln Thr Ala 165 170 175	528
caa aat gcg ggc ttt cct gtg gta gcg ctc tac aat gat cat gac ggc Gln Asn Ala Gly Phe Pro Val Val Ala Leu Tyr Asn Asp His Asp Gly 180 185 190	576
cgc gac ccc gtc ttt atg cgc gag cac tct aac atc ttt gcc cac acc Arg Asp Pro Val Phe Met Arg Glu His Ser Asn Ile Phe Ala His Thr 195 200 205	624
tac ggc gag ctg tgc ctt ctg cgc ctt cag gac tac gag cgc cct ctg Tyr Gly Glu Leu Ser Leu Leu Arg Leu Gln Asp Tyr Glu Arg Pro Leu 210 215 220	672
acc gca gcg cct tct ggc gag aaa ccc ctt gag gtc ctt gtt gtg ggc Thr Ala Ala Pro Ser Gly Glu Lys Pro Leu Glu Val Leu Val Val Gly 225 230 235 240	720
gga tcc cca gag gcg gtt tca cac acg acg ctg tct acc tgc gcc caa Gly Ser Pro Glu Ala Val Ser His Thr Thr Leu Ser Thr Cys Ala Gln	768

ES 2 806 732 T3

	245	250	255	
	agc gct gac tac ctg ata gcg gtt gac cat ggt gca gat gca tgt cac			816
	Ser Ala Asp Tyr Leu Ile Ala Val Asp His Gly Ala Asp Ala Cys His			
	260	265	270	
	gct gcc ggc gtg att cca cag ctt gcg ctt gga gac ttt gac tcg gct			864
	Ala Ala Gly Val Ile Pro Gln Leu Ala Leu Gly Asp Phe Asp Ser Ala			
	275	280	285	
	aca cca gaa act ctg gct tgg ctc aaa gag cag cag gta cct tgc atg			912
	Thr Pro Glu Thr Leu Ala Trp Leu Lys Glu Gln Gln Val Pro Cys Met			
	290	295	300	
	aag ttt aat gcg gac aag tac gat acc gac ctg gct ctt gct tta aag			960
	Lys Phe Asn Ala Asp Lys Tyr Asp Thr Asp Leu Ala Leu Ala Leu Lys			
	305	310	315	320
	tcc gcc gag tac gag gct att cgt aga gat agc aag ctc tct ctt acg			1008
	Ser Ala Glu Tyr Glu Ala Ile Arg Arg Asp Ser Lys Leu Ser Leu Thr			
	325	330	335	
	gtt gtc tcc aca tct ggc gga cac ctt gat cac cag ctt gta gtg ctt			1056
	Val Val Ser Thr Ser Gly Gly His Leu Asp His Gln Leu Val Val Leu			
	340	345	350	
	ggc ctt ctc gcc acg tgg gca aag acg ggc aag gca agt gtt cga gtt			1104
	Gly Leu Leu Ala Thr Trp Ala Lys Thr Gly Lys Ala Ser Val Arg Val			
	355	360	365	
	att gaa aat gac ttt gag atg cgc ttt tta act gca ggc cag gtt gat			1152
	Ile Glu Asn Asp Phe Glu Met Arg Phe Leu Thr Ala Gly Gln Val Asp			
	370	375	380	
	tct tgg cag ctg agc gca act gat gta ggt aaa aag atg tcc ctt gtg			1200
	Ser Trp Gln Leu Ser Ala Thr Asp Val Gly Lys Lys Met Ser Leu Val			
	385	390	395	400
	gct ttg tca gag gag tgc gag gtt tct gag gcc ggc atg aag tgg aat			1248
	Ala Leu Ser Glu Glu Cys Glu Val Ser Glu Ala Gly Met Lys Trp Asn			
	405	410	415	
	ctt gat cac cag aag ttc acc ttg ctg gga gac gac ggt att tca aat			1296
	Leu Asp His Gln Lys Phe Thr Leu Leu Gly Asp Asp Gly Ile Ser Asn			
	420	425	430	
	atc gtc gaa tca gac aat tcc tgg gta agg tgc gag aag ggc tgt ctt			1344
	Ile Val Glu Ser Asp Asn Ser Trp Val Arg Cys Glu Lys Gly Cys Leu			
	435	440	445	
	ttg gtg cag ctt tgg aac taa			1365
	Leu Val Gln Leu Trp Asn			
	450			

<210> 58

<211> 454

5 <212> PRT

<213> Especie *Atopobium*

<400> 58

10 Met Gln Val Thr Gly Ala Ile Phe Asp Cys Asp Gly Thr Leu Val Asp

ES 2 806 732 T3

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

ES 2 806 732 T3

Ser Ala Asp Tyr Leu Ile Ala Val Asp His Gly Ala Asp Ala Cys His
260 265 270

Ala Ala Gly Val Ile Pro Gln Leu Ala Leu Gly Asp Phe Asp Ser Ala
275 280 285

Thr Pro Glu Thr Leu Ala Trp Leu Lys Glu Gln Gln Val Pro Cys Met
290 295 300

Lys Phe Asn Ala Asp Lys Tyr Asp Thr Asp Leu Ala Leu Ala Leu Lys
305 310 315 320

Ser Ala Glu Tyr Glu Ala Ile Arg Arg Asp Ser Lys Leu Ser Leu Thr
325 330 335

Val Val Ser Thr Ser Gly Gly His Leu Asp His Gln Leu Val Val Leu
340 345 350

Gly Leu Leu Ala Thr Trp Ala Lys Thr Gly Lys Ala Ser Val Arg Val
355 360 365

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

Ser Trp Gln Leu Ser Ala Thr Asp Val Gly Lys Lys Met Ser Leu Val
385 390 395 400

Ala Leu Ser Glu Glu Cys Glu Val Ser Glu Ala Gly Met Lys Trp Asn
405 410 415

Leu Asp His Gln Lys Phe Thr Leu Leu Gly Asp Asp Gly Ile Ser Asn
420 425 430

Ile Val Glu Ser Asp Asn Ser Trp Val Arg Cys Glu Lys Gly Cys Leu
435 440 445

Leu Val Gln Leu Trp Asn
450

<210> 59

<211> 1365

5 <212> ADN

<213> *Atopobium parvulum*

<220>

<221> CDS

10 <222> (1)..(1365)

<223> gen de *Atopobium parvulum* que codifica para TMP fosfatasa [WP_035433109]

<400> 59

ES 2 806 732 T3

atg cag gtg acc ggt gca att ttt gat tgc gat gga act ctt gtt gat Met Gln Val Thr Gly Ala Ile Phe Asp Cys Asp Gly Thr Leu Val Asp 1 5 10 15	48
tca atg cac gtt tgg cac aac gtt ttt ggc gct gtt ctt cct aaa tac Ser Met His Val Trp His Asn Val Phe Gly Ala Val Leu Pro Lys Tyr 20 25 30	96
ggc aag act att gat tgc gat att ttt gac cgc gta gag gct gtt tcc Gly Lys Thr Ile Asp Ser Asp Ile Phe Asp Arg Val Glu Ala Val Ser 35 40 45	144
ctc att ggt gga tgt cag att tgc gtt gat gag ctg gat ttg cct att Leu Ile Gly Gly Cys Gln Ile Cys Val Asp Glu Leu Asp Leu Pro Ile 50 55 60	192
aca gcg gaa gct tta tat gaa gag ttc tgc gcg tac gta act gat cag Thr Ala Glu Ala Leu Tyr Glu Glu Phe Cys Ala Tyr Val Thr Asp Gln 65 70 75 80	240
tac cga cat cat gtt tca atc att ccc ggt gca aag gag ttc tta cag Tyr Arg His His Val Ser Ile Ile Pro Gly Ala Lys Glu Phe Leu Gln 85 90 95	288
gaa ctc cac gac gca ggc att cct atg gcc gtt gct tgc tca act ccc Glu Leu His Asp Ala Gly Ile Pro Met Ala Val Ala Ser Ser Thr Pro 100 105 110	336
gtg cga gaa gtt cgt gca gct ctg gca gct caa ggt att gag cac ctc Val Arg Glu Val Arg Ala Ala Leu Ala Ala Gln Gly Ile Glu His Leu 115 120 125	384
ttt aaa aca gtg gtc tca acg gaa gat gtg ggg gga gtg gac aag gtt Phe Lys Thr Val Val Ser Thr Glu Asp Val Gly Gly Val Asp Lys Val 130 135 140	432
gag cca gat gtt tac ctt gag gct ctt cgc cgt ctt ggc act gat aag Glu Pro Asp Val Tyr Leu Glu Ala Leu Arg Arg Leu Gly Thr Asp Lys 145 150 155 160	480
gca act acc tgg gtc ttc gag gat gct ccg ttt ggc gca cag aca gca Ala Thr Thr Trp Phe Glu Asp Ala Pro Phe Gly Ala Gln Thr Ala 165 170 175	528
caa aat gca ggc ttt cct gtg gct gta ctc tac aac gac cac gat ggc Gln Asn Ala Gly Phe Pro Val Ala Val Leu Tyr Asn Asp His Asp Gly 180 185 190	576
cgc gac ccc gtc ttt atg cgc gag cac tct aac atc ttt gcc cac acc Arg Asp Pro Val Phe Met Arg Glu His Ser Asn Ile Phe Ala His Thr 195 200 205	624
tac ggc gag ctg tgc ctt ctg cgc ctt cag gac tac gag cgc cct ctg Tyr Gly Glu Leu Ser Leu Leu Arg Leu Gln Asp Tyr Glu Arg Pro Leu 210 215 220	672
acc gca gcg cct tct ggc gag aaa ccc ctt gag gtc ctt gtt gtg ggc Thr Ala Ala Pro Ser Gly Glu Lys Pro Leu Glu Val Leu Val Val Gly 225 230 235 240	720
gga tcc cca gag gcg gtt tgc cac acg acg ctg tct acc tgc gcc caa Gly Ser Pro Glu Ala Val Ser His Thr Thr Leu Ser Thr Cys Ala Gln	768

ES 2 806 732 T3

	245	250	255	
	agc gct gac tac ctg ata gcg gtt gac cat ggc gca gat gtc tgt cac			816
	Ser Ala Asp Tyr Leu Ile Ala Val Asp His Gly Ala Asp Val Cys His			
	260	265	270	
	gct gcc ggc gtg att cca caa ctt gcg ctt gga gac ttt gac tcc gct			864
	Ala Ala Gly Val Ile Pro Gln Leu Ala Leu Gly Asp Phe Asp Ser Ala			
	275	280	285	
	aca cca gaa act ctg gct tgg ctc aaa gag cag cag gta cct tgc atg			912
	Thr Pro Glu Thr Leu Ala Trp Leu Lys Glu Gln Gln Val Pro Cys Met			
	290	295	300	
	aag ttt aat gcg gac aag tac gat acc gac ctg gcg cta gca ttg aaa			960
	Lys Phe Asn Ala Asp Lys Tyr Asp Thr Asp Leu Ala Leu Ala Leu Lys			
	305	310	315	320
	tca gct gaa tat gag gca att cgt aga gat agc aag ctc tct ctg acg			1008
	Ser Ala Glu Tyr Glu Ala Ile Arg Arg Asp Ser Lys Leu Ser Leu Thr			
	325	330	335	
	gtt gtc tcc aca tct ggc ggc cac ctt gat cac cag ctt gta gtg ctt			1056
	Val Val Ser Thr Ser Gly Gly His Leu Asp His Gln Leu Val Val Leu			
	340	345	350	
	ggc ctt ctc gcc acg tgg gca aag acg ggc aag gca agc gtt cga gtt			1104
	Gly Leu Leu Ala Thr Trp Ala Lys Thr Gly Lys Ala Ser Val Arg Val			
	355	360	365	
	att gag aat gac ttt gag atg cgc ttt tta gtt gct ggc cag gtg gat			1152
	Ile Glu Asn Asp Phe Glu Met Arg Phe Leu Val Ala Gly Gln Val Asp			
	370	375	380	
	tct tgg cag ctg aac act atc aat gta ggt aaa aag att tct ctt gta			1200
	Ser Trp Gln Leu Asn Thr Ile Asn Val Gly Lys Lys Ile Ser Leu Val			
	385	390	395	400
	gct ttg tca gag gag tgc gag gtt tct gag gcc ggc atg aag tgg aat			1248
	Ala Leu Ser Glu Glu Cys Glu Val Ser Glu Ala Gly Met Lys Trp Asn			
	405	410	415	
	ctt gat cac cag aag ttc acc ttg ctg gga gac gac ggt att tca aac			1296
	Leu Asp His Gln Lys Phe Thr Leu Leu Gly Asp Asp Gly Ile Ser Asn			
	420	425	430	
	ata gtt gaa tca gac aat tcc tgg gta agg tgc gag aag ggc tgt ctt			1344
	Ile Val Glu Ser Asp Asn Ser Trp Val Arg Cys Glu Lys Gly Cys Leu			
	435	440	445	
	ttg gtg cag ctt tgg aac taa			1365
	Leu Val Gln Leu Trp Asn			
	450			

<210> 60

<211> 454

5 <212> PRT

<213> *Atopobium parvulum*

<400> 60

10 Met Gln Val Thr Gly Ala Ile Phe Asp Cys Asp Gly Thr Leu Val Asp

ES 2 806 732 T3

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

ES 2 806 732 T3

Ser Ala Asp Tyr Leu Ile Ala Val Asp His Gly Ala Asp Val Cys His
260 265 270

Ala Ala Gly Val Ile Pro Gln Leu Ala Leu Gly Asp Phe Asp Ser Ala
275 280 285

Thr Pro Glu Thr Leu Ala Trp Leu Lys Glu Gln Gln Val Pro Cys Met
290 295 300

Lys Phe Asn Ala Asp Lys Tyr Asp Thr Asp Leu Ala Leu Ala Leu Lys
305 310 315 320

Ser Ala Glu Tyr Glu Ala Ile Arg Arg Asp Ser Lys Leu Ser Leu Thr
325 330 335

Val Val Ser Thr Ser Gly Gly His Leu Asp His Gln Leu Val Val Leu
340 345 350

Gly Leu Leu Ala Thr Trp Ala Lys Thr Gly Lys Ala Ser Val Arg Val
355 360 365

Ile Glu Asn Asp Phe Glu Met Arg Phe Leu Val Ala Gly Gln Val Asp
370 375 380

Ser Trp Gln Leu Asn Thr Ile Asn Val Gly Lys Lys Ile Ser Leu Val
385 390 395 400

Ala Leu Ser Glu Glu Cys Glu Val Ser Glu Ala Gly Met Lys Trp Asn
405 410 415

Leu Asp His Gln Lys Phe Thr Leu Leu Gly Asp Asp Gly Ile Ser Asn
420 425 430

Ile Val Glu Ser Asp Asn Ser Trp Val Arg Cys Glu Lys Gly Cys Leu
435 440 445

Leu Val Gln Leu Trp Asn
450

<210> 61

<211> 1383

5 <212> ADN

<213> *Atopobium rimae*

<220>

<221> CDS

10 <222> (1)..(1383)

<223> gen de *Atopobium rimae* que codifica para TMP fosfatasa [WP_003148415]

<400> 61

ES 2 806 732 T3

atg cag ata acg ggt gca atc ttt gat ctt gat ggg aca ctg gtt gac Met Gln Ile Thr Gly Ala Ile Phe Asp Leu Asp Gly Thr Leu Val Asp 1 5 10 15	48
tcc atg tgg atg tgg aga aga tcg ttc gga gat gtt tta gaa gac ctg Ser Met Trp Met Trp Arg Arg Ser Phe Gly Asp Val Leu Glu Asp Leu 20 25 30	96
cat atc aat atg act ccg gat ttt ttt aaa agg gtc gag gcc att tcg His Ile Asn Met Thr Pro Asp Phe Phe Lys Arg Val Glu Ala Ile Ser 35 40 45	144
ctt tac gat ggt tgc gta gcg tgt att gag gaa ttt aat ctt cct tta Leu Tyr Asp Gly Cys Val Ala Cys Ile Glu Glu Phe Asn Leu Pro Leu 50 55 60	192
tcc gca gaa gag ctg tat gaa aag ttc ctt ttg tat gta caa acg gta Ser Ala Glu Glu Leu Tyr Glu Lys Phe Leu Leu Tyr Val Gln Thr Val 65 70 75 80	240
tat tcg cac gat att aaa agc att gcg ggg gct acc gac ttt ctc cag Tyr Ser His Asp Ile Lys Ser Ile Ala Gly Ala Thr Asp Phe Leu Gln 85 90 95	288
gaa ctt ttt gac gca gga ata cct ctt gcc att gct tct tct acg cca Glu Leu Phe Asp Ala Gly Ile Pro Leu Ala Ile Ala Ser Ser Thr Pro 100 105 110	336
tct cgt gcc ata cat gtt gct ctt gaa gcc caa ggt atg gag aag ttt Ser Arg Ala Ile His Val Ala Leu Glu Ala Gln Gly Met Glu Lys Phe 115 120 125	384
ttt aaa gcg gtt gtg tgt acc gaa gac gtc ggg ggt gtc gat aaa gca Phe Lys Ala Val Val Cys Thr Glu Asp Val Gly Gly Val Asp Lys Ala 130 135 140	432
aaa ccc gat gtc tat ctt gag gct ctc aga cgc ctg ggc acc gat aaa Lys Pro Asp Val Tyr Leu Glu Ala Leu Arg Arg Leu Gly Thr Asp Lys 145 150 155 160	480
gca cac acg tgg gtc ttt gag gac gct gag ttt ggt gta cat acg gca Ala His Thr Trp Val Phe Glu Asp Ala Glu Phe Gly Val His Thr Ala 165 170 175	528
caa acc gag ggc ttt ccc gtt gtt gcg ctg ttc aat ggc aaa gac ggc Gln Thr Glu Gly Phe Pro Val Val Ala Leu Phe Asn Gly Lys Asp Gly 180 185 190	576
cgt gat ctt gag tat atg aag gcg cac tct gat ctt ctc gca cat gat Arg Asp Leu Glu Tyr Met Lys Ala His Ser Asp Leu Leu Ala His Asp 195 200 205	624
tat cga gaa ctc tct ctt gcc cgc att tac gat tat gaa cgg gtg acg Tyr Arg Glu Leu Ser Leu Ala Arg Ile Tyr Asp Tyr Glu Arg Val Thr 210 215 220	672
aat cag cca cat ctg ggc gcc tca tcg gct cag aag gtc ttt tcg gtt Asn Gln Pro His Leu Gly Ala Ser Ser Ala Gln Lys Val Phe Ser Val 225 230 235 240	720
ctc gtt gtt gat gga tct ccc acg cca agt tca gcc gcg ctg gtt tca Leu Val Val Asp Gly Ser Pro Thr Pro Ser Ser Ala Ala Leu Val Ser	768

ES 2 806 732 T3

	245	250	255	
	gaa ctt tca tca tgc tcg gat tat gtc gtt gct gca gat cgc ggg gca	816		
	Glu Leu Ser Ser Cys Ser Asp Tyr Val Val Ala Ala Asp Arg Gly Ala			
	260 265 270			
	tat atc tgc aag gag gcc ggt gtc gtt cct gat att gcg tgc gga gac	864		
	Tyr Ile Cys Lys Glu Ala Gly Val Val Pro Asp Ile Ala Cys Gly Asp			
	275 280 285			
	ttt gat tcc gtg gga gaa gag aca ctc tct tgg atc cat gca caa aag	912		
	Phe Asp Ser Val Gly Glu Glu Thr Leu Ser Trp Ile His Ala Gln Lys			
	290 295 300			
	gtg cac acg att gct tat cct caa gat aag tac gag acc gat ttg tct	960		
	Val His Thr Ile Ala Tyr Pro Gln Asp Lys Tyr Glu Thr Asp Leu Ser			
	305 310 315 320			
	ctt gca ctc aat gcc gct tgc cat gaa gca acc cgt caa gca ctt ccg	1008		
	Leu Ala Leu Asn Ala Ala Cys His Glu Ala Thr Arg Gln Ala Leu Pro			
	325 330 335			
	ctg tca ctg aca ctt acc tgc gct tcc ggc ggc agg ctt gat cat gag	1056		
	Leu Ser Leu Thr Leu Thr Cys Ala Ser Gly Gly Arg Leu Asp His Glu			
	340 345 350			
	ctt ggt gta gtg ggg ctt ctg gct cga tta agc act gcc tca tgg agg	1104		
	Leu Gly Val Val Gly Leu Leu Ala Arg Leu Ser Thr Ala Ser Trp Arg			
	355 360 365			
	gtg cgg att gtt gag gat gcc ttt gaa gca agg att ctt tcg gca gat	1152		
	Val Arg Ile Val Glu Asp Ala Phe Glu Ala Arg Ile Leu Ser Ala Asp			
	370 375 380			
	acg tat gcg gcg tgg agg ctc tca gaa aaa gat cga gga aag aca ctg	1200		
	Thr Tyr Ala Ala Trp Arg Leu Ser Glu Lys Asp Arg Gly Lys Thr Leu			
	385 390 395 400			
	tcg gtg ctt ccg ctt cag gaa gaa acg gtg att acc gag atc ggt atg	1248		
	Ser Val Leu Pro Leu Gln Glu Glu Thr Val Ile Thr Glu Ile Gly Met			
	405 410 415			
	caa tgg gac ctt gcc tca cga act ttg ctg ctc ctg tct gat gaa gga	1296		
	Gln Trp Asp Leu Ala Ser Arg Thr Leu Leu Leu Leu Ser Asp Glu Gly			
	420 425 430			
	att tcc aat gtg gta caa acg gat gtg gca caa ata cat tgc gag aag	1344		
	Ile Ser Asn Val Val Gln Thr Asp Val Ala Gln Ile His Cys Glu Lys			
	435 440 445			
	ggc aag gcg ctc gtg gtg ctt ctc gca aat gaa tcg tga	1383		
	Gly Lys Ala Leu Val Val Leu Leu Ala Asn Glu Ser			
	450 455 460			
5	<210> 62			
	<211> 460			
	<212> PRT			
	<213> <i>Atopobium rimae</i>			
	<400> 62			
10	Met Gln Ile Thr Gly Ala Ile Phe Asp Leu Asp Gly Thr Leu Val Asp			

ES 2 806 732 T3

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

ES 2 806 732 T3

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

<210> 63

<211> 1380

5 <212> ADN

<213> *Olsenella uli*

<220>

<221> CDS

10 <222> (1)..(1380)

<223> gen de *Olsenella uli* que codifica para TMP fosfatasa [WP_013251930]

<400> 63

ES 2 806 732 T3

atg ccc atc aag gcc gcc atc ttc gac tgt gac gga acg ctg gtc gac Met Pro Ile Lys Ala Ala Ile Phe Asp Cys Asp Gly Thr Leu Val Asp 1 5 10 15	48
tcc atg ccc ctg tgg cat gac gtg acg gtc gaa ctg ctg cgc cgc cac Ser Met Pro Leu Trp His Asp Val Thr Val Glu Leu Leu Arg Arg His 20 25 30	96
cat gtc gcc gac gcc gag gag gcg ttc gtc cgc acc gag tcg ctt ccc His Val Ala Asp Ala Glu Glu Ala Phe Val Arg Thr Glu Ser Leu Pro 35 40 45	144
atg gtc gag atg tgc cat gcc ttc cac gac gag tgg ggc gtt gag gcc Met Val Glu Met Cys His Ala Phe His Asp Glu Trp Gly Val Glu Ala 50 55 60	192
gag ggc gag gag ctg gtg cgc gag ctg gtc gat atg gtc cgc gag ggg Glu Gly Glu Glu Leu Val Arg Glu Leu Val Asp Met Val Arg Glu Gly 65 70 75 80	240
tat cgc agc cgg gtt agc ctg ctg ccg ggc tgc cgg gcg ttt ctg gac Tyr Arg Ser Arg Val Ser Leu Leu Pro Gly Cys Arg Ala Phe Leu Asp 85 90 95	288
gag ctg gcg tct gcg ggc gtc cgc atg gtc gtc gcg tcg tcg acg gct Glu Leu Ala Ser Ala Gly Val Arg Met Val Val Ala Ser Ser Thr Ala 100 105 110	336
ccg gag gag ctc tcc gtc gcg cta tcg gcg cag ggg gtc gac ggc tac Pro Glu Glu Leu Ser Val Ala Leu Ser Ala Gln Gly Val Asp Gly Tyr 115 120 125	384
ttc gag cgg gtc ttc tcc acg gga ggc ccc ata cgc agc aag gac tac Phe Glu Arg Val Phe Ser Thr Gly Gly Pro Ile Arg Ser Lys Asp Tyr 130 135 140	432
ccg gac atc tgg gag ctg gtc ctg gac tac ctg ggc acc gac ccg gct Pro Asp Ile Trp Glu Leu Val Leu Asp Tyr Leu Gly Thr Asp Pro Ala 145 150 155 160	480
gac acc tgg gtc ttc gag gac gcc ccg ttt ggg atg cgg acg gcc cga Asp Thr Trp Val Phe Glu Asp Ala Pro Phe Gly Met Arg Thr Ala Arg 165 170 175	528
tcg gtc ggc gcc aac acc gtc tgc ctg ttc agc cca cac ggg gac cgc Ser Val Gly Ala Asn Thr Val Cys Leu Phe Ser Pro His Gly Asp Arg 180 185 190	576
gac ctt gcg gcc tgc gag cgc tac gct gac ata ctg gtc cac agc tac Asp Leu Ala Ala Cys Glu Arg Tyr Ala Asp Ile Leu Val His Ser Tyr 195 200 205	624
cac gag cta tcg ctc gcc ctg ctg gac gac tac gcc cgt ccg ccg caa His Glu Leu Ser Leu Ala Leu Leu Asp Asp Tyr Ala Arg Pro Pro Gln 210 215 220	672
gcg tcc ccc tcg gcc cac cct cgc ctc gcg ccg ctt cgc gtc ctc gtc Ala Ser Pro Ser Ala His Pro Arg Leu Ala Pro Leu Arg Val Leu Val 225 230 235 240	720
gtg ggc gcc tcg ccc gag cgc ccg tct tcg gcg ctg ctc cgc tcc ctg Val Gly Ala Ser Pro Glu Arg Pro Ser Ser Ala Leu Leu Arg Ser Leu 245 250 255 260	768

ES 2 806 732 T3

	245	250	255	
	gcc gcc agt acc gac tac gtc atc gcc gcc gac gcc ggg gcc gac gcg Ala Ala Ser Thr Asp Tyr Val Ile Ala Ala Asp Ala Gly Ala Asp Ala 260 265 270			816
	ctg cgc tcc tgt ggc atc gcc ccc gac gtc ttc tgc ggc gac gcc gac Leu Arg Ser Cys Gly Ile Ala Pro Asp Val Phe Cys Gly Asp Ala Asp 275 280 285			864
	tcg gca acg ggc gaa tcg gct gcg tgg gcc cgc tcg gtc gcc cgt gcg Ser Ala Thr Gly Glu Ser Ala Ala Trp Ala Arg Ser Val Ala Arg Ala 290 295 300			912
	gac ata gag ttt ccc tcc gag aag tac gcg acc gac ctc gcc ctc gcc Asp Ile Glu Phe Pro Ser Glu Lys Tyr Ala Thr Asp Leu Ala Leu Ala 305 310 315 320			960
	atc tcc tgc gcc cgc cat gag gcc gct cga cgc aac gcg cgg ctg gag Ile Ser Cys Ala Arg His Glu Ala Ala Arg Arg Asn Ala Arg Leu Glu 325 330 335			1008
	ctc acg ctg acc ggc gtc acg ggc ggc agg ccc gac cac gcc ctt gcc Leu Thr Leu Thr Gly Val Thr Gly Arg Pro Asp His Ala Leu Ala 340 345 350			1056
	gtc gtg ggt cag ctc gcg cgg aac gct gac gcc tcg ccg cgc atc gtg Val Val Gly Gln Leu Ala Arg Asn Ala Asp Ala Ser Pro Arg Ile Val 355 360 365			1104
	gag gac ggc ttc gag tgc cga ctg ctc agc ccc tct ggc act gcg tgc Glu Asp Gly Phe Glu Cys Arg Leu Leu Ser Pro Ser Gly Thr Ala Cys 370 375 380			1152
	tgg gag ctg ggt ggg gcc cac gtg cca gcc gcc ggg gtc gag ggg acg Trp Glu Leu Gly Gly Ala His Val Pro Ala Ala Gly Val Glu Gly Thr 385 390 395 400			1200
	ctc ttc tcg gcc att ccc gtg gca gag ggg acc atg ctc tcc gag cgg Leu Phe Ser Ala Ile Pro Val Ala Glu Gly Thr Met Leu Ser Glu Arg 405 410 415			1248
	ggc ttc aag tgg gag ctg gat cat cgt gag ctg ccc ctt ctg ggg gat Gly Phe Lys Trp Glu Leu Asp His Arg Glu Leu Pro Leu Leu Gly Asp 420 425 430			1296
	gag gga atc tcg aac gtg gtc acg tcc gcg acg gcc agc gtc gag tgc Glu Gly Ile Ser Asn Val Val Thr Ser Ala Thr Ala Ser Val Glu Cys 435 440 445			1344
	cat gcc ggc gca gtt gcg gcg ttc ctg ttg gca tag His Ala Gly Ala Val Ala Ala Phe Leu Leu Ala 450 455			1380

<210> 64
 <211> 459
 <212> PRT
 <213> *Olsenella uli*

<400> 64

Met Pro Ile Lys Ala Ala Ile Phe Asp Cys Asp Gly Thr Leu Val Asp

ES 2 806 732 T3

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

ES 2 806 732 T3

Ala Ala Ser Thr Asp Tyr Val Ile Ala Ala Asp Ala Gly Ala Asp Ala
260 265 270

Leu Arg Ser Cys Gly Ile Ala Pro Asp Val Phe Cys Gly Asp Ala Asp
275 280 285

Ser Ala Thr Gly Glu Ser Ala Ala Trp Ala Arg Ser Val Ala Arg Ala
290 295 300

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

Ile Ser Cys Ala Arg His Glu Ala Ala Arg Arg Asn Ala Arg Leu Glu
325 330 335

Leu Thr Leu Thr Gly Val Thr Gly Gly Arg Pro Asp His Ala Leu Ala
340 345 350

Val Val Gly Gln Leu Ala Arg Asn Ala Asp Ala Ser Pro Arg Ile Val
355 360 365

Glu Asp Gly Phe Glu Cys Arg Leu Leu Ser Pro Ser Gly Thr Ala Cys
370 375 380

Trp Glu Leu Gly Gly Ala His Val Pro Ala Ala Gly Val Glu Gly Thr
385 390 395 400

Leu Phe Ser Ala Ile Pro Val Ala Glu Gly Thr Met Leu Ser Glu Arg
405 410 415

Gly Phe Lys Trp Glu Leu Asp His Arg Glu Leu Pro Leu Leu Gly Asp
420 425 430

Glu Gly Ile Ser Asn Val Val Thr Ser Ala Thr Ala Ser Val Glu Cys
435 440 445

His Ala Gly Ala Val Ala Ala Phe Leu Leu Ala
450 455

<210> 65

<211> 762

5 <212> ADN

<213> *Atopobium minutum*

<220>

<221> CDS

10 <222> (1)..(762)

<223> gen de *Atopobium minutum* que codifica para TMP fosfatasa [KRN55115]

<400> 65

ES 2 806 732 T3

atg tgg gct aaa acc tct cga cat tgt acg caa aaa ggc ttt acc atg	48
Met Trp Ala Lys Thr Ser Arg His Cys Thr Gln Lys Gly Phe Thr Met	
1 5 10 15	
aac cct gca cgc att tta ttt gat gga gga act tgt atg gca ata agc	96
Asn Pro Ala Arg Ile Leu Phe Asp Gly Gly Thr Cys Met Ala Ile Ser	
20 25 30	
ggc gca atc ttt gac tgt gac ggc acg ctg gtt gat tct atg tat atg	144
Gly Ala Ile Phe Asp Cys Asp Gly Thr Leu Val Asp Ser Met Tyr Met	
35 40 45	
tgg tgg gac gcc ttt ccc cgc ctg ctt gcc agc cat ggc ttt gct atg	192
Trp Trp Asp Ala Phe Pro Arg Leu Leu Ala Ser His Gly Phe Ala Met	
50 55 60	
acg cct cag atc gag aaa atc ttg cat gag tgt gag gcg gtc agc ttg	240
Thr Pro Gln Ile Glu Lys Ile Leu His Glu Cys Glu Ala Val Ser Leu	
65 70 75 80	
gat gaa gag atc cat acg ctg cgc aac gct ctt gct att ccc gct tct	288
Asp Glu Glu Ile His Thr Leu Arg Asn Ala Leu Ala Ile Pro Ala Ser	
85 90 95	
gcc gag cag cta gca caa gaa tta tcc cag aat att agc aat gcg tat	336
Ala Glu Gln Leu Ala Gln Glu Leu Ser Gln Asn Ile Ser Asn Ala Tyr	
100 105 110	
gcc tca gag atc aaa gca tgg cct gcc gtt aag ccg ttc ttg gat cag	384
Ala Ser Glu Ile Lys Ala Trp Pro Ala Val Lys Pro Phe Leu Asp Gln	
115 120 125	
ctc aaa gac gca ggt atc ccc atg atc att tgt act tct acc gga gcc	432
Leu Lys Asp Ala Gly Ile Pro Met Ile Ile Cys Thr Ser Thr Gly Ala	
130 135 140	
aaa gaa gtt ggt ctg tgc atg gat cat ctt ggt ttg tcc aag ttt ttt	480
Lys Glu Val Gly Leu Cys Met Asp His Leu Gly Leu Ser Lys Phe Phe	
145 150 155 160	
gta gat att gtc agc gcg gaa gaa aac aat ttc acc aaa act gag cca	528
Val Asp Ile Val Ser Ala Glu Glu Asn Asn Phe Thr Lys Thr Glu Pro	
165 170 175	
gat atc tat tac tat gcg cta aaa aag ctt ggt acc act aaa gag aca	576
Asp Ile Tyr Tyr Tyr Ala Leu Lys Lys Leu Gly Thr Thr Lys Glu Thr	
180 185 190	
acc tgg gta ttt gag gat gct ccg ttt ggc ctt act acc tct gag cgt	624
Thr Trp Val Phe Glu Asp Ala Pro Phe Gly Leu Thr Ser Glu Arg	
195 200 205	
gca gga ttt cct aat gtg tgc gtc ttt aat gcg cac gat aag cgc gat	672
Ala Gly Phe Pro Asn Val Cys Val Phe Asn Ala His Asp Lys Arg Asp	
210 215 220	
gag gac ttt ttg cgt ctt cat gct acg ttg ttt acg cac ata tat gag	720
Glu Asp Phe Leu Arg Leu His Ala Thr Leu Phe Thr His Ile Tyr Glu	
225 230 235 240	
gat att tcc ctt gcg gat ttg cag tcg tac ccc acc aag taa	762
Asp Ile Ser Leu Ala Asp Leu Gln Ser Tyr Pro Thr Lys	
245 250	

<210> 66

5 <211> 253

<212> PRT

<213> *Atopobium minutum*

<400> 66

10

ES 2 806 732 T3

```

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

Asn Pro Ala Arg Ile Leu Phe Asp Gly Gly Thr Cys Met Ala Ile Ser
          20          25          30

Gly Ala Ile Phe Asp Cys Asp Gly Thr Leu Val Asp Ser Met Tyr Met
          35          40          45

Trp Trp Asp Ala Phe Pro Arg Leu Leu Ala Ser His Gly Phe Ala Met
          50          55          60

Thr Pro Gln Ile Glu Lys Ile Leu His Glu Cys Glu Ala Val Ser Leu
65          70          75          80

Asp Glu Glu Ile His Thr Leu Arg Asn Ala Leu Ala Ile Pro Ala Ser
          85          90          95

Ala Glu Gln Leu Ala Gln Glu Leu Ser Gln Asn Ile Ser Asn Ala Tyr
          100          105          110

Ala Ser Glu Ile Lys Ala Trp Pro Ala Val Lys Pro Phe Leu Asp Gln
          115          120          125

Leu Lys Asp Ala Gly Ile Pro Met Ile Ile Cys Thr Ser Thr Gly Ala
130          135          140

Lys Glu Val Gly Leu Cys Met Asp His Leu Gly Leu Ser Lys Phe Phe
145          150          155          160

Val Asp Ile Val Ser Ala Glu Glu Asn Asn Phe Thr Lys Thr Glu Pro
          165          170          175

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

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

Ala Gly Phe Pro Asn Val Cys Val Phe Asn Ala His Asp Lys Arg Asp
          210          215          220

Glu Asp Phe Leu Arg Leu His Ala Thr Leu Phe Thr His Ile Tyr Glu
225          230          235          240

Asp Ile Ser Leu Ala Asp Leu Gln Ser Tyr Pro Thr Lys
          245          250

```

<210> 67

5 <211> 648

<212> ADN

<213> *Syntrophomonas wolfei*

<220>

10 <221> CDS

<222> (1)..(648)

<223> gen de *Syntrophomonas wolfei* que codifica para TMP fosfatasa [WP_011640074]

<400> 67

atg gga gag aaa tta ata att ttt atg gat ttc gat ggc act att tct	48
Met Gly Glu Lys Leu Ile Ile Phe Met Asp Phe Asp Gly Thr Ile Ser	
1 5 10 15	
cgg gag gat gtc tgc aat aag atg gca gcc agg tat gcc ggc agg gac	96
Arg Glu Asp Val Cys Asn Lys Met Ala Ala Arg Tyr Ala Gly Arg Asp	
20 25 30	
tgg gag gaa ata aac cgc ctc tgg gaa gag gga ggt att act act gga	144
Trp Glu Glu Ile Asn Arg Leu Trp Glu Glu Gly Gly Ile Thr Thr Gly	
35 40 45	
gag tgc gcc agt cgt att ctt tca tca atg gag gta ggg gcg gct gaa	192
Glu Cys Ala Ser Arg Ile Leu Ser Ser Met Glu Val Gly Ala Ala Glu	
50 55 60	
ttg gag gcc ttt ttt cag gct cag gaa gta gac ccc ggc ttt tcc cct	240
Leu Glu Ala Phe Phe Gln Ala Gln Glu Val Asp Pro Gly Phe Ser Pro	
65 70 75 80	
ttc ctg gac tgg gta caa aaa aat cag cac ctc ccc att ata ttg agc	288
Phe Leu Asp Trp Val Gln Lys Asn Gln His Leu Pro Ile Ile Leu Ser	
85 90 95	
gat ggt tat gac cgc tat ata aaa agc ata tta cgg ggc cag ggc tgg	336
Asp Gly Tyr Asp Arg Tyr Ile Lys Ser Ile Leu Arg Gly Gln Gly Trp	
100 105 110	
gaa atc gag ttt tat gcc aat aaa tta tac tgg gat gac gcc tgg cgg	384
Glu Ile Glu Phe Tyr Ala Asn Lys Leu Tyr Trp Asp Asp Ala Trp Arg	
115 120 125	
atg gaa tcg ccc tac ctg gat gaa gaa tgc ttt aaa tgt ggg gta tgc	432
Met Glu Ser Pro Tyr Leu Asp Glu Glu Cys Phe Lys Cys Gly Val Cys	
130 135 140	
aag agc aag ata atc cag gaa aga agt tta ccc ggc tat ctc aca gta	480
Lys Ser Lys Ile Ile Gln Glu Arg Ser Leu Pro Gly Tyr Leu Thr Val	
145 150 155 160	
tat atc gga gat ggc tac tcc gat ttc tgc ccg gcg gcc tct tgt gat	528
Tyr Ile Gly Asp Gly Tyr Ser Asp Phe Cys Pro Ala Ala Ser Cys Asp	
165 170 175	
att gtt ttt gcc aaa aat gaa ctg gcc ggc tac tgc cag aaa gag ggt	576
Ile Val Phe Ala Lys Asn Glu Leu Ala Gly Tyr Cys Gln Lys Glu Gly	
180 185 190	
tta act tac tac ccc tac cgg gat ttt cac gat att ctc cag caa ctg	624
Leu Thr Tyr Tyr Pro Tyr Arg Asp Phe His Asp Ile Leu Gln Gln Leu	
195 200 205	
ccg agg att gtt agc agg atg tag	648
Pro Arg Ile Val Ser Arg Met	
210 215	

<210> 68

<211> 215

<212> PRT

<213> *Syntrophomonas wolfei*

<400> 68

ES 2 806 732 T3

Met Gly Glu Lys Leu Ile Ile Phe Met Asp Phe Asp Gly Thr Ile Ser
1 5 10 15

Arg Glu Asp Val Cys Asn Lys Met Ala Ala Arg Tyr Ala Gly Arg Asp
20 25 30

Trp Glu Glu Ile Asn Arg Leu Trp Glu Glu Gly Gly Ile Thr Thr Gly
35 40 45

Glu Cys Ala Ser Arg Ile Leu Ser Ser Met Glu Val Gly Ala Ala Glu
50 55 60

Leu Glu Ala Phe Phe Gln Ala Gln Glu Val Asp Pro Gly Phe Ser Pro
65 70 75 80

Phe Leu Asp Trp Val Gln Lys Asn Gln His Leu Pro Ile Ile Leu Ser
85 90 95

Asp Gly Tyr Asp Arg Tyr Ile Lys Ser Ile Leu Arg Gly Gln Gly Trp
100 105 110

Glu Ile Glu Phe Tyr Ala Asn Lys Leu Tyr Trp Asp Asp Ala Trp Arg
115 120 125

Met Glu Ser Pro Tyr Leu Asp Glu Glu Cys Phe Lys Cys Gly Val Cys
130 135 140

Lys Ser Lys Ile Ile Gln Glu Arg Ser Leu Pro Gly Tyr Leu Thr Val
145 150 155 160
Tyr Ile Gly Asp Gly Tyr Ser Asp Phe Cys Pro Ala Ala Ser Cys Asp
165 170 175

Ile Val Phe Ala Lys Asn Glu Leu Ala Gly Tyr Cys Gln Lys Glu Gly
180 185 190

Leu Thr Tyr Tyr Pro Tyr Arg Asp Phe His Asp Ile Leu Gln Gln Leu
195 200 205

Pro Arg Ile Val Ser Arg Met
210 215

<210> 69

5 <211> 666

<212> ADN

<213> *Desulfitobacterium hafniense*

<220>

10 <221> CDS

<222> (1)..(666)

<223> gen de *Desulfitobacterium hafniense* que codifica para TMP fosfatasa [WP_018212876]

<400> 69

15

ES 2 806 732 T3

```

atg gag gaa ttg aac agt att ttc ttc gtg gat ttt gac ggc acc atc      48
Met Glu Glu Leu Asn Ser Ile Phe Phe Val Asp Phe Asp Gly Thr Ile
1          5          10          15

gtc act cag gat atg tgt gca gtc ctc gtt gaa acc ttg gcc ggg gaa      96
Val Thr Gln Asp Met Cys Ala Val Leu Val Glu Thr Leu Ala Gly Glu
          20          25          30

gga tgg cgg gag att aat gaa ctt tgg gaa aga aaa gag ctt tcc acc     144
Gly Trp Arg Glu Ile Asn Glu Leu Trp Glu Arg Lys Glu Leu Ser Thr
          35          40          45

ctg gag tgc gcc cgc cgg acc ttt aaa ctc ttt aac agc aat gac ccg     192
Leu Glu Cys Ala Arg Arg Phe Lys Leu Phe Asn Ser Asn Asp Pro
          50          55          60

gaa gtt ttt cgc cag ctt atc ggg cag gcg gtg ttc gat ccc gga ttt     240
Glu Val Phe Arg Gln Leu Ile Gly Gln Ala Val Phe Asp Pro Gly Phe
65          70          75          80

tta gat ttt gcc gct ttt tgt gaa cag aga gga ttt ccc ctc atc att     288
Leu Asp Phe Ala Ala Phe Cys Glu Gln Arg Gly Phe Pro Leu Ile Ile
          85          90          95

ctc agc gac gga tat gat ttc tat att gag tac ctc ttg caa aga gag     336
Leu Ser Asp Gly Tyr Asp Phe Tyr Ile Glu Tyr Leu Leu Gln Arg Glu
          100          105          110

gga ttg aac ctg cca tac tat gcc aac aaa ttg ctg ttt gct ccc caa     384
Gly Leu Asn Leu Pro Tyr Tyr Ala Asn Lys Leu Leu Phe Ala Pro Gln
          115          120          125

ctt gac gta gaa acc ccc tac agc tcc gcc gaa tgt gat cta tgc ggg     432
Leu Asp Val Glu Thr Pro Tyr Ser Ser Gly Glu Cys Asp Leu Cys Gly
          130          135          140

gtc tgc aaa ctg cag ctg atg gaa aaa ttg ctt aaa ccc ggt tgc cga     480
Val Cys Lys Leu Gln Leu Met Glu Lys Leu Leu Lys Pro Gly Cys Arg
145          150          155          160

tcc gtc tat atc gga gat ggg act tcc gat ttt tgc ccg gcg gaa agg     528
Ser Val Tyr Ile Gly Asp Gly Thr Ser Asp Phe Cys Pro Ala Glu Arg
          165          170          175

gcg gat aag gtc ttt gcc agg agc agg ctt tat cag cat tgc cag gag     576
Ala Asp Lys Val Phe Ala Arg Ser Arg Leu Tyr Gln His Cys Gln Glu
          180          185          190

gtg ggc aaa gaa gcc cag cta ttc caa tcg ttt cag gat att ctt cag     624
Val Gly Lys Glu Ala Gln Leu Phe Gln Ser Phe Gln Asp Ile Leu Gln
          195          200          205

aca gtt gaa cat tgg gga agg gaa gag gag gaa ggg act tga           666
Thr Val Glu His Trp Gly Arg Glu Glu Glu Glu Gly Thr
          210          215          220

```

<210> 70

5 <211> 221

<212> PRT

<213> *Desulfitobacterium hafniense*

<400> 70

10

ES 2 806 732 T3

Met Glu Glu Leu Asn Ser Ile Phe Phe Val Asp Phe Asp Gly Thr Ile
1 5 10 15

Val Thr Gln Asp Met Cys Ala Val Leu Val Glu Thr Leu Ala Gly Glu
20 25 30

Gly Trp Arg Glu Ile Asn Glu Leu Trp Glu Arg Lys Glu Leu Ser Thr
35 40 45

Leu Glu Cys Ala Arg Arg Thr Phe Lys Leu Phe Asn Ser Asn Asp Pro
50 55 60

Glu Val Phe Arg Gln Leu Ile Gly Gln Ala Val Phe Asp Pro Gly Phe
65 70 75 80

Leu Asp Phe Ala Ala Phe Cys Glu Gln Arg Gly Phe Pro Leu Ile Ile
85 90 95

Leu Ser Asp Gly Tyr Asp Phe Tyr Ile Glu Tyr Leu Leu Gln Arg Glu
100 105 110

Gly Leu Asn Leu Pro Tyr Tyr Ala Asn Lys Leu Leu Phe Ala Pro Gln
115 120 125
Leu Asp Val Glu Thr Pro Tyr Ser Ser Gly Glu Cys Asp Leu Cys Gly
130 135 140

Val Cys Lys Leu Gln Leu Met Glu Lys Leu Leu Lys Pro Gly Cys Arg
145 150 155 160

Ser Val Tyr Ile Gly Asp Gly Thr Ser Asp Phe Cys Pro Ala Glu Arg
165 170 175

Ala Asp Lys Val Phe Ala Arg Ser Arg Leu Tyr Gln His Cys Gln Glu
180 185 190

Val Gly Lys Glu Ala Gln Leu Phe Gln Ser Phe Gln Asp Ile Leu Gln
195 200 205

Thr Val Glu His Trp Gly Arg Glu Glu Glu Glu Gly Thr
210 215 220

<210> 71

5 <211> 642

<212> ADN

<213> *Pelotomaculum thermopropionicum*

<220>

10 <221> CDS

<222> (1)..(642)

<223> gen de *Pelotomaculum thermopropionicum* que codifica para TMP fosfatasa [WP_012032097]

<400> 71

15

ES 2 806 732 T3

atg gaa aaa gtt ttt ttt gtt gat ttt gac ggg acg gta acc aaa aag Met Glu Lys Val Phe Phe Val Asp Phe Asp Gly Thr Val Thr Lys Lys 1 5 10 15	48
gat acc tgc gtg gcc atg atc gag gcc ttt gcc ggc ggc aac tgg aga Asp Thr Cys Val Ala Met Ile Glu Ala Phe Ala Gly Gly Asn Trp Arg 20 25 30	96
gag att aac gag gcg tgg gaa aga aaa gaa att tcc acg gaa gaa tgt Glu Ile Asn Glu Ala Trp Glu Arg Lys Glu Ile Ser Thr Glu Glu Cys 35 40 45	144
gca aac atg atc ttc agg ctt ttc cgc gcc ggc att gaa gac atc agg Ala Asn Met Ile Phe Arg Leu Phe Arg Ala Gly Ile Glu Asp Ile Arg 50 55 60	192
aag ctt ttg gac ggt atc gag ata gac ggc cat ttt aaa gat ttt ctt Lys Leu Leu Asp Gly Ile Glu Ile Asp Gly His Phe Lys Asp Phe Leu 65 70 75 80	240
tct ttt tgc cgg gaa aga ggc tat aaa ata tac atc ctc agc gac ggt Ser Phe Cys Arg Glu Arg Gly Tyr Lys Ile Tyr Ile Leu Ser Asp Gly 85 90 95	288
tac gac ttt tgc att gag acg gtg ttt aaa aaa cac gga ata gag ctg Tyr Asp Phe Cys Ile Glu Thr Val Phe Lys Lys His Gly Ile Glu Leu 100 105 110	336
ccg tac tat gcc aac aaa atg gtt tac ggc aat ggt ttt aaa ata gaa Pro Tyr Tyr Ala Asn Lys Met Val Tyr Gly Asn Gly Phe Lys Ile Glu 115 120 125	384
tgc ttc agg ccc aac ccg gcc tgc ggt att tgc ggg acc tgc aag acc Cys Phe Arg Pro Asn Pro Ala Cys Gly Ile Cys Gly Thr Cys Lys Thr 130 135 140	432
aag ctg att gag gag ctt aaa ggg gac ggc agc cag gtt att tac att Lys Leu Ile Glu Glu Leu Lys Gly Asp Gly Ser Gln Val Ile Tyr Ile 145 150 155 160	480
ggc gac gga tat tgc gac aca tgc ccg gcc atg aaa gcc gat gtg gtt Gly Asp Gly Tyr Ser Asp Thr Cys Pro Ala Met Lys Ala Asp Val Val 165 170 175	528
ttt gcc aag gga gta ttg tac agg cat tgc cgg gaa aac ggc aaa aag Phe Ala Lys Gly Val Leu Tyr Arg His Cys Arg Glu Asn Gly Lys Lys 180 185 190	576
gct att tat tat aat aac ttt ggt gat att att aat tat ttt ttc caa Ala Ile Tyr Tyr Asn Asn Phe Gly Asp Ile Ile Asn Tyr Phe Phe Gln 195 200 205	624
ata aaa aaa agt ttg taa Ile Lys Lys Ser Leu 210	642

<210> 72

5 <211> 213

<212> PRT

<213> *Pelotomaculum thermopropionicum*

<400> 72

10

ES 2 806 732 T3

Met Glu Lys Val Phe Phe Val Asp Phe Asp Gly Thr Val Thr Lys Lys
1 5 10 15

Asp Thr Cys Val Ala Met Ile Glu Ala Phe Ala Gly Gly Asn Trp Arg
20 25 30

Glu Ile Asn Glu Ala Trp Glu Arg Lys Glu Ile Ser Thr Glu Glu Cys
35 40 45

Ala Asn Met Ile Phe Arg Leu Phe Arg Ala Gly Ile Glu Asp Ile Arg
50 55 60

Lys Leu Leu Asp Gly Ile Glu Ile Asp Gly His Phe Lys Asp Phe Leu
65 70 75 80

Ser Phe Cys Arg Glu Arg Gly Tyr Lys Ile Tyr Ile Leu Ser Asp Gly
85 90 95
Tyr Asp Phe Cys Ile Glu Thr Val Phe Lys Lys His Gly Ile Glu Leu
100 105 110

Pro Tyr Tyr Ala Asn Lys Met Val Tyr Gly Asn Gly Phe Lys Ile Glu
115 120 125

Cys Phe Arg Pro Asn Pro Ala Cys Gly Ile Cys Gly Thr Cys Lys Thr
130 135 140

Lys Leu Ile Glu Glu Leu Lys Gly Asp Gly Ser Gln Val Ile Tyr Ile
145 150 155 160

Gly Asp Gly Tyr Ser Asp Thr Cys Pro Ala Met Lys Ala Asp Val Val
165 170 175

Phe Ala Lys Gly Val Leu Tyr Arg His Cys Arg Glu Asn Gly Lys Lys
180 185 190

Ala Ile Tyr Tyr Asn Asn Phe Gly Asp Ile Ile Asn Tyr Phe Phe Gln
195 200 205

Ile Lys Lys Ser Leu
210

<210> 73

5 <211> 651

<212> ADN

<213> *Desulfotomaculum ruminis*

<220>

10 <221> CDS

<222> (1)..(651)

<223> gen de *Desulfotomaculum ruminis* que codifica para TMP fosfatasa [WP_013840216]

<400> 73

15

ES 2 806 732 T3

atg gaa acc att ctt ttt ctg gat ttt gac ggc acc att acc gag cag	48
Met Glu Thr Ile Leu Phe Leu Asp Phe Asp Gly Thr Ile Thr Glu Gln	
1 5 10 15	
gat acc tgc gat atg ctg atg gag cgc tac ggc aat gcg gaa tgt ctg	96
Asp Thr Cys Asp Met Leu Met Glu Arg Tyr Gly Asn Ala Glu Cys Leu	
20 25 30	
gaa ttg aac cgg cgc tgg gaa cgc aag gaa att tcc acc atg gaa tgt	144
Glu Leu Asn Arg Arg Trp Glu Arg Lys Glu Ile Ser Thr Met Glu Cys	
35 40 45	
gcc cgg cag tcc ttc cgg caa atg cag gta act ccc gag gtt cta aag	192
Ala Arg Gln Ser Phe Arg Gln Met Gln Val Thr Pro Glu Val Leu Lys	
50 55 60	
cgg ttg gtg cag gag gtg aag gta gac cct cat ttg aaa gaa ttg ctc	240
Arg Leu Val Gln Glu Val Lys Val Asp Pro His Leu Lys Glu Leu Leu	
65 70 75 80	
cgt ttc tgt gag cag gag aat tac ccc gcc tat att ttg agc gat ggg	288
Arg Phe Cys Glu Gln Glu Asn Tyr Pro Ala Tyr Ile Leu Ser Asp Gly	
85 90 95	
tat gaa ccc atc att cag ggg gta ctg cag cgg gaa gga ata aaa ata	336
Tyr Glu Pro Ile Ile Gln Gly Val Leu Gln Arg Glu Gly Ile Lys Ile	
100 105 110	
tct tgt ttt tgc aac ggg ttg tcc ttt gac ggc cag tac cgg gtc atg	384
Ser Cys Phe Cys Asn Gly Leu Ser Phe Asp Gly Gln Tyr Arg Val Met	
115 120 125	
gcg cct cac tat aat ccc cgg tgc ggc cgg tgc gga acc tgt aaa caa	432
Ala Pro His Tyr Asn Pro Arg Cys Gly Arg Cys Gly Thr Cys Lys Gln	
130 135 140	
aag ctg gtg gaa cgc ctg ggt cag ccg ggc gcc cgg aag att ttt gtg	480
Lys Leu Val Glu Arg Leu Gly Gln Pro Gly Ala Arg Lys Ile Phe Val	
145 150 155 160	
gga gac ggt tat tgc gat ttc tgt gcc gca gag tcc tgc agt aag gtc	528
Gly Asp Gly Tyr Ser Asp Phe Cys Ala Ala Glu Ser Cys Ser Lys Val	
165 170 175	
ttt gct aaa aaa aat tta ttg aag tat tgc ctg gaa aac cag att ccg	576
Phe Ala Lys Lys Asn Leu Leu Lys Tyr Cys Leu Glu Asn Gln Ile Pro	
180 185 190	
gcc cac ccc tat gaa acc ctg gga gag gtt tta cag tgg ctg aga gga	624
Ala His Pro Tyr Glu Thr Leu Gly Glu Val Leu Gln Trp Leu Arg Gly	
195 200 205	
gag gct gaa cat gga cat ccg gtt taa	651
Glu Ala Glu His Gly His Pro Val	
210 215	

<210> 74

5 <211> 216

<212> PRT

<213> *Desulfotomaculum ruminis*

<400> 74

10

ES 2 806 732 T3

Met Glu Thr Ile Leu Phe Leu Asp Phe Asp Gly Thr Ile Thr Glu Gln
1 5 10 15

Asp Thr Cys Asp Met Leu Met Glu Arg Tyr Gly Asn Ala Glu Cys Leu
20 25 30

Glu Leu Asn Arg Arg Trp Glu Arg Lys Glu Ile Ser Thr Met Glu Cys
35 40 45

Ala Arg Gln Ser Phe Arg Gln Met Gln Val Thr Pro Glu Val Leu Lys
50 55 60

Arg Leu Val Gln Glu Val Lys Val Asp Pro His Leu Lys Glu Leu Leu
65 70 75 80

Arg Phe Cys Glu Gln Glu Asn Tyr Pro Ala Tyr Ile Leu Ser Asp Gly
85 90 95

Tyr Glu Pro Ile Ile Gln Gly Val Leu Gln Arg Glu Gly Ile Lys Ile
100 105 110

Ser Cys Phe Cys Asn Gly Leu Ser Phe Asp Gly Gln Tyr Arg Val Met
115 120 125

Ala Pro His Tyr Asn Pro Arg Cys Gly Arg Cys Gly Thr Cys Lys Gln
130 135 140

Lys Leu Val Glu Arg Leu Gly Gln Pro Gly Ala Arg Lys Ile Phe Val
145 150 155 160

Gly Asp Gly Tyr Ser Asp Phe Cys Ala Ala Glu Ser Cys Ser Lys Val
165 170 175

Phe Ala Lys Lys Asn Leu Leu Lys Tyr Cys Leu Glu Asn Gln Ile Pro
180 185 190

Ala His Pro Tyr Glu Thr Leu Gly Glu Val Leu Gln Trp Leu Arg Gly
195 200 205

Glu Ala Glu His Gly His Pro Val
210 215

<210> 75

5 <211> 1896

<212> ADN

<213> *Escherichia coli*

<220>

10 <221> CDS

<222> (1)..(1896)

<223> gen ThiC de *E. coli* que codifica para HMP-P sintasa

<400> 75

15

ES 2 806 732 T3

atg tct gca aca aaa ctg acc cgc cgc gaa caa cgc gcc cgg gcc caa Met Ser Ala Thr Lys Leu Thr Arg Arg Glu Gln Arg Ala Arg Ala Gln 1 5 10 15	48
cat ttt atc gac acc ctg gaa ggc acc gcc ttt ccc aac tca aaa cgc His Phe Ile Asp Thr Leu Glu Gly Thr Ala Phe Pro Asn Ser Lys Arg 20 25 30	96
att tat atc act ggc aca cac ccc ggc gtg cgc gtg ccg atg cgt gag Ile Tyr Ile Thr Gly Thr His Pro Gly Val Arg Val Pro Met Arg Glu 35 40 45	144
atc cag ctt agc ccg acg cta att ggc ggt agc aaa gaa cag ccg cag Ile Gln Leu Ser Pro Thr Leu Ile Gly Gly Ser Lys Glu Gln Pro Gln 50 55 60	192
tac gaa gaa aac gaa gcg att ccg gtc tac gac acc tcc ggc ccg tat Tyr Glu Glu Asn Glu Ala Ile Pro Val Tyr Asp Thr Ser Gly Pro Tyr 65 70 75 80	240
ggt gat ccg cag att gcc att aac gtg cag caa ggg ctg gca aaa cta Gly Asp Pro Gln Ile Ala Ile Asn Val Gln Gln Gly Leu Ala Lys Leu 85 90 95	288
cgc cag ccg tgg atc gat gcg cgc ggc gat acc gaa gaa ctt acc gtg Arg Gln Pro Trp Ile Asp Ala Arg Gly Asp Thr Glu Glu Leu Thr Val 100 105 110	336
cgc agt tcc gat tac act aaa gcg cgc ctg gca gat gat ggc ctc gac Arg Ser Ser Asp Tyr Thr Lys Ala Arg Leu Ala Asp Asp Gly Leu Asp 115 120 125	384
gaa ctg cgt ttt agc ggc gta cta aca cca aaa cgc gcc aaa gca gga Glu Leu Arg Phe Ser Gly Val Leu Thr Pro Lys Arg Ala Lys Ala Gly 130 135 140	432
cgc cgt gtc acc caa ctg cac tac gcc cgc cag ggc atc atc acg ccg Arg Arg Val Thr Gln Leu His Tyr Ala Arg Gln Gly Ile Ile Thr Pro 145 150 155 160	480
gaa atg gaa ttc atc gcc atc cgc gag aat atg ggc cgc gag cgc atc Glu Met Glu Phe Ile Ala Ile Arg Glu Asn Met Gly Arg Glu Arg Ile 165 170 175	528
cgt agc gag gtt tta cgc cac cag cat ccg gga atg agc ttt ggc gca Arg Ser Glu Val Leu Arg His Gln His Pro Gly Met Ser Phe Gly Ala 180 185 190	576
cat ctg ccg gaa aat atc act gcg gaa ttt gtc cgt gat gaa gtt gct His Leu Pro Glu Asn Ile Thr Ala Glu Phe Val Arg Asp Glu Val Ala 195 200 205	624
gcc gga cgt gcg att atc ccg gcc aac att aat cat ccg gaa tcg gag Ala Gly Arg Ala Ile Ile Pro Ala Asn Ile Asn His Pro Glu Ser Glu 210 215 220	672
ccg atg att att ggt cgc aat ttc ctg gta aaa gtt aac gcc aat atc Pro Met Ile Ile Gly Arg Asn Phe Leu Val Lys Val Asn Ala Asn Ile 225 230 235 240	720
ggc aac tcg gcg gtc acc tct tcc atc gaa gaa gaa gtg gaa aag ctg Gly Asn Ser Ala Val Thr Ser Ser Ile Glu Glu Glu Val Glu Lys Leu 245 250 255	768
gta tgg tcc acg cgc tgg gga gcg gat acg gtg atg gat ctc tcc acc Val Trp Ser Thr Arg Trp Gly Ala Asp Thr Val Met Asp Leu Ser Thr 260 265 270	816
ggt cgc tat att cac gaa acc cgc gag tgg att ttg cgt aac agc ccg Gly Arg Tyr Ile His Glu Thr Arg Glu Trp Ile Leu Arg Asn Ser Pro 275 280 285	864
gtg ccg atc ggt aca gtg ccg atc tac cag gcg ctg gag aag gtt aac Val Pro Ile Gly Thr Val Pro Ile Tyr Gln Ala Leu Glu Lys Val Asn 290 295 300	912

ES 2 806 732 T3

ggg atc gcc gaa gat ctt acc tgg gaa gcg ttc cgc gac acg ctg ctg Gly Ile Ala Glu Asp Leu Thr Trp Glu Ala Phe Arg Asp Thr Leu Leu 305 310 315 320	960
gaa cag gcc gag caa ggt gtg gat tac ttc act atc cat gcg ggc gta Glu Gln Ala Glu Gln Gly Val Asp Tyr Phe Thr Ile His Ala Gly Val 325 330 335	1008
ctg ctg cgc tat gtg ccg atg acc gcg aaa cgc ctg acc ggt atc gtc Leu Leu Arg Tyr Val Pro Met Thr Ala Lys Arg Leu Thr Gly Ile Val 340 345 350	1056
tct cgc ggc ggt tgc att atg gcg aaa tgg tgc ctc tcc cat cat cag Ser Arg Gly Gly Ser Ile Met Ala Lys Trp Cys Leu Ser His His Gln 355 360 365	1104
gaa aat ttc ctc tat caa cac ttc cgc gaa att tgt gaa atc tgt gcc Glu Asn Phe Leu Tyr Gln His Phe Arg Glu Ile Cys Glu Ile Cys Ala 370 375 380	1152
gct tat gac gtt tgc ctg tgc ctg ggc gac ggt ctg cgc ccc ggt tct Ala Tyr Asp Val Ser Leu Ser Leu Gly Asp Gly Leu Arg Pro Gly Ser 385 390 395 400	1200
att cag gac gcc aac gat gaa gcg cag ttt gcc gag ctg cat acg ctg Ile Gln Asp Ala Asn Asp Glu Ala Gln Phe Ala Glu Leu His Thr Leu 405 410 415	1248
ggc gaa ctg acc aaa att gcc tgg gaa tat gac gtg cag gtg atg att Gly Glu Leu Thr Lys Ile Ala Trp Glu Tyr Asp Val Gln Val Met Ile 420 425 430	1296
gaa ggc cca ggc cac gtg ccg atg cag atg atc cgc cgc aat atg acc Glu Gly Pro Gly His Val Pro Met Gln Met Ile Arg Arg Asn Met Thr 435 440 445	1344
gag gag tta gag cac tgc cac gaa gcg ccg ttt tac act ctg ggg ccg Glu Glu Leu Glu His Cys His Glu Ala Pro Phe Tyr Thr Leu Gly Pro 450 455 460	1392
cta act acc gat att gcg ccg ggc tat gac cac ttc acg tgc ggg att Leu Thr Thr Asp Ile Ala Pro Gly Tyr Asp His Phe Thr Ser Gly Ile 465 470 475 480	1440
ggt gcg gcg atg att ggc tgg ttt ggc tgc gcg atg ctc tgt tac gta Gly Ala Ala Met Ile Gly Trp Phe Gly Cys Ala Met Leu Cys Tyr Val 485 490 495	1488
acg cca aaa gag cat ctg ggt ctg ccc aat aaa gaa gat gtt aag cag Thr Pro Lys Glu His Leu Gly Leu Pro Asn Lys Glu Asp Val Lys Gln 500 505 510	1536
ggg ctt atc acc tat aag att gct gcc cac gcc gct gac ctg gcg aaa Gly Leu Ile Thr Tyr Lys Ile Ala Ala His Ala Ala Asp Leu Ala Lys 515 520 525	1584
ggg cat ccg ggc gcg caa att cgc gat aac gcc atg tgc aaa gcc cgc Gly His Pro Gly Ala Gln Ile Arg Asp Asn Ala Met Ser Lys Ala Arg 530 535 540	1632
ttc gaa ttt cgc tgg gaa gac cag ttt aat ctg gcc ctc gac ccg ttt Phe Glu Phe Arg Trp Glu Asp Gln Phe Asn Leu Ala Leu Asp Pro Phe	1680

ES 2 806 732 T3

545	550	555	560	
acc gcc cgc gct tat	cac gat gaa acc	ctg ccg caa gag	tca ggt aaa	1728
Thr Ala Arg Ala Tyr	His Asp Glu Thr	Leu Pro Gln Glu	Ser Gly Lys	
	565	570	575	
gtc gcc cat ttt tgc	tcc atg tgt ggg	ccg aaa ttc tgc	tcg atg aaa	1776
Val Ala His Phe Cys	Ser Met Cys Gly	Pro Lys Phe Cys	Ser Met Lys	
	580	585	590	
atc agc cag gaa gtg	cgt gat tac gcc	gcc acg caa act	att gaa atg	1824
Ile Ser Gln Glu Val	Arg Asp Tyr Ala	Ala Thr Gln Thr	Ile Glu Met	
	595	600	605	
gga atg gcg gat atg	tcg gag aac ttc	cgt gcc aga ggc	gga gaa atc	1872
Gly Met Ala Asp Met	Ser Glu Asn Phe	Arg Ala Arg Gly	Gly Glu Ile	
	610	615	620	
tac ctg cgt aag gag	gaa gcg tga			1896
Tyr Leu Arg Lys Glu	Glu Ala			
	625	630		

<210> 76

<211> 631

5 <212> PRT

<213> *Escherichia coli*

<400> 76

Met Ser Ala Thr Lys	Leu Thr Arg Arg	Glu Gln Arg Ala	Arg Ala Gln
1	5	10	15
His Phe Ile Asp Thr	Leu Glu Gly Thr	Ala Phe Pro Asn	Ser Lys Arg
	20	25	30
Ile Tyr Ile Thr Gly	Thr His Pro Gly	Val Arg Val Pro	Met Arg Glu
	35	40	45
Ile Gln Leu Ser Pro	Thr Leu Ile Gly	Gly Ser Lys Glu	Gln Pro Gln
	50	55	60
Tyr Glu Glu Asn Glu	Ala Ile Pro Val	Tyr Asp Thr Ser	Gly Pro Tyr
	65	70	75
Gly Asp Pro Gln Ile	Ala Ile Asn Val	Gln Gln Gly Leu	Ala Lys Leu
	85	90	95
Arg Gln Pro Trp Ile	Asp Ala Arg Gly	Asp Thr Glu Glu	Leu Thr Val
	100	105	110
Arg Ser Ser Asp Tyr	Thr Lys Ala Arg	Leu Ala Asp Asp	Gly Leu Asp
	115	120	125
Glu Leu Arg Phe Ser	Gly Val Leu Thr	Pro Lys Arg Ala	Lys Ala Gly

10

ES 2 806 732 T3

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

ES 2 806 732 T3

Ala Tyr Asp Val Ser Leu Ser Leu Gly Asp Gly Leu Arg Pro Gly Ser
385 390 395 400

Ile Gln Asp Ala Asn Asp Glu Ala Gln Phe Ala Glu Leu His Thr Leu
405 410 415

Gly Glu Leu Thr Lys Ile Ala Trp Glu Tyr Asp Val Gln Val Met Ile
420 425 430

Glu Gly Pro Gly His Val Pro Met Gln Met Ile Arg Arg Asn Met Thr
435 440 445

Glu Glu Leu Glu His Cys His Glu Ala Pro Phe Tyr Thr Leu Gly Pro
450 455 460

Leu Thr Thr Asp Ile Ala Pro Gly Tyr Asp His Phe Thr Ser Gly Ile
465 470 475 480

Gly Ala Ala Met Ile Gly Trp Phe Gly Cys Ala Met Leu Cys Tyr Val
485 490 495

Thr Pro Lys Glu His Leu Gly Leu Pro Asn Lys Glu Asp Val Lys Gln
500 505 510

Gly Leu Ile Thr Tyr Lys Ile Ala Ala His Ala Ala Asp Leu Ala Lys
515 520 525

Gly His Pro Gly Ala Gln Ile Arg Asp Asn Ala Met Ser Lys Ala Arg
530 535 540

Phe Glu Phe Arg Trp Glu Asp Gln Phe Asn Leu Ala Leu Asp Pro Phe
545 550 555 560

Thr Ala Arg Ala Tyr His Asp Glu Thr Leu Pro Gln Glu Ser Gly Lys
565 570 575

Val Ala His Phe Cys Ser Met Cys Gly Pro Lys Phe Cys Ser Met Lys
580 585 590

Ile Ser Gln Glu Val Arg Asp Tyr Ala Ala Thr Gln Thr Ile Glu Met
595 600 605

Gly Met Ala Asp Met Ser Glu Asn Phe Arg Ala Arg Gly Gly Glu Ile
610 615 620

Tyr Leu Arg Lys Glu Glu Ala
625 630

<210> 77

<211> 1371

5 <212> ADN

<213> *Synechococcus elongatus*

<220>

<221> CDS

10 <222> (1)..(1371)

<223> gen ThiC de *Synechococcus elongatus_PCC_7942:_[NC_007604]* que codifica para una HMP-P sintasa

ES 2 806 732 T3

<400> 77

atg cgc agc gac tgg atc gca ccc cgc cga ggc caa gcc aac gtc act	48
Met Arg Ser Asp Trp Ile Ala Pro Arg Arg Gly Gln Ala Asn Val Thr	
1 5 10 15	
caa atg cac tac gcc cgc caa ggc gtg atc acc gaa gaa atg gac ttc	96
Gln Met His Tyr Ala Arg Gln Gly Val Ile Thr Glu Glu Met Asp Phe	
20 25 30	
gtg gcg cgg cgc gaa aat ctg cca gcc gat cta att cgg gat gaa gtg	144
Val Ala Arg Arg Glu Asn Leu Pro Ala Asp Leu Ile Arg Asp Glu Val	
35 40 45	
gca cgg ggt cgg atg att atc ccc gcc aac atc aac cac acc aat ttg	192
Ala Arg Gly Arg Met Ile Ile Pro Ala Asn Ile Asn His Thr Asn Leu	
50 55 60	
gag ccg atg gcg atc ggc att gcc tcc aag tgc aag gtc aac gcc aac	240
Glu Pro Met Ala Ile Gly Ile Ala Ser Lys Cys Lys Val Asn Ala Asn	
65 70 75 80	
atc ggt gct tcg cct aac gcc tcc aac atc gat gaa gaa gtc gag aag	288
Ile Gly Ala Ser Pro Asn Ala Ser Asn Ile Asp Glu Glu Val Glu Lys	
85 90 95	
ctg aag ctc gcg gtc aaa tac ggt gcc gat acc gtc atg gac ctc tcg	336
Leu Lys Leu Ala Val Lys Tyr Gly Ala Asp Thr Val Met Asp Leu Ser	
100 105 110	
acc gcc gcc gcc aac ctc gat gag att cgc acc gcg atc atc aat gct	384
Thr Gly Gly Gly Asn Leu Asp Glu Ile Arg Thr Ala Ile Ile Asn Ala	
115 120 125	
tcg ccg gta ccg atc ggc acc gtg ccg gtc tac caa gcc ctg gaa tcc	432
Ser Pro Val Pro Ile Gly Thr Val Pro Val Tyr Gln Ala Leu Glu Ser	
130 135 140	
gtt cac ggg cgc atc gaa aaa ctc agc gcc gac gac ttc ttg cat gtg	480
Val His Gly Arg Ile Glu Lys Leu Ser Ala Asp Asp Phe Leu His Val	
145 150 155 160	
atc gaa aag cac tgc gaa cag gcc gtc gac tac caa acc atc cac gcc	528
Ile Glu Lys His Cys Glu Gln Gly Val Asp Tyr Gln Thr Ile His Ala	
165 170 175	
ggg ctg ctg att gaa cac ctg ccc aag gtc aag agc cgg atc acc ggg	576
Gly Leu Leu Ile Glu His Leu Pro Lys Val Lys Ser Arg Ile Thr Gly	
180 185 190	
att gtt tcg ccg gcc gcc gcc atc att gcc cag tgg atg ctc tac cac	624

ES 2 806 732 T3

Ile	Val	Ser	Arg	Gly	Gly	Gly	Ile	Ile	Ala	Gln	Trp	Met	Leu	Tyr	His		
		195					200					205					
cac	aag	caa	aac	ccg	ctc	tat	acc	cac	ttt	cgc	gac	atc	atc	gaa	atc	672	
His	Lys	Gln	Asn	Pro	Leu	Tyr	Thr	His	Phe	Arg	Asp	Ile	Ile	Glu	Ile		
	210					215				220							
ttc	aag	cgc	tac	gac	tgt	agc	ttc	agc	ttg	ggc	gac	tcg	ctg	cgg	ccg	720	
Phe	Lys	Arg	Tyr	Asp	Cys	Ser	Phe	Ser	Leu	Gly	Asp	Ser	Leu	Arg	Pro		
225					230					235					240		
ggc	tgc	ctg	cac	gat	gct	agc	gac	gat	gcc	cag	ctc	agc	gag	ctg	aag	768	
Gly	Cys	Leu	His	Asp	Ala	Ser	Asp	Asp	Ala	Gln	Leu	Ser	Glu	Leu	Lys		
				245					250					255			
act	ctc	ggc	caa	ctg	acg	cgg	gtt	gct	tgg	gaa	cac	gac	gtg	caa	gtc	816	
Thr	Leu	Gly	Gln	Leu	Thr	Arg	Val	Ala	Trp	Glu	His	Asp	Val	Gln	Val		
			260					265					270				
atg	gtc	gaa	ggg	cca	ggc	cac	gtt	ccc	atg	gac	cag	atc	gag	ttc	aac	864	
Met	Val	Glu	Gly	Pro	Gly	His	Val	Pro	Met	Asp	Gln	Ile	Glu	Phe	Asn		
		275					280					285					
gtc	cgc	aag	caa	atg	gaa	gag	tgc	tca	gaa	gct	ccc	ttc	tac	gtc	ttg	912	
Val	Arg	Lys	Gln	Met	Glu	Gly	Cys	Ser	Glu	Ala	Pro	Phe	Tyr	Val	Leu		
	290					295					300						
ggc	ccc	ctc	gtg	acc	gac	att	gca	ccg	ggc	tat	gac	cac	atc	acc	agc	960	
Gly	Pro	Leu	Val	Thr	Asp	Ile	Ala	Pro	Gly	Tyr	Asp	His	Ile	Thr	Ser		
305					310					315					320		
gcg	atc	ggg	gca	gca	atg	gcg	ggc	tgg	tat	ggc	acg	gca	atg	ctc	tgc	1008	
Ala	Ile	Gly	Ala	Ala	Met	Ala	Gly	Trp	Tyr	Gly	Thr	Ala	Met	Leu	Cys		
				325					330					335			
tac	gtc	acg	ccc	aaa	gag	cac	ttg	ggc	ctg	ccc	aat	gcg	gaa	gat	gtg	1056	
Tyr	Val	Thr	Pro	Lys	Glu	His	Leu	Gly	Leu	Pro	Asn	Ala	Glu	Asp	Val		
			340					345					350				
cgc	aat	ggc	ttg	atc	gcc	tac	aaa	att	gcg	gct	cat	gca	gca	gat	atc	1104	
Arg	Asn	Gly	Leu	Ile	Ala	Tyr	Lys	Ile	Ala	Ala	His	Ala	Ala	Asp	Ile		
		355					360					365					
gct	cgc	cac	cgt	ccg	ggc	gct	cgc	gat	cgc	gat	gat	gaa	ctg	agt	cgg	1152	
Ala	Arg	His	Arg	Pro	Gly	Ala	Arg	Asp	Arg	Asp	Asp	Glu	Leu	Ser	Arg		
	370				375						380						
gca	cgc	tac	gcc	ttc	gac	tgg	aac	aag	caa	ttt	gac	ttg	agc	ctc	gat	1200	
Ala	Arg	Tyr	Ala	Phe	Asp	Trp	Asn	Lys	Gln	Phe	Asp	Leu	Ser	Leu	Asp		
	385				390					395					400		
cca	gag	cgg	gcg	cgg	gaa	tac	cac	gac	gaa	act	ctg	cca	gca	gat	atc	1248	
Pro	Glu	Arg	Ala	Arg	Glu	Tyr	His	Asp	Glu	Thr	Leu	Pro	Ala	Asp	Ile		
				405					410					415			
tac	aaa	acg	gca	gaa	ttc	tgt	tcg	atg	tgt	gga	ccg	aag	cac	tgt	ccg	1296	
Tyr	Lys	Thr	Ala	Glu	Phe	Cys	Ser	Met	Cys	Gly	Pro	Lys	His	Cys	Pro		
			420					425					430				
atg	caa	acc	aag	atc	acc	gag	gaa	gat	cta	acc	gag	ttg	gaa	aaa	ttc	1344	
Met	Gln	Thr	Lys	Ile	Thr	Glu	Glu	Asp	Leu	Thr	Glu	Leu	Glu	Lys	Phe		
		435					440					445					
ctc	gag	aaa	gat	agc	gct	ctg	gcg	tag								1371	
Leu	Glu	Lys	Asp	Ser	Ala	Leu	Ala										
	450					455											

<210> 78

5 <211> 456

<212> PRT

<213> *Synechococcus elongatus*

<400> 78

10

ES 2 806 732 T3

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

ES 2 806 732 T3

His Lys Gln Asn Pro Leu Tyr Thr His Phe Arg Asp Ile Ile Glu Ile
210 215 220

Phe Lys Arg Tyr Asp Cys Ser Phe Ser Leu Gly Asp Ser Leu Arg Pro
225 230 235 240

Gly Cys Leu His Asp Ala Ser Asp Asp Ala Gln Leu Ser Glu Leu Lys
245 250 255

Thr Leu Gly Gln Leu Thr Arg Val Ala Trp Glu His Asp Val Gln Val
260 265 270

Met Val Glu Gly Pro Gly His Val Pro Met Asp Gln Ile Glu Phe Asn
275 280 285

Val Arg Lys Gln Met Glu Glu Cys Ser Glu Ala Pro Phe Tyr Val Leu
290 295 300

Gly Pro Leu Val Thr Asp Ile Ala Pro Gly Tyr Asp His Ile Thr Ser
305 310 315 320

Ala Ile Gly Ala Ala Met Ala Gly Trp Tyr Gly Thr Ala Met Leu Cys
325 330 335

Tyr Val Thr Pro Lys Glu His Leu Gly Leu Pro Asn Ala Glu Asp Val
340 345 350

Arg Asn Gly Leu Ile Ala Tyr Lys Ile Ala Ala His Ala Ala Asp Ile
355 360 365

Ala Arg His Arg Pro Gly Ala Arg Asp Arg Asp Asp Glu Leu Ser Arg
370 375 380

Ala Arg Tyr Ala Phe Asp Trp Asn Lys Gln Phe Asp Leu Ser Leu Asp
385 390 395 400

Pro Glu Arg Ala Arg Glu Tyr His Asp Glu Thr Leu Pro Ala Asp Ile
405 410 415

Tyr Lys Thr Ala Glu Phe Cys Ser Met Cys Gly Pro Lys His Cys Pro
420 425 430

Met Gln Thr Lys Ile Thr Glu Glu Asp Leu Thr Glu Leu Glu Lys Phe
435 440 445

Leu Glu Lys Asp Ser Ala Leu Ala
450 455

<210> 79

<211> 1764

5 <212> ADN

<213> *Corynebacterium glutamicum*

<220>

<221> CDS

10 <222> (1)..(1764)

<223> gen ThiC de *Corynebacterium glutamicum* que codifica para una HMP-P sintasa

ES 2 806 732 T3

<400> 79

atg acg cct acc caa aat gag atc cac ccg aaa cat agc tac tcc ccc Met Thr Pro Thr Gln Asn Glu Ile His Pro Lys His Ser Tyr Ser Pro 1 5 10 15	48
atc cgc aag gac ggt ctc gag gtc ccg gag acc gaa atc cgc ctc gat Ile Arg Lys Asp Gly Leu Glu Val Pro Glu Thr Glu Ile Arg Leu Asp 20 25 30	96
gac tcg cca agc ggc ccc aac gaa ccc ttc cgc atc tac cgc acc cgt Asp Ser Pro Ser Gly Pro Asn Glu Pro Phe Arg Ile Tyr Arg Thr Arg 35 40 45	144
ggc cca gaa acc aac ccc aag cag gga ctt ccg cgg ctg cgc gag tca Gly Pro Glu Thr Asn Pro Lys Gln Gly Leu Pro Arg Leu Arg Glu Ser 50 55 60	192
tgg atc acc gcc cgc ggc gac gtt gcc acc tat cag ggg cgc gag cgt Trp Ile Thr Ala Arg Gly Asp Val Ala Thr Tyr Gln Gly Arg Glu Arg 65 70 75 80	240
ttg ctt atc gac gac ggc cgc tcg gca atg cgt cga ggt caa gct tcg Leu Leu Ile Asp Asp Gly Arg Ser Ala Met Arg Arg Gly Gln Ala Ser 85 90 95	288
gct gag tgg aaa ggc caa aaa cca gct cct ttg aag gcg cta cct ggc Ala Glu Trp Lys Gly Gln Lys Pro Ala Pro Leu Lys Ala Leu Pro Gly 100 105 110	336
aaa aga gtc acc caa atg gcc tat gca cgt gct ggc gtg att act cgt Lys Arg Val Thr Gln Met Ala Tyr Ala Arg Ala Gly Val Ile Thr Arg 115 120 125	384
gaa atg gag ttt gta gcg ctg cgc gaa cac gtt gat gcg gag ttt gtg Glu Met Glu Phe Val Ala Leu Arg Glu His Val Asp Ala Glu Phe Val 130 135 140	432
cgc tct gag gtg gcg cgc ggt cgg gcc att att ccc aac aac gtc aac Arg Ser Glu Val Ala Arg Gly Arg Ala Ile Ile Pro Asn Asn Val Asn 145 150 155 160	480
cac ccc gaa tct gaa ccg atg att att ggt cgc aaa ttt ttg acc aaa His Pro Glu Ser Glu Pro Met Ile Ile Gly Arg Lys Phe Leu Thr Lys 165 170 175	528
atc aac gcc aat att ggc aat tct gcg gtc acc tct tca atc gag gaa Ile Asn Ala Asn Ile Gly Asn Ser Ala Val Thr Ser Ser Ile Glu Glu 180 185 190	576

ES 2 806 732 T3

gag gtg tcc aag ctg cag tgg gcc acg cgc tgg ggt gcc gat acc gtg Glu Val Ser Lys Leu Gln Trp Ala Thr Arg Trp Gly Ala Asp Thr Val 195 200 205	624
atg gat cta tcc acc ggc gat gat att cac acc acc cgc gaa tgg att Met Asp Leu Ser Thr Gly Asp Asp Ile His Thr Thr Arg Glu Trp Ile 210 215 220	672
atc cgc aac tcc ccc gtt cct atc ggc acc gtc ccg atc tac caa gcg Ile Arg Asn Ser Pro Val Pro Ile Gly Thr Val Pro Ile Tyr Gln Ala 225 230 235 240	720
ctg gaa aaa gta aat ggc gtg gcc gca gac ctt aac tgg gaa gta ttc Leu Glu Lys Val Asn Gly Val Ala Ala Asp Leu Asn Trp Glu Val Phe 245 250 255	768
cgc gat acc atc att gag cag tgt gaa caa ggc gtg gac tat atg acc Arg Asp Thr Ile Ile Glu Gln Cys Glu Gln Gly Val Asp Tyr Met Thr 260 265 270	816
atc cac gcc ggc gtc ctg ctg gct tat atc cca ctg act acc cgt cgt Ile His Ala Gly Val Leu Leu Ala Tyr Ile Pro Leu Thr Thr Arg Arg 275 280 285	864
gtc acc ggc att gtc tcc cgc ggc gga tcc att atg gcc ggt tgg tgt Val Thr Gly Ile Val Ser Arg Gly Gly Ser Ile Met Ala Gly Trp Cys 290 295 300	912
ctg gcg cat cac cgc gaa tca ttc ctc tac gag cat ttc gac gag ctg Leu Ala His His Arg Glu Ser Phe Leu Tyr Glu His Phe Asp Glu Leu 305 310 315 320	960
tgc gaa atc ttt gca caa tat gac gtc gca ttc tcc ctc ggt gat ggc Cys Glu Ile Phe Ala Gln Tyr Asp Val Ala Phe Ser Leu Gly Asp Gly 325 330 335	1008
cta cgc ccc gga tcg ctt gcc gat gcc aac gac gcc gcg caa ttc gcc Leu Arg Pro Gly Ser Leu Ala Asp Ala Asn Asp Ala Ala Gln Phe Ala 340 345 350	1056
gag ctg aaa acc att ggt gag ctc acc caa cgc gcc tgg gaa tac gat Glu Leu Lys Thr Ile Gly Glu Leu Thr Gln Arg Ala Trp Glu Tyr Asp 355 360 365	1104
gta caa gta atg gtc gaa gga cct gga cac gtg cca cta aac atg atc Val Gln Val Met Val Glu Gly Pro Gly His Val Pro Leu Asn Met Ile 370 375 380	1152
cag gaa aac aac gag ctg gaa caa aag tgg gca gcg gac gca cct ttt Gln Glu Asn Asn Glu Leu Glu Gln Lys Trp Ala Ala Asp Ala Pro Phe 385 390 395 400	1200
tac act ctt gga cca cta gtt acc gac atc gct cca ggt tat gac cac Tyr Thr Leu Gly Pro Leu Val Thr Asp Ile Ala Pro Gly Tyr Asp His 405 410 415	1248
atc act tct gcc att ggt gca gct cac atc gcc atg ggt ggc acc gcc Ile Thr Ser Ala Ile Gly Ala Ala His Ile Ala Met Gly Gly Thr Ala 420 425 430	1296
atg ctg tgt tat gtc acc ccg aaa gaa cac ctt ggc ctg ccc aac cgt Met Leu Cys Tyr Val Thr Pro Lys Glu His Leu Gly Leu Pro Asn Arg 435 440 445	1344

ES 2 806 732 T3

gac gac gtc aaa acc ggc gta atc acc tac aag ctc gct gcc cac gca 1392
 Asp Asp Val Lys Thr Gly Val Ile Thr Tyr Lys Leu Ala Ala His Ala
 450 455 460

gca gat gtg gcc aag ggt cat ccc ggc gcg cgt gcc tgg gac gac gcc 1440
 Ala Asp Val Ala Lys Gly His Pro Gly Ala Arg Ala Trp Asp Asp Ala
 465 470 475 480

atg agt aaa gcg cgt ttt gaa ttc cgt tgg aat gat cag ttt gcg ctc 1488
 Met Ser Lys Ala Arg Phe Glu Phe Arg Trp Asn Asp Gln Phe Ala Leu
 485 490 495

tcc ctc gac ccc gac act gca atc gct tat cac gac gaa acc ctg ccg 1536
 Ser Leu Asp Pro Asp Thr Ala Ile Ala Tyr His Asp Glu Thr Leu Pro
 500 505 510

gca gag cct gcg aaa acc gca cac ttc tgt tca atg tgt ggc ccg aag 1584
 Ala Glu Pro Ala Lys Thr Ala His Phe Cys Ser Met Cys Gly Pro Lys
 515 520 525

ttc tgc tcc atg cga att agc cag gac att cgc gat atg ttt ggc gat 1632
 Phe Cys Ser Met Arg Ile Ser Gln Asp Ile Arg Asp Met Phe Gly Asp
 530 535 540

caa atc gcg gaa ttg ggg atg cct ggg gtt ggg gat tct tct agt gct 1680
 Gln Ile Ala Glu Leu Gly Met Pro Gly Val Gly Asp Ser Ser Ser Ala
 545 550 555 560

gtt gct tct agt ggg gca cgg gag ggg atg gct gag aaa tcc cgg gaa 1728
 Val Ala Ser Ser Gly Ala Arg Glu Gly Met Ala Glu Lys Ser Arg Glu
 565 570 575

ttt att gct ggt ggt gcg gag gtt tat cgg cgt tag 1764
 Phe Ile Ala Gly Gly Ala Glu Val Tyr Arg Arg
 580 585

<210> 80

<211> 587

5 <212> PRT

<213> *Corynebacterium glutamicum*

<400> 80

Met Thr Pro Thr Gln Asn Glu Ile His Pro Lys His Ser Tyr Ser Pro
 1 5 10 15

Ile Arg Lys Asp Gly Leu Glu Val Pro Glu Thr Glu Ile Arg Leu Asp
 20 25 30

Asp Ser Pro Ser Gly Pro Asn Glu Pro Phe Arg Ile Tyr Arg Thr Arg
 35 40 45

Gly Pro Glu Thr Asn Pro Lys Gln Gly Leu Pro Arg Leu Arg Glu Ser
 50 55 60

Trp Ile Thr Ala Arg Gly Asp Val Ala Thr Tyr Gln Gly Arg Glu Arg
 65 70 75 80

10

ES 2 806 732 T3

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

ES 2 806 732 T3

```

325              330              335

Leu Arg Pro Gly Ser Leu Ala Asp Ala Asn Asp Ala Ala Gln Phe Ala
      340              345              350

Glu Leu Lys Thr Ile Gly Glu Leu Thr Gln Arg Ala Trp Glu Tyr Asp
      355              360              365

Val Gln Val Met Val Glu Gly Pro Gly His Val Pro Leu Asn Met Ile
      370              375              380

Gln Glu Asn Asn Glu Leu Glu Gln Lys Trp Ala Ala Asp Ala Pro Phe
385              390              395              400

Tyr Thr Leu Gly Pro Leu Val Thr Asp Ile Ala Pro Gly Tyr Asp His
      405              410              415

Ile Thr Ser Ala Ile Gly Ala Ala His Ile Ala Met Gly Gly Thr Ala
      420              425              430

Met Leu Cys Tyr Val Thr Pro Lys Glu His Leu Gly Leu Pro Asn Arg
      435              440              445

Asp Asp Val Lys Thr Gly Val Ile Thr Tyr Lys Leu Ala Ala His Ala
450              455              460

Ala Asp Val Ala Lys Gly His Pro Gly Ala Arg Ala Trp Asp Asp Ala
465              470              475              480

Met Ser Lys Ala Arg Phe Glu Phe Arg Trp Asn Asp Gln Phe Ala Leu
      485              490              495

Ser Leu Asp Pro Asp Thr Ala Ile Ala Tyr His Asp Glu Thr Leu Pro
      500              505              510

Ala Glu Pro Ala Lys Thr Ala His Phe Cys Ser Met Cys Gly Pro Lys
      515              520              525

Phe Cys Ser Met Arg Ile Ser Gln Asp Ile Arg Asp Met Phe Gly Asp
530              535              540

Gln Ile Ala Glu Leu Gly Met Pro Gly Val Gly Asp Ser Ser Ser Ala
545              550              555              560

Val Ala Ser Ser Gly Ala Arg Glu Gly Met Ala Glu Lys Ser Arg Glu
      565              570              575
Phe Ile Ala Gly Gly Ala Glu Val Tyr Arg Arg
      580              585

```

<210> 81
 5 <211> 1869
 <212> ADN
 <213> *Candidatus Baumannia cicadellinicola*

<220>
 10 <221> CDS
 <222> (1)..(1869)

ES 2 806 732 T3

<223> gen ThiC de *Candidatus Baumannia cicadellinicola* [WP_011520252] que codifica para HMP-P sintasa

<400> 81

atg tca aga tca tca ata cct gct tca cgc cga gtg agc cgt gca aaa	48
Met Ser Arg Ser Ser Ile Pro Ala Ser Arg Val Ser Arg Ala Lys	
1 5 10 15	
gca cag gct ttt atg gat agc tta aca ggt agt agc tat ttt cct aac	96
Ala Gln Ala Phe Met Asp Ser Leu Thr Gly Ser Ser Tyr Phe Pro Asn	
20 25 30	
tca aga agg ata tat tta caa ggt aaa aca cct tca gta cat gta cca	144
Ser Arg Arg Ile Tyr Leu Gln Gly Lys Thr Pro Ser Val His Val Pro	
35 40 45	
atg cgt gaa att aag cta cat cct aca ttg atc ggt aaa aac ggt gaa	192
Met Arg Glu Ile Lys Leu His Pro Thr Leu Ile Gly Lys Asn Gly Glu	
50 55 60	
cat tat gag gat aat caa cct ata cca gtt tat gat act tca ggt cct	240
His Tyr Glu Asp Asn Gln Pro Ile Pro Val Tyr Asp Thr Ser Gly Pro	
65 70 75 80	
tac ggt gat cct act ata gca att aac gta cgt aca ggt ctt aac cgg	288
Tyr Gly Asp Pro Thr Ile Ala Ile Asn Val Arg Thr Gly Leu Asn Arg	
85 90 95	
tta cgc gag ata tgg att ctt gca cga caa gat agt gag cca ata agt	336
Leu Arg Glu Ile Trp Ile Leu Ala Arg Gln Asp Ser Glu Pro Ile Ser	
100 105 110	
aat aat aat aac gat cgt cag agt tca gat aaa cag tta agt ttt act	384
Asn Asn Asn Asn Asp Arg Gln Ser Ser Asp Lys Gln Leu Ser Phe Thr	
115 120 125	
act aac tat aat cca cgc cga gct agc tat gga cgc tgt att aca caa	432
Thr Asn Tyr Asn Pro Arg Arg Ala Ser Tyr Gly Arg Cys Ile Thr Gln	
130 135 140	
tta cat tac gca cgt gcc ggt atc ata acg cca gaa atg gag ttt ata	480
Leu His Tyr Ala Arg Ala Gly Ile Ile Thr Pro Glu Met Glu Phe Ile	
145 150 155 160	
gct tta cgt gaa aat atg ggc cga gaa cgt att agt agc aac gtg cta	528
Ala Leu Arg Glu Asn Met Gly Arg Glu Arg Ile Ser Ser Asn Val Leu	
165 170 175	
cat cag cag cat tta ggt tct aac ttt ggt gct aaa aaa gct gat cat	576
His Gln Gln His Leu Gly Ser Asn Phe Gly Ala Lys Lys Ala Asp His	

5

ES 2 806 732 T3

180	185	190	
att aca gca gaa ttt gtc cgg cag gaa gta gca gca gga cgt gct att Ile Thr Ala Glu Phe Val Arg Gln Glu Val Ala Ala Gly Arg Ala Ile 195 200 205			624
ata cct agt aat att aat cat cca gaa tct gag cca atg atc att ggc Ile Pro Ser Asn Ile Asn His Pro Glu Ser Glu Pro Met Ile Ile Gly 210 215 220			672
cgt aat ttt ctc gta aaa gta aat gca aat att ggt aac tca gca gta Arg Asn Phe Leu Val Lys Val Asn Ala Asn Ile Gly Asn Ser Ala Val 225 230 235 240			720
aca tct tct att gag gaa gaa gtc gaa aag tta gta tgg gct act cgt Thr Ser Ser Ile Glu Glu Val Glu Lys Leu Val Trp Ala Thr Arg 245 250 255			768
tgg gga gct gat aca gtc atg gac tta tct act ggt agt tat att cac Trp Gly Ala Asp Thr Val Met Asp Leu Ser Thr Gly Ser Tyr Ile His 260 265 270			816
gaa act aga gaa tgg ata tta cgt aat agc cca gta cct ata ggt act Glu Thr Arg Glu Trp Ile Leu Arg Asn Ser Pro Val Pro Ile Gly Thr 275 280 285			864
gta cct atc tat caa gcg tta gaa aaa gta aat gga gtc ata gaa aat Val Pro Ile Tyr Gln Ala Leu Lys Val Asn Gly Val Ile Glu Asn 290 295 300			912
ctt aat tgg gat att ttc tac gag aca tta tta gaa caa gct aac caa Leu Asn Trp Asp Ile Phe Tyr Glu Thr Leu Leu Glu Gln Ala Asn Gln 305 310 315 320			960
gga gta gat tat ttt acg att cat gct ggc gta tta aaa cgt tat gtt Gly Val Asp Tyr Phe Thr Ile His Ala Gly Val Leu Lys Arg Tyr Val 325 330 335			1008
cta cta aca gct agt agg tta act ggt atc gta tcg cgt ggt ggc tct Leu Leu Thr Ala Ser Arg Leu Thr Gly Ile Val Ser Arg Gly Gly Ser 340 345 350			1056
att atg gct caa tgg agt tta gta cat aat cag gaa aac ttc ctt tat Ile Met Ala Gln Trp Ser Leu Val His Asn Gln Glu Asn Phe Leu Tyr 355 360 365			1104
gag cat ttt agt gaa att tgc aag ctt tgt gct gct tat gat att gct Glu His Phe Ser Glu Ile Cys Lys Leu Cys Ala Ala Tyr Asp Ile Ala 370 375 380			1152
cta tct ctt gga gat ggt cta aga ccc ggt tcc gta caa gat gct aat Leu Ser Leu Gly Asp Gly Leu Arg Pro Gly Ser Val Gln Asp Ala Asn 385 390 395 400			1200
gat gaa gca caa ttt tct gag tta cat aca cta ggc gaa tta act aaa Asp Glu Ala Gln Phe Ser Glu Leu His Thr Leu Gly Glu Leu Thr Lys 405 410 415			1248
att gcc tgg gaa tat gat gtg caa gta atg atc gaa gga cct ggt cat Ile Ala Trp Glu Tyr Asp Val Gln Val Met Ile Glu Gly Pro Gly His 420 425 430			1296
att cca cta cat atg att gag cgt aat atg act gat caa ctt aaa tat			1344

ES 2 806 732 T3

Ile	Pro	Leu	His	Met	Ile	Glu	Arg	Asn	Met	Thr	Asp	Gln	Leu	Lys	Tyr		
		435					440					445					
tgc	cac	gaa	gca	cca	ttc	tac	act	ctc	gga	cca	ctc	aca	aca	gat	att	1392	
Cys	His	Glu	Ala	Pro	Phe	Tyr	Thr	Leu	Gly	Pro	Leu	Thr	Thr	Asp	Ile		
		450				455					460						
gct	cct	ggt	tat	gac	cac	ttt	act	tca	ggt	att	ggt	gcc	gca	cta	ata	1440	
Ala	Pro	Gly	Tyr	Asp	His	Phe	Thr	Ser	Gly	Ile	Gly	Ala	Ala	Leu	Ile		
465					470				475						480		
ggc	tgg	ttt	gga	tgt	gct	atg	ctg	tgc	tat	gta	act	cct	aaa	gag	cat	1488	
Gly	Trp	Phe	Gly	Cys	Ala	Met	Leu	Cys	Tyr	Val	Thr	Pro	Lys	Glu	His		
				485					490					495			
cta	ggt	tta	cct	aat	aag	gaa	gac	gta	aaa	cag	ggt	tta	att	gcc	tat	1536	
Leu	Gly	Leu	Pro	Asn	Lys	Glu	Asp	Val	Lys	Gln	Gly	Leu	Ile	Ala	Tyr		
			500					505					510				
aaa	att	gcc	gca	cat	gct	gca	gat	cta	gct	aaa	gga	cat	cct	ggt	gct	1584	
Lys	Ile	Ala	Ala	His	Ala	Ala	Asp	Leu	Ala	Lys	Gly	His	Pro	Gly	Ala		
		515					520					525					
caa	ata	cgt	gat	aat	gct	atg	tca	aaa	gct	cgt	ttc	gaa	ttt	cgc	tgg	1632	
Gln	Ile	Arg	Asp	Asn	Ala	Met	Ser	Lys	Ala	Arg	Phe	Glu	Phe	Arg	Trp		
		530				535					540						
gaa	gat	caa	ttt	aac	tta	gct	tta	gat	cct	ttt	acg	gcg	cgt	atg	tat	1680	
Glu	Asp	Gln	Phe	Asn	Leu	Ala	Leu	Asp	Pro	Phe	Thr	Ala	Arg	Met	Tyr		
545					550					555					560		
cac	gat	gaa	act	ata	ccg	caa	aca	gca	gga	aaa	tta	gca	aat	ttt	tgc	1728	
His	Asp	Glu	Thr	Ile	Pro	Gln	Thr	Ala	Gly	Lys	Leu	Ala	Asn	Phe	Cys		
				565					570					575			
tcg	atg	tgt	ggt	cct	aag	ttt	tgt	tct	atg	aag	cta	tca	aaa	aaa	ata	1776	
Ser	Met	Cys	Gly	Pro	Lys	Phe	Cys	Ser	Met	Lys	Leu	Ser	Lys	Lys	Ile		
			580					585					590				
cgt	aat	tac	act	aat	atg	aaa	aat	ata	aaa	act	att	agt	aat	agt	ttc	1824	
Arg	Asn	Tyr	Thr	Asn	Met	Lys	Asn	Ile	Lys	Thr	Ile	Ser	Asn	Ser	Phe		
		595					600					605					
atg	aat	aaa	tta	gat	aat	agc	ggt	att	aaa	aat	gct	gac	cga	taa		1869	
Met	Asn	Lys	Leu	Asp	Asn	Ser	Gly	Ile	Lys	Asn	Ala	Asp	Arg				
		610				615					620						

<210> 82

<211> 622

5 <212> PRT

<213> *Candidatus Baumannia cicadellinicola*

<400> 82

Met	Ser	Arg	Ser	Ser	Ile	Pro	Ala	Ser	Arg	Arg	Val	Ser	Arg	Ala	Lys
1				5					10					15	

Ala	Gln	Ala	Phe	Met	Asp	Ser	Leu	Thr	Gly	Ser	Ser	Tyr	Phe	Pro	Asn
			20					25					30		

10

ES 2 806 732 T3

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

ES 2 806 732 T3

Val Pro Ile Tyr Gln Ala Leu Glu Lys Val Asn Gly Val Ile Glu Asn
 290 295 300
 Leu Asn Trp Asp Ile Phe Tyr Glu Thr Leu Leu Glu Gln Ala Asn Gln
 305 310 315 320
 Gly Val Asp Tyr Phe Thr Ile His Ala Gly Val Leu Lys Arg Tyr Val
 325 330 335
 Leu Leu Thr Ala Ser Arg Leu Thr Gly Ile Val Ser Arg Gly Gly Ser
 340 345 350
 Ile Met Ala Gln Trp Ser Leu Val His Asn Gln Glu Asn Phe Leu Tyr
 355 360 365
 Glu His Phe Ser Glu Ile Cys Lys Leu Cys Ala Ala Tyr Asp Ile Ala
 370 375 380
 Leu Ser Leu Gly Asp Gly Leu Arg Pro Gly Ser Val Gln Asp Ala Asn
 385 390 395 400
 Asp Glu Ala Gln Phe Ser Glu Leu His Thr Leu Gly Glu Leu Thr Lys
 405 410 415
 Ile Ala Trp Glu Tyr Asp Val Gln Val Met Ile Glu Gly Pro Gly His
 420 425 430
 Ile Pro Leu His Met Ile Glu Arg Asn Met Thr Asp Gln Leu Lys Tyr
 435 440 445
 Cys His Glu Ala Pro Phe Tyr Thr Leu Gly Pro Leu Thr Thr Asp Ile
 450 455 460
 Ala Pro Gly Tyr Asp His Phe Thr Ser Gly Ile Gly Ala Ala Leu Ile
 465 470 475 480
 Gly Trp Phe Gly Cys Ala Met Leu Cys Tyr Val Thr Pro Lys Glu His
 485 490 495
 Leu Gly Leu Pro Asn Lys Glu Asp Val Lys Gln Gly Leu Ile Ala Tyr
 500 505 510
 Lys Ile Ala Ala His Ala Ala Asp Leu Ala Lys Gly His Pro Gly Ala
 515 520 525
 Gln Ile Arg Asp Asn Ala Met Ser Lys Ala Arg Phe Glu Phe Arg Trp
 530 535 540

ES 2 806 732 T3

Glu Asp Gln Phe Asn Leu Ala Leu Asp Pro Phe Thr Ala Arg Met Tyr
545 550 555 560

His Asp Glu Thr Ile Pro Gln Thr Ala Gly Lys Leu Ala Asn Phe Cys
565 570 575

Ser Met Cys Gly Pro Lys Phe Cys Ser Met Lys Leu Ser Lys Lys Ile
580 585 590

Arg Asn Tyr Thr Asn Met Lys Asn Ile Lys Thr Ile Ser Asn Ser Phe
595 600 605

Met Asn Lys Leu Asp Asn Ser Gly Ile Lys Asn Ala Asp Arg
610 615 620

<210> 83

<211> 636

5 <212> ADN

<213> *Escherichia coli*

<220>

<221> CDS

10 <222> (1)..(636)

<223> gen ThiE de *E. coli* que codifica para tiamina fosfato sintasa

<400> 83

atg tat cag cct gat ttt cct cct gta cct ttt cgt tca gga ctg tac 48
Met Tyr Gln Pro Asp Phe Pro Pro Val Pro Phe Arg Ser Gly Leu Tyr
1 5 10 15

cag gtg gtg gac agc gta cag tgg atc gaa cgt ctg ttg gat gca ggc 96
Pro Val Val Asp Ser Val Gln Trp Ile Glu Arg Leu Leu Asp Ala Gly
20 25 30

gta cgt act ctc cag cta cgc atc aaa gat cgg cgc gat gaa gag gtg 144
Val Arg Thr Leu Gln Leu Arg Ile Lys Asp Arg Arg Asp Glu Glu Val
35 40 45

gaa gcc gat gtc gtg gcg gca att gcg ctg ggc cgc cgc tat aac gcg 192
Glu Ala Asp Val Val Ala Ala Ile Ala Leu Gly Arg Arg Tyr Asn Ala
50 55 60

cga ttg ttt atc aac gat tac tgg cgg ctg gcg atc aag cat cag gcg 240
Arg Leu Phe Ile Asn Asp Tyr Trp Arg Leu Ala Ile Lys His Gln Ala
65 70 75 80

tat ggc gtc cat ttg ggg cag gaa gat ttg caa gcc acc gat ctc aat 288
Tyr Gly Val His Leu Gly Gln Glu Asp Leu Gln Ala Thr Asp Leu Asn
85 90 95

gcc atc cgc gcg gca ggc ctg cgg ctg ggc gtt tcg aca cat gac gat 336
Ala Ile Arg Ala Ala Gly Leu Arg Leu Gly Val Ser Thr His Asp Asp
100 105 110

15 atg gaa atc gac gtc gcg ctg gca gca cgc ccc tct tat atc gcg ctg 384

ES 2 806 732 T3

Met	Glu	Ile	Asp	Val	Ala	Leu	Ala	Ala	Arg	Pro	Ser	Tyr	Ile	Ala	Leu	
		115					120					125				
gga	cat	gtg	ttc	ccg	acg	caa	acc	aaa	cag	atg	cct	tct	gca	ccg	cag	432
Gly	His	Val	Phe	Pro	Thr	Gln	Thr	Lys	Gln	Met	Pro	Ser	Ala	Pro	Gln	
		130				135					140					
ggg	ctg	gaa	cag	ctg	gca	cgg	cat	gtt	gag	cga	ctg	gcg	gat	tat	ccc	480
Gly	Leu	Glu	Gln	Leu	Ala	Arg	His	Val	Glu	Arg	Leu	Ala	Asp	Tyr	Pro	
					150					155					160	
acc	gtg	gcg	att	ggc	ggt	atc	agt	ctg	gca	cgc	gcg	cct	gcg	gtg	ata	528
Thr	Val	Ala	Ile	Gly	Gly	Ile	Ser	Leu	Ala	Arg	Ala	Pro	Ala	Val	Ile	
				165					170					175		
gca	acg	ggt	gtc	ggc	agt	atc	gcc	gtc	gtc	agc	gcc	att	act	caa	gcc	576
Ala	Thr	Gly	Val	Gly	Ser	Ile	Ala	Val	Val	Ser	Ala	Ile	Thr	Gln	Ala	
			180					185					190			
gca	gac	tgg	cgt	ttg	gca	acg	gca	cag	ttg	ctg	gaa	att	gca	gga	gtt	624
Ala	Asp	Trp	Arg	Leu	Ala	Thr	Ala	Gln	Leu	Leu	Glu	Ile	Ala	Gly	Val	
		195					200					205				
ggc	gat	gaa	tga													636
Gly	Asp	Glu														
		210														

<210> 84
 <211> 211
 5 <212> PRT
 <213> *Escherichia coli*

<400> 84

Met	Tyr	Gln	Pro	Asp	Phe	Pro	Pro	Val	Pro	Phe	Arg	Ser	Gly	Leu	Tyr	
1				5					10					15		
Pro	Val	Val	Asp	Ser	Val	Gln	Trp	Ile	Glu	Arg	Leu	Leu	Asp	Ala	Gly	
			20					25					30			
Val	Arg	Thr	Leu	Gln	Leu	Arg	Ile	Lys	Asp	Arg	Arg	Asp	Glu	Glu	Val	
		35					40					45				
Glu	Ala	Asp	Val	Val	Ala	Ala	Ile	Ala	Leu	Gly	Arg	Arg	Tyr	Asn	Ala	
		50				55					60					
Arg	Leu	Phe	Ile	Asn	Asp	Tyr	Trp	Arg	Leu	Ala	Ile	Lys	His	Gln	Ala	
65				70					75					80		
Tyr	Gly	Val	His	Leu	Gly	Gln	Glu	Asp	Leu	Gln	Ala	Thr	Asp	Leu	Asn	
				85				90						95		
Ala	Ile	Arg	Ala	Ala	Gly	Leu	Arg	Leu	Gly	Val	Ser	Thr	His	Asp	Asp	
			100					105					110			

10

ES 2 806 732 T3

Met Glu Ile Asp Val Ala Leu Ala Ala Arg Pro Ser Tyr Ile Ala Leu
115 120 125

Gly His Val Phe Pro Thr Gln Thr Lys Gln Met Pro Ser Ala Pro Gln
130 135 140

Gly Leu Glu Gln Leu Ala Arg His Val Glu Arg Leu Ala Asp Tyr Pro
145 150 155 160

Thr Val Ala Ile Gly Gly Ile Ser Leu Ala Arg Ala Pro Ala Val Ile
165 170 175

Ala Thr Gly Val Gly Ser Ile Ala Val Val Ser Ala Ile Thr Gln Ala
180 185 190

Ala Asp Trp Arg Leu Ala Thr Ala Gln Leu Leu Glu Ile Ala Gly Val
195 200 205

Gly Asp Glu
210

<210> 85

<211> 756

5 <212> ADN

<213> *Escherichia coli*

<220>

<221> CDS

10 <222> (1)..(756)

<223> gen ThiF de *E. coli* que codifica para una adenililtransferasa de ThiS

<400> 85

atg aat gac cgt gac ttt atg cgt tat agc cgc caa atc ctg ctc gac 48
Met Asn Asp Arg Asp Phe Met Arg Tyr Ser Arg Gln Ile Leu Leu Asp
1 5 10 15

gat atc gct ctg gac ggg cag caa aaa ctg ctc gac agc cag gtg ctg 96
Asp Ile Ala Leu Asp Gly Gln Gln Lys Leu Leu Asp Ser Gln Val Leu
20 25 30

att atc ggt ctg ggc ggg ctg ggt aca cct gct gcg ctg tac ctg gcg 144
Ile Ile Gly Leu Gly Gly Leu Gly Thr Pro Ala Ala Leu Tyr Leu Ala
35 40 45

ggc gct ggc gtc ggg acg ctg gta ctg gca gat gac gac gat gtg cat 192
Gly Ala Gly Val Gly Thr Leu Val Leu Ala Asp Asp Asp Asp Val His
50 55 60

tta agc aat ctg caa cga caa atc ctc ttt acc act gaa gat atc gat 240
Leu Ser Asn Leu Gln Arg Gln Ile Leu Phe Thr Thr Glu Asp Ile Asp
65 70 75 80

cgc ccg aaa tcg cag gtc agc caa cag cga ctg aca cag ttg aat ccc 288
Arg Pro Lys Ser Gln Val Ser Gln Gln Arg Leu Thr Gln Leu Asn Pro
85 90 95

15

ES 2 806 732 T3

gac att caa ctg aca gca tta caa caa cgg tta acg ggt gag gcg tta	336
Asp Ile Gln Leu Thr Ala Leu Gln Gln Arg Leu Thr Gly Glu Ala Leu	
100 105 110	
aaa gat gcg gtt gca cgg gcc gat gtg gtg ctc gac tgt acc gac aat	384
Lys Asp Ala Val Ala Arg Ala Asp Val Val Leu Asp Cys Thr Asp Asn	
115 120 125	
atg gcg act cgc cag gag att aat gcc gcc tgc gtg gca ctc aac acg	432
Met Ala Thr Arg Gln Glu Ile Asn Ala Ala Cys Val Ala Leu Asn Thr	
130 135 140	
ccg ctt atc acc gcc agc gcg gtc gga ttt ggc ggt cag ttg atg gta	480
Pro Leu Ile Thr Ala Ser Ala Val Gly Phe Gly Gly Gln Leu Met Val	
145 150 155 160	
ctg acg ccg ccc tgg gag cag ggg tgt tac cgc tgc ctg tgg cca gat	528
Leu Thr Pro Pro Trp Glu Gln Gly Cys Tyr Arg Cys Leu Trp Pro Asp	
165 170 175	
aac cag gag cca gaa cgc aac tgc cgc acg gcg ggc gtg gtt ggc ccg	576
Asn Gln Glu Pro Glu Arg Asn Cys Arg Thr Ala Gly Val Val Gly Pro	
180 185 190	
gtg gtc ggg gtt atg ggc act ttg cag gca ctg gaa gcc att aag tta	624
Val Val Gly Val Met Gly Thr Leu Gln Ala Leu Glu Ala Ile Lys Leu	
195 200 205	
tta agc ggt ata gag aca cct gcg gga gaa ctc cga ctg ttc gac ggt	672
Leu Ser Gly Ile Glu Thr Pro Ala Gly Glu Leu Arg Leu Phe Asp Gly	
210 215 220	
aaa tcg agc cag tgg cgc agc ctg gcg ttg cgc cgc gcc agt ggt tgc	720
Lys Ser Ser Gln Trp Arg Ser Leu Ala Leu Arg Arg Ala Ser Gly Cys	
225 230 235 240	
ccg gta tgc gga gga agc aat gca gat cct gtt taa	756
Pro Val Cys Gly Gly Ser Asn Ala Asp Pro Val	
245 250	

<210> 86

<211> 251

5 <212> PRT

<213> *Escherichia coli*

<400> 86

Met Asn Asp Arg Asp Phe Met Arg Tyr Ser Arg Gln Ile Leu Leu Asp
1 5 10 15

Asp Ile Ala Leu Asp Gly Gln Gln Lys Leu Leu Asp Ser Gln Val Leu
20 25 30

Ile Ile Gly Leu Gly Gly Leu Gly Thr Pro Ala Ala Leu Tyr Leu Ala
35 40 45

Gly Ala Gly Val Gly Thr Leu Val Leu Ala Asp Asp Asp Val His
50 55 60

10

ES 2 806 732 T3

Leu Ser Asn Leu Gln Arg Gln Ile Leu Phe Thr Thr Glu Asp Ile Asp
 65 70 75 80
 Arg Pro Lys Ser Gln Val Ser Gln Gln Arg Leu Thr Gln Leu Asn Pro
 85 90 95
 Asp Ile Gln Leu Thr Ala Leu Gln Gln Arg Leu Thr Gly Glu Ala Leu
 100 105 110
 Lys Asp Ala Val Ala Arg Ala Asp Val Val Leu Asp Cys Thr Asp Asn
 115 120 125
 Met Ala Thr Arg Gln Glu Ile Asn Ala Ala Cys Val Ala Leu Asn Thr
 130 135 140
 Pro Leu Ile Thr Ala Ser Ala Val Gly Phe Gly Gly Gln Leu Met Val
 145 150 155 160
 Leu Thr Pro Pro Trp Glu Gln Gly Cys Tyr Arg Cys Leu Trp Pro Asp
 165 170 175
 Asn Gln Glu Pro Glu Arg Asn Cys Arg Thr Ala Gly Val Val Gly Pro
 180 185 190
 Val Val Gly Val Met Gly Thr Leu Gln Ala Leu Glu Ala Ile Lys Leu
 195 200 205
 Leu Ser Gly Ile Glu Thr Pro Ala Gly Glu Leu Arg Leu Phe Asp Gly
 210 215 220
 Lys Ser Ser Gln Trp Arg Ser Leu Ala Leu Arg Arg Ala Ser Gly Cys
 225 230 235 240
 Pro Val Cys Gly Gly Ser Asn Ala Asp Pro Val
 245 250

<210> 87

<211> 201

5 <212> ADN

<213> *Escherichia coli*

<220>

<221> CDS

10 <222> (1)..(201)

<223> gen ThiS de *E. coli* que codifica para proteína portadora de azufre

<400> 87

15 atg cag atc ctg ttt aac gat caa gcg atg cag tgc gcc gcc ggg caa

48

ES 2 806 732 T3

	Met	Gln	Ile	Leu	Phe	Asn	Asp	Gln	Ala	Met	Gln	Cys	Ala	Ala	Gly	Gln	
	1				5					10					15		
	act	gtt	cac	gaa	cta	ctg	gag	caa	ctc	gac	caa	cga	caa	gcg	ggc	gcg	96
	Thr	Val	His	Glu	Leu	Leu	Glu	Gln	Leu	Asp	Gln	Arg	Gln	Ala	Gly	Ala	
				20					25					30			
	gct	ctg	gcg	att	aat	cag	caa	atc	gtc	ccg	cgt	gag	cag	tgg	gcg	caa	144
	Ala	Leu	Ala	Ile	Asn	Gln	Gln	Ile	Val	Pro	Arg	Glu	Gln	Trp	Ala	Gln	
				35				40					45				
	cat	atc	gtg	cag	gat	ggc	gac	cag	atc	ctg	ctt	ttt	cag	gtt	att	gca	192
	His	Ile	Val	Gln	Asp	Gly	Asp	Gln	Ile	Leu	Leu	Phe	Gln	Val	Ile	Ala	
		50					55					60					
	ggg	ggt	tga														201
	Gly	Gly															
	65																
	<210>	88															
	<211>	66															
5	<212>	PRT															
	<213>	<i>Escherichia coli</i>															
	<400>	88															
	Met	Gln	Ile	Leu	Phe	Asn	Asp	Gln	Ala	Met	Gln	Cys	Ala	Ala	Gly	Gln	
	1				5					10					15		
	Thr	Val	His	Glu	Leu	Leu	Glu	Gln	Leu	Asp	Gln	Arg	Gln	Ala	Gly	Ala	
				20					25					30			
	Ala	Leu	Ala	Ile	Asn	Gln	Gln	Ile	Val	Pro	Arg	Glu	Gln	Trp	Ala	Gln	
			35					40					45				
	His	Ile	Val	Gln	Asp	Gly	Asp	Gln	Ile	Leu	Leu	Phe	Gln	Val	Ile	Ala	
		50					55					60					
	Gly	Gly															
10	65																
	<210>	89															
	<211>	771															
	<212>	ADN															
15	<213>	<i>Escherichia coli</i>															
	<220>																
	<221>	CDS															
	<222>	(1)..(771)															
20	<223>	gen ThiG de <i>E. coli</i> que codifica para tiazol sintasa															
	<400>	89															
	atg	tta	cgt	att	gcg	gac	aaa	acg	ttt	gat	tca	cat	ctg	ttt	acc	ggc	48
	Met	Leu	Arg	Ile	Ala	Asp	Lys	Thr	Phe	Asp	Ser	His	Leu	Phe	Thr	Gly	
	1			5						10					15		

ES 2 806 732 T3

aca ggg aaa ttc gct tct tca caa ctg atg gtg gag gcg atc cgc gct Thr Gly Lys Phe Ala Ser Ser Gln Leu Met Val Glu Ala Ile Arg Ala 20 25 30	96
tcc ggc agc cag ctg gtg aca ctg gcg atg aaa cgt gtc gac ttg cgc Ser Gly Ser Gln Leu Val Thr Leu Ala Met Lys Arg Val Asp Leu Arg 35 40 45	144
cag cac aac gac gct atc ctc gaa ccg ctt atc gcg gcg ggt gtg acc Gln His Asn Asp Ala Ile Leu Glu Pro Leu Ile Ala Ala Gly Val Thr 50 55 60	192
ctg ctg cca aat aca tcc ggg gcg aaa aca gcg gaa gaa gcc att ttc Leu Leu Pro Asn Thr Ser Gly Ala Lys Thr Ala Glu Glu Ala Ile Phe 65 70 75 80	240
gcc gcc cat ctg gct cgt gaa gcg tta ggc aca aac tgg tta aaa tta Ala Ala His Leu Ala Arg Glu Ala Leu Gly Thr Asn Trp Leu Lys Leu 85 90 95	288
gag att cac cct gac gcc cgc tgg ctg ttg ccc gat ccc atc gaa acc Glu Ile His Pro Asp Ala Arg Trp Leu Leu Pro Asp Pro Ile Glu Thr 100 105 110	336
ctg aaa gcc gcc gaa acg ctg gta caa cag gga ttt gtc gtg ctg cct Leu Lys Ala Ala Glu Thr Leu Val Gln Gln Gly Phe Val Val Leu Pro 115 120 125	384
tac tgc ggg gcc gat ccg gta ttg tgt aaa cgt ctg gaa gaa gtc ggc Tyr Cys Gly Ala Asp Pro Val Leu Cys Lys Arg Leu Glu Glu Val Gly 130 135 140	432
tgt gca gcg gtg atg ccg ctc ggc gcg ccg att ggc tcg aat cag gga Cys Ala Ala Val Met Pro Leu Gly Ala Pro Ile Gly Ser Asn Gln Gly 145 150 155 160	480
ctg gaa acc cgc gcc atg ctg gag att att atc cag cag gcc aca gtg Leu Glu Thr Arg Ala Met Leu Glu Ile Ile Ile Gln Gln Ala Thr Val 165 170 175	528
ccg gtg gtt gtc gat gct ggc atc ggc gtt ccc agc cat gcc gcg cag Pro Val Val Val Asp Ala Gly Ile Gly Val Pro Ser His Ala Ala Gln 180 185 190	576
gcg ctg gaa atg ggg gcc gac gcg gtg tta gtg aat acg gcg att gcc Ala Leu Glu Met Gly Ala Asp Ala Val Leu Val Asn Thr Ala Ile Ala 195 200 205	624
gtc gcg gac gat ccc gtc aac atg gcg aag gca ttt cgt ctg gcg gta Val Ala Asp Asp Pro Val Asn Met Ala Lys Ala Phe Arg Leu Ala Val 210 215 220	672
gaa gca ggc cta ctg gca cgt cag tcc gga ccg ggc agc cgc agt tat Glu Ala Gly Leu Leu Ala Arg Gln Ser Gly Pro Gly Ser Arg Ser Tyr 225 230 235 240	720
ttt gct cat gcc acc agc ccg ctg acc gga ttt ctg gag gca tcg gca Phe Ala His Ala Thr Ser Pro Leu Thr Gly Phe Leu Glu Ala Ser Ala 245 250 255	768
tga	771

<210> 90

<211> 256

5 <212> PRT

<213> *Escherichia coli*

<400> 90

ES 2 806 732 T3

```

Met Leu Arg Ile Ala Asp Lys Thr Phe Asp Ser His Leu Phe Thr Gly
1          5          10          15

Thr Gly Lys Phe Ala Ser Ser Gln Leu Met Val Glu Ala Ile Arg Ala
          20          25          30

Ser Gly Ser Gln Leu Val Thr Leu Ala Met Lys Arg Val Asp Leu Arg
          35          40          45

Gln His Asn Asp Ala Ile Leu Glu Pro Leu Ile Ala Ala Gly Val Thr
          50          55          60

Leu Leu Pro Asn Thr Ser Gly Ala Lys Thr Ala Glu Glu Ala Ile Phe
65          70          75          80

Ala Ala His Leu Ala Arg Glu Ala Leu Gly Thr Asn Trp Leu Lys Leu
          85          90          95

Glu Ile His Pro Asp Ala Arg Trp Leu Leu Pro Asp Pro Ile Glu Thr
          100          105          110

Leu Lys Ala Ala Glu Thr Leu Val Gln Gln Gly Phe Val Val Leu Pro
          115          120          125

Tyr Cys Gly Ala Asp Pro Val Leu Cys Lys Arg Leu Glu Glu Val Gly
          130          135          140

Cys Ala Ala Val Met Pro Leu Gly Ala Pro Ile Gly Ser Asn Gln Gly
          145          150          155          160

Leu Glu Thr Arg Ala Met Leu Glu Ile Ile Ile Gln Gln Ala Thr Val
          165          170          175

Pro Val Val Val Asp Ala Gly Ile Gly Val Pro Ser His Ala Ala Gln
          180          185          190

Ala Leu Glu Met Gly Ala Asp Ala Val Leu Val Asn Thr Ala Ile Ala
          195          200          205

Val Ala Asp Asp Pro Val Asn Met Ala Lys Ala Phe Arg Leu Ala Val
          210          215          220
Glu Ala Gly Leu Leu Ala Arg Gln Ser Gly Pro Gly Ser Arg Ser Tyr
          225          230          235          240

Phe Ala His Ala Thr Ser Pro Leu Thr Gly Phe Leu Glu Ala Ser Ala
          245          250          255

```

<210> 91
 5 <211> 1134
 <212> ADN
 <213> *Escherichia coli*

<220>
 10 <221> CDS
 <222> (1)..(1134)
 <223> gen ThiH de *E. coli* que codifica para 2-iminoacetato sintasa

ES 2 806 732 T3

<400> 91

atg aaa acc ttc agc gat cgc tgg cga caa ctg gac tgg gac gac atc Met Lys Thr Phe Ser Asp Arg Trp Arg Gln Leu Asp Trp Asp Asp Ile 1 5 10 15	48
cgc ctg cgt atc aac ggc aaa acg gct gct gac gta gag cgg gcg cta Arg Leu Arg Ile Asn Gly Lys Thr Ala Ala Asp Val Glu Arg Ala Leu 20 25 30	96
aat gcc tcg caa ctc acc cgc gac gac atg atg gcg ctg tta tcg cct Asn Ala Ser Gln Leu Thr Arg Asp Asp Met Met Ala Leu Leu Ser Pro 35 40 45	144
gcc gcc agt ggc tat ctg gaa caa ctg gcc caa cgg gcg cag cgt ctg Ala Ala Ser Gly Tyr Leu Glu Gln Leu Ala Gln Arg Ala Gln Arg Leu 50 55 60	192
acc cgt cag cga ttt ggc aac aca gtt agt ttc tac gtc ccg ctt tat Thr Arg Gln Arg Phe Gly Asn Thr Val Ser Phe Tyr Val Pro Leu Tyr 65 70 75 80	240
ctt tcc aat ctt tgc gct aac gac tgc acg tac tgt gga ttt tcc atg Leu Ser Asn Leu Cys Ala Asn Asp Cys Thr Tyr Cys Gly Phe Ser Met 85 90 95	288
agt aat cgc atc aag cgc aaa acg ctg gat gaa gcg gat att gcc agg Ser Asn Arg Ile Lys Arg Lys Thr Leu Asp Glu Ala Asp Ile Ala Arg 100 105 110	336
gaa agt gcc gct ata cgg gag atg ggc ttt gaa cat ctg ctg tta gtc Glu Ser Ala Ala Ile Arg Glu Met Gly Phe Glu His Leu Leu Leu Val 115 120 125	384
act ggt gaa cat cag gcg aaa gtg ggg atg gat tac ttt cgt cgt cat Thr Gly Glu His Gln Ala Lys Val Gly Met Asp Tyr Phe Arg Arg His 130 135 140	432
ctc cct gcc ctt cgt gaa cag ttc tct tca cta cag atg gaa gtg caa Leu Pro Ala Leu Arg Glu Gln Phe Ser Ser Leu Gln Met Glu Val Gln 145 150 155 160	480
ccg ctg gcg gag acg gaa tac gcc gag tta aag caa ctt ggt ctg gat Pro Leu Ala Glu Thr Glu Tyr Ala Glu Leu Lys Gln Leu Gly Leu Asp 165 170 175	528

ES 2 806 732 T3

ggc gtg atg gtt tat cag gag aca tat cac gag gcg act tat gcc cgc 576
 Gly Val Met Val Tyr Gln Glu Thr Tyr His Glu Ala Thr Tyr Ala Arg
 180 185 190

cat cat ctg aaa ggc aaa aaa cag gac ttc ttc tgg cgg ctg gaa acg 624
 His His Leu Lys Gly Lys Lys Gln Asp Phe Phe Trp Arg Leu Glu Thr
 195 200 205

ccg gat cgg ctg ggg cgt gcg ggg att gat aag ata ggc ctc ggc gcg 672
 Pro Asp Arg Leu Gly Arg Ala Gly Ile Asp Lys Ile Gly Leu Gly Ala
 210 215 220

cta att ggc ctt tcc gac aac tgg cgc gtt gac agc tat atg gtt gcc 720
 Leu Ile Gly Leu Ser Asp Asn Trp Arg Val Asp Ser Tyr Met Val Ala
 225 230 235 240

gaa cat ttg cta tgg ctg caa cag cat tac tgg caa agc cgt tac tct 768
 Glu His Leu Leu Trp Leu Gln Gln His Tyr Trp Gln Ser Arg Tyr Ser
 245 250 255

gtc tcc ttt ccg cgc ctg cgc ccg tgt act ggc ggc att gag cct gcg 816
 Val Ser Phe Pro Arg Leu Arg Pro Cys Thr Gly Gly Ile Glu Pro Ala
 260 265 270

tcg att atg gat gaa cgc cag tta gtg caa acc atc tgc gcc ttc cga 864
 Ser Ile Met Asp Glu Arg Gln Leu Val Gln Thr Ile Cys Ala Phe Arg
 275 280 285

ctg ctt gca ccg gag att gaa ctg tca ctc tcc acg cgg gaa tca ccg 912
 Leu Leu Ala Pro Glu Ile Glu Leu Ser Leu Ser Thr Arg Glu Ser Pro
 290 295 300

tgg ttt cgc gat cgc gtt att ccg ctg gcg atc aat aac gtc agc gcc 960
 Trp Phe Arg Asp Arg Val Ile Pro Leu Ala Ile Asn Asn Val Ser Ala
 305 310 315 320

ttc tcg aaa acg cag cca ggt ggc tat gcc gat aat cac ccc gag ttg 1008
 Phe Ser Lys Thr Gln Pro Gly Gly Tyr Ala Asp Asn His Pro Glu Leu
 325 330 335

gaa cag ttc tca ccg cac gac gat cgc aga ccg gaa gcg gtt gct gcc 1056
 Glu Gln Phe Ser Pro His Asp Asp Arg Arg Pro Glu Ala Val Ala Ala
 340 345 350

gcg tta acc gct cag ggt ttg cag ccg gta tgg aaa gac tgg gac agc 1104
 Ala Leu Thr Ala Gln Gly Leu Gln Pro Val Trp Lys Asp Trp Asp Ser
 355 360 365

tat ctg gga cgc gcc tcg caa aga cta tga 1134
 Tyr Leu Gly Arg Ala Ser Gln Arg Leu
 370 375

<210> 92
 <211> 377
 5 <212> PRT
 <213> *Escherichia coli*

<400> 92

10 Met Lys Thr Phe Ser Asp Arg Trp Arg Gln Leu Asp Trp Asp Asp Ile
 1 5 10 15

ES 2 806 732 T3

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

ES 2 806 732 T3

	260		265		270	
	Ser Ile Met Asp Glu Arg Gln Leu Val Gln Thr Ile Cys Ala Phe Arg					
	275		280		285	
	Leu Leu Ala Pro Glu Ile Glu Leu Ser Leu Ser Thr Arg Glu Ser Pro					
	290		295		300	
	Trp Phe Arg Asp Arg Val Ile Pro Leu Ala Ile Asn Asn Val Ser Ala					
	305		310		315	320
	Phe Ser Lys Thr Gln Pro Gly Gly Tyr Ala Asp Asn His Pro Glu Leu					
		325		330		335
	Glu Gln Phe Ser Pro His Asp Asp Arg Arg Pro Glu Ala Val Ala Ala					
		340		345		350
	Ala Leu Thr Ala Gln Gly Leu Gln Pro Val Trp Lys Asp Trp Asp Ser					
		355		360		365
	Tyr Leu Gly Arg Ala Ser Gln Arg Leu					
	370		375			
	<210> 93					
	<211> 1110					
5	<212> ADN					
	<213> <i>Escherichia coli</i>					
	<220>					
	<221> CDS					
10	<222> (1)..(1110)					
	<223> gen Tio de <i>E. coli</i> que codifica para glicina oxidasa					
	<400> 93					
	atg aaa agg cat tat gaa gca gtg gtg att gga ggc gga att atc ggt					48
	Met Lys Arg His Tyr Glu Ala Val Val Ile Gly Gly Gly Ile Ile Gly					
	1	5	10	15		
	tcc gca att gct tat tat ttg gca aag gaa aac aaa aac acc gca ttg					96
	Ser Ala Ile Ala Tyr Tyr Leu Ala Lys Glu Asn Lys Asn Thr Ala Leu					
		20	25	30		
	ttt gaa agc gga aca atg ggc ggc aga acg aca agt gcc gct gcc gga					144
	Phe Glu Ser Gly Thr Met Gly Gly Arg Thr Thr Ser Ala Ala Ala Gly					
		35	40	45		
	atg ctg ggc gcc cat gcc gaa tgc gag gaa cgt gac gcg ttt ttt gat					192
	Met Leu Gly Ala His Ala Glu Cys Glu Glu Arg Asp Ala Phe Phe Asp					
		50	55	60		
	ttc gct atg cac agt cag cgt ctg tac aaa ggt ctt gga gaa gag ctt					240
15	Phe Ala Met His Ser Gln Arg Leu Tyr Lys Gly Leu Gly Glu Glu Leu					
	65	70	75	80		

ES 2 806 732 T3

tat gca tta tcc ggt gtg gat atc agg cag cat aac ggc ggt atg ttt Tyr Ala Leu Ser Gly Val Asp Ile Arg Gln His Asn Gly Gly Met Phe 85 90 95	288
aag ctt gca ttt tct gaa gaa gat gtg ctg cag ctg aga cag atg gac Lys Leu Ala Phe Ser Glu Glu Asp Val Leu Gln Leu Arg Gln Met Asp 100 105 110	336
gat ttg gac tct gtc agc tgg tat tca aaa gaa gag gtg tta gaa aaa Asp Leu Asp Ser Val Ser Trp Tyr Ser Lys Glu Glu Val Leu Glu Lys 115 120 125	384
gag ccg tat gcg tct ggt gac atc ttt ggt gca tct ttt att cag gat Glu Pro Tyr Ala Ser Gly Asp Ile Phe Gly Ala Ser Phe Ile Gln Asp 130 135 140	432
gat gtg cat gtg gag cct tat ttt gtt tgc aag gca tat gtg aaa gca Asp Val His Val Glu Pro Tyr Phe Val Cys Lys Ala Tyr Val Lys Ala 145 150 155 160	480
gca aaa atg ctt ggg gcg gag att ttt gag cat acg ccc gtc ctg cat Ala Lys Met Leu Gly Ala Glu Ile Phe Glu His Thr Pro Val Leu His 165 170 175	528
gtc gaa cgt gac ggt gaa gcc ctg ttc atc aag acc cct agc gga gac Val Glu Arg Asp Gly Glu Ala Leu Phe Ile Lys Thr Pro Ser Gly Asp 180 185 190	576
gta tgg gct aat cat gtt gtc gtt gcc agc ggg gtg tgg agc gga atg Val Trp Ala Asn His Val Val Val Ala Ser Gly Val Trp Ser Gly Met 195 200 205	624
ttt ttt aaa cag ctt gga ctg aac aat gct ttt ctc cct gta aaa ggg Phe Phe Lys Gln Leu Gly Leu Asn Asn Ala Phe Leu Pro Val Lys Gly 210 215 220	672
gag tgc ctg tcc gtt tgg aat gat gat atc ccg ctg aca aaa acg ctt Glu Cys Leu Ser Val Trp Asn Asp Asp Ile Pro Leu Thr Lys Thr Leu 225 230 235 240	720
tac cat gat cac tgc tat atc gta ccg aga aaa agc gga aga ctg gtt Tyr His Asp His Cys Tyr Ile Val Pro Arg Lys Ser Gly Arg Leu Val 245 250 255	768
gtc ggc gcg aca atg aag ccg ggg gac tgg agt gaa aca ccg gat ctt Val Gly Ala Thr Met Lys Pro Gly Asp Trp Ser Glu Thr Pro Asp Leu 260 265 270	816
ggc gga ttg gaa tct gtt atg aaa aaa gca aaa acg atg ctg ccg gct Gly Gly Leu Glu Ser Val Met Lys Lys Ala Lys Thr Met Leu Pro Ala 275 280 285	864
ata cag aat atg aag gtg gat cgt ttt tgg gcg gga ctc cgt ccg gga Ile Gln Asn Met Lys Val Asp Arg Phe Trp Ala Gly Leu Arg Pro Gly 290 295 300	912
aca aag gat gga aaa ccg tac atc ggc aga cat cct gag gac agc cgt Thr Lys Asp Gly Lys Pro Tyr Ile Gly Arg His Pro Glu Asp Ser Arg 305 310 315 320	960
att tta ttt gcg gct ggc cat ttc aga aac ggg atc ctg ctt gct ccc Ile Leu Phe Ala Ala Gly His Phe Arg Asn Gly Ile Leu Leu Ala Pro 325 330 335	1008
gca acg ggc gct ttg atc agt gat ctc atc atg aat aaa gag gtc aac Ala Thr Gly Ala Leu Ile Ser Asp Leu Ile Met Asn Lys Glu Val Asn 340 345 350	1056
caa gac tgg ctg cac gca ttc cga att gat cgc aag gag gcg gtt cag Gln Asp Trp Leu His Ala Phe Arg Ile Asp Arg Lys Glu Ala Val Gln 355 360 365	1104
ata tga Ile	1110

<210> 94

ES 2 806 732 T3

<211> 369
 <212> PRT
 <213> *Escherichia coli*

5 <400> 94

```

Met Lys Arg His Tyr Glu Ala Val Val Ile Gly Gly Gly Ile Ile Gly
 1              5              10              15

Ser Ala Ile Ala Tyr Tyr Leu Ala Lys Glu Asn Lys Asn Thr Ala Leu
      20              25              30

Phe Glu Ser Gly Thr Met Gly Gly Arg Thr Thr Ser Ala Ala Ala Gly
      35              40              45

Met Leu Gly Ala His Ala Glu Cys Glu Glu Arg Asp Ala Phe Phe Asp
 50              55              60

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

Tyr Ala Leu Ser Gly Val Asp Ile Arg Gln His Asn Gly Gly Met Phe
      85              90              95

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

Asp Leu Asp Ser Val Ser Trp Tyr Ser Lys Glu Glu Val Leu Glu Lys
      115              120              125

Glu Pro Tyr Ala Ser Gly Asp Ile Phe Gly Ala Ser Phe Ile Gln Asp
      130              135              140

Asp Val His Val Glu Pro Tyr Phe Val Cys Lys Ala Tyr Val Lys Ala
      145              150              155              160

Ala Lys Met Leu Gly Ala Glu Ile Phe Glu His Thr Pro Val Leu His
      165              170              175
  
```

ES 2 806 732 T3

Val Glu Arg Asp Gly Glu Ala Leu Phe Ile Lys Thr Pro Ser Gly Asp
180 185 190

Val Trp Ala Asn His Val Val Val Ala Ser Gly Val Trp Ser Gly Met
195 200 205

Phe Phe Lys Gln Leu Gly Leu Asn Asn Ala Phe Leu Pro Val Lys Gly
210 215 220

Glu Cys Leu Ser Val Trp Asn Asp Asp Ile Pro Leu Thr Lys Thr Leu
225 230 235 240

Tyr His Asp His Cys Tyr Ile Val Pro Arg Lys Ser Gly Arg Leu Val
245 250 255

Val Gly Ala Thr Met Lys Pro Gly Asp Trp Ser Glu Thr Pro Asp Leu
260 265 270

Gly Gly Leu Glu Ser Val Met Lys Lys Ala Lys Thr Met Leu Pro Ala
275 280 285

Ile Gln Asn Met Lys Val Asp Arg Phe Trp Ala Gly Leu Arg Pro Gly
290 295 300

Thr Lys Asp Gly Lys Pro Tyr Ile Gly Arg His Pro Glu Asp Ser Arg
305 310 315 320

Ile Leu Phe Ala Ala Gly His Phe Arg Asn Gly Ile Leu Leu Ala Pro
325 330 335

Ala Thr Gly Ala Leu Ile Ser Asp Leu Ile Met Asn Lys Glu Val Asn
340 345 350

Gln Asp Trp Leu His Ala Phe Arg Ile Asp Arg Lys Glu Ala Val Gln
355 360 365

Ile

<210> 95

<211> 1098

5 <212> ADN

<213> *Pseudomonas putida*

<220>

<221> CDS

10 <222> (1)..(1098)

<223> gen Tio de *Pseudomonas putida* que codifica para una glicina oxidasa

<400> 95

ES 2 806 732 T3

atg agc aag caa gta gtg gtg gtc ggt ggc ggg gtc att ggc ctg ctg Met Ser Lys Gln Val Val Val Val Gly Gly Gly Val Ile Gly Leu Leu 1 5 10 15	48
acg gca ttc aac ctg gcg gcg agc gtc gac cag gtg gtg gta tgc gac Thr Ala Phe Asn Leu Ala Ala Ser Val Asp Gln Val Val Val Cys Asp 20 25 30	96
cag ggc gaa gta ggg cgc gag tcc tcc tgg gct ggg ggc ggt atc gtc Gln Gly Glu Val Gly Arg Glu Ser Ser Trp Ala Gly Gly Gly Ile Val 35 40 45	144
tcg ccc ctg tat cct tgg cgc tac agc ccg gca gtg acc gcc ctg gcg Ser Pro Leu Tyr Pro Trp Arg Tyr Ser Pro Ala Val Thr Ala Leu Ala 50 55 60	192
cat tgg tcg cag gac ttt tac cca cag ttg ggc gag cgc ttg ttc gcc His Trp Ser Gln Asp Phe Tyr Pro Gln Leu Gly Glu Arg Leu Phe Ala 65 70 75 80	240
agc acg ggc ctg gat cct gag gtg cat acc acc ggg ctt tac tgg ctc Ser Thr Gly Leu Asp Pro Glu Val His Thr Thr Gly Leu Tyr Trp Leu 85 90 95	288
gac ctg gat gac caa gcc cag gcc ttg gcg tgg gca ggc cgt cag cag Asp Leu Asp Asp Gln Ala Gln Ala Leu Ala Trp Ala Gly Arg Gln Gln 100 105 110	336
cgt ccg ctg agc gcc gtg gat att tca gcg gtg tac gac gca gtc cct Arg Pro Leu Ser Ala Val Asp Ile Ser Ala Val Tyr Asp Ala Val Pro 115 120 125	384
gtg ctg ggg cca ggc ttt gag cga gcc ctc tac atg gaa ggc gtg gcc Val Leu Gly Pro Gly Phe Glu Arg Ala Leu Tyr Met Glu Gly Val Ala 130 135 140	432
aat gtg cgc aac ccg cgc ctg gtc aaa tcg ctg aag gcg gcg ttg ctg Asn Val Arg Asn Pro Arg Leu Val Lys Ser Leu Lys Ala Ala Leu Leu 145 150 155 160	480
gca ttg ccc aat gtg agc gtg cgc gag cac tgc cag atc acg ggc ttc Ala Leu Pro Asn Val Ser Val Arg Glu His Cys Gln Ile Thr Gly Phe 165 170 175	528
gtg cag cag ggc gct cgt atc att ggg gtg agc acc gct gaa ggc gag Val Gln Gln Gly Ala Arg Ile Ile Gly Val Ser Thr Ala Glu Gly Glu 180 185 190	576
ctg gcc gcc gac gaa gtc gta ctg agc gcc ggt gcc tgg agc ggc gaa Leu Ala Ala Asp Glu Val Val Leu Ser Ala Gly Ala Trp Ser Gly Glu 195 200 205	624
ctg ctg cgc cac ttg ggc ctt gag ctt cca gtc gag ccg gta aaa ggg Leu Leu Arg His Leu Gly Leu Glu Leu Pro Val Glu Pro Val Lys Gly 210 215 220	672
cag atg atc ctg ttc aaa tgc gct gaa gat ttt ctg cca agc atg gtg Gln Met Ile Leu Phe Lys Cys Ala Glu Asp Phe Leu Pro Ser Met Val 225 230 235 240	720

ES 2 806 732 T3

ctt gcc aaa ggt cgt tat gca att ccg cgt cgg gat ggt cac att ctg 768
 Leu Ala Lys Gly Arg Tyr Ala Ile Pro Arg Arg Asp Gly His Ile Leu
 245 250 255
 gtg gcc agc acg ctg gag cat gcc gcc tac gac aag aca ccc acc gat 816
 Val Gly Ser Thr Leu Glu His Ala Gly Tyr Asp Lys Thr Pro Thr Asp
 260 265 270
 gag gcc ttg gcc agc ctc aag gca tcg gcg gtg gat ctg ctg ccc ggc 864
 Glu Ala Leu Ala Ser Leu Lys Ala Ser Ala Val Asp Leu Leu Pro Gly
 275 280 285
 ctg gaa gcc gcg cac gtg gtt gcc cac tgg gcc ggg ctg cgg cca ggt 912
 Leu Glu Gly Ala His Val Val Ala His Trp Ala Gly Leu Arg Pro Gly
 290 295 300
 tcg cca gaa gcc gtt ccg ttt atc ggg ccg gta ccc gcc ttc gat ggc 960
 Ser Pro Glu Gly Val Pro Phe Ile Gly Pro Val Pro Gly Phe Asp Gly
 305 310 315 320
 tta tgg ctg aac tgc gcc cat tac cga aac ggg ctg gtg ctg gcg ccc 1008
 Leu Trp Leu Asn Cys Gly His Tyr Arg Asn Gly Leu Val Leu Ala Pro
 325 330 335
 gct tcg tgc caa ctg ctg gcc gat ttg ctc aat gcc gcc gag ccc atc 1056
 Ala Ser Cys Gln Leu Leu Ala Asp Leu Leu Asn Gly Ala Glu Pro Ile
 340 345 350
 atc gac ccg tca ccc tac gcc ccg tct ggg cgc ctt gcc taa 1098
 Ile Asp Pro Ser Pro Tyr Ala Pro Ser Gly Arg Leu Gly
 355 360 365

<210> 96

<211> 365

5 <212> PRT

<213> *Pseudomonas putida*

<400> 96

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

10

ES 2 806 732 T3

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

<210> 97
 5 <211> 1140
 <212> ADN
 <213> *Synechococcus elongatus*

<220>
 10 <221> CDS
 <222> (1)..(1140)

ES 2 806 732 T3

<223> gen Tio de *Synechococcus elongatus* que codifica para una glicina oxidasa

<400> 97

atg gcg ttc gag gta gcc gtc ttt ggg ggc ggc gtc att ggc ttg gcg	48
Met Ala Phe Glu Val Ala Val Phe Gly Gly Gly Val Ile Gly Leu Ala	
1 5 10 15	
atc gcg cta gaa ctg cga tcg cga ggc ggc atg gtg cag gtc tac agt	96
Ile Ala Leu Glu Leu Arg Ser Arg Gly Ala Met Val Gln Val Tyr Ser	
20 25 30	
caa aac act cag gcg gcg gca ggt cgt gtg gca gca ggg atg ttg gcg	144
Gln Asn Thr Gln Ala Ala Ala Gly Arg Val Ala Ala Gly Met Leu Ala	
35 40 45	
ccc cag tcg gaa ggc atc gaa gtc ggg ccc atg ctg gat ctg ggg ctg	192
Pro Gln Ser Glu Gly Ile Glu Val Gly Pro Met Leu Asp Leu Gly Leu	
50 55 60	
cgc agc cga tcg ctc tac gcc cgc tgg acc cag caa ctc gaa caa ctc	240
Arg Ser Arg Ser Leu Tyr Ala Arg Trp Thr Gln Gln Leu Glu Gln Leu	
65 70 75 80	
agc ggt caa gac agt ggc tac tgg ccc tgc ggc att ttg gtg ccc ctg	288
Ser Gly Gln Asp Ser Gly Tyr Trp Pro Cys Gly Ile Leu Val Pro Leu	
85 90 95	
agt gag gcc aaa aat cgc gat cgc tat cct cat cca gca gaa tct ccg	336
Ser Glu Ala Lys Asn Arg Asp Arg Tyr Pro His Pro Ala Glu Ser Pro	
100 105 110	
ggg caa tgg ctc tcg gca gcg gac tta cga gac ttt cag ccc gca cta	384
Gly Gln Trp Leu Ser Ala Ala Asp Leu Arg Asp Phe Gln Pro Ala Leu	
115 120 125	
tgc tct gac cta atc ggt ggc tgg tgg ttt tcc caa gaa ggg caa gtt	432
Cys Ser Asp Leu Ile Gly Gly Trp Trp Phe Ser Gln Glu Gly Gln Val	
130 135 140	
gat agt cgc cgt gcc ctg tat cca gcg ctg cga gcc gcc gcg atc gcc	480
Asp Ser Arg Arg Ala Leu Tyr Pro Ala Leu Arg Ala Ala Ala Ile Ala	
145 150 155 160	
agt ggc gtc acg atc cat gaa agc gtg gcg ctg cgg gag tta tct gta	528
Ser Gly Val Thr Ile His Glu Ser Val Ala Leu Arg Glu Leu Ser Val	
165 170 175	
5 aca ggc gat cgc ctg caa tcc gcg atg acc gat cgc ggg cca gtt caa	576

ES 2 806 732 T3

Thr	Gly	Asp	Arg	Leu	Gln	Ser	Ala	Met	Thr	Asp	Arg	Gly	Pro	Val	Gln		
			180					185					190				
gct	gac	gcc	tac	gtt	ctg	gca	acc	ggc	gct	tgg	tcc	ggc	gac	tgg	cta		624
Ala	Asp	Ala	Tyr	Val	Leu	Ala	Thr	Gly	Ala	Trp	Ser	Gly	Asp	Trp	Leu		
		195					200					205					
caa	ctg	ccg	gtc	tat	ccc	gtt	aaa	ggc	caa	atg	ttc	tcg	ctg	caa	gct		672
Gln	Leu	Pro	Val	Tyr	Pro	Val	Lys	Gly	Gln	Met	Phe	Ser	Leu	Gln	Ala		
	210					215					220						
gac	ccg	cgt	ttg	ctg	aac	cac	gtt	ttg	ttt	ggg	gag	cgg	gtg	tat	att		720
Asp	Pro	Arg	Leu	Leu	Asn	His	Val	Leu	Phe	Gly	Glu	Arg	Val	Tyr	Ile		
225					230					235					240		
gtg	ccg	cgc	cga	gat	ggg	ctg	att	gtg	gtc	ggg	gcc	acc	atg	gaa	gcg		768
Val	Pro	Arg	Arg	Asp	Gly	Leu	Ile	Val	Val	Gly	Ala	Thr	Met	Glu	Ala		
			245						250					255			
acg	gcg	gga	ttc	agg	act	ggc	aac	acc	gct	ggc	ccc	tta	cag	agc	ttg		816
Thr	Ala	Gly	Phe	Arg	Thr	Gly	Asn	Thr	Ala	Gly	Pro	Leu	Gln	Ser	Leu		
		260					265						270				
atg	gcc	gag	gcg	atc	gcc	ctc	gtt	ccg	gct	ctg	gcg	gac	tgt	cca	ctg		864
Met	Ala	Glu	Ala	Ile	Ala	Leu	Val	Pro	Ala	Leu	Ala	Asp	Cys	Pro	Leu		
	275					280						285					
gtt	gaa	act	tgg	tgg	gga	tac	cgt	ccc	gcg	aca	cca	gat	gaa	tgg	ccg		912
Val	Glu	Thr	Trp	Trp	Gly	Tyr	Arg	Pro	Ala	Thr	Pro	Asp	Glu	Trp	Pro		
	290					295					300						
atc	ctg	ggg	caa	ggc	ccc	gct	gag	aac	tta	ttc	ttg	gcg	acc	ggc	cac		960
Ile	Leu	Gly	Gln	Gly	Pro	Ala	Glu	Asn	Leu	Phe	Leu	Ala	Thr	Gly	His		
305					310					315					320		
tac	cgc	aac	ggg	atg	ctg	ctc	gcc	cca	att	acc	gct	cag	cta	ctc	gct		1008
Tyr	Arg	Asn	Gly	Met	Leu	Leu	Ala	Pro	Ile	Thr	Ala	Gln	Leu	Leu	Ala		
			325						330					335			
gac	caa	att	ctc	gac	cac	tgc	acg	gat	caa	ctg	ctt	cat	gcc	ttc	cgt		1056
Asp	Gln	Ile	Leu	Asp	His	Cys	Thr	Asp	Gln	Leu	Leu	His	Ala	Phe	Arg		
		340					345						350				
tac	gac	cgc	ttc	tcc	agc	cat	gac	tcc	agc	acc	cat	caa	ccc	tta	ccc		1104
Tyr	Asp	Arg	Phe	Ser	Ser	His	Asp	Ser	Ser	Thr	His	Gln	Pro	Leu	Pro		
	355					360						365					
gct	ctt	gca	ggc	ttg	tca	gcg	tca	acg	ggg	cag	tga						1140
Ala	Leu	Ala	Gly	Leu	Ser	Ala	Ser	Thr	Gly	Gln							
	370					375											

<210> 98

<211> 379

5 <212> PRT

<213> *Synechococcus elongatus*

<400> 98

Met	Ala	Phe	Glu	Val	Ala	Val	Phe	Gly	Gly	Gly	Val	Ile	Gly	Leu	Ala
10	1			5				10						15	

ES 2 806 732 T3

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

ES 2 806 732 T3

Met Ala Glu Ala Ile Ala Leu Val Pro Ala Leu Ala Asp Cys Pro Leu
275 280 285

Val Glu Thr Trp Trp Gly Tyr Arg Pro Ala Thr Pro Asp Glu Trp Pro
290 295 300

Ile Leu Gly Gln Gly Pro Ala Glu Asn Leu Phe Leu Ala Thr Gly His
305 310 315 320

Tyr Arg Asn Gly Met Leu Leu Ala Pro Ile Thr Ala Gln Leu Leu Ala
325 330 335

Asp Gln Ile Leu Asp His Cys Thr Asp Gln Leu Leu His Ala Phe Arg
340 345 350

Tyr Asp Arg Phe Ser Ser His Asp Ser Ser Thr His Gln Pro Leu Pro
355 360 365

Ala Leu Ala Gly Leu Ser Ala Ser Thr Gly Gln
370 375

<210> 99

<211> 801

5 <212> ADN

<213> *Escherichia coli*

<220>

<221> CDS

10 <222> (1)..(801)

<223> gen ThiD de *E. coli* que codifica para fosfohidroximetilpirimidina cinasa

<400> 99

atg aaa cga att aac gct ctg acg att gcc ggt act gat ccg agt ggt 48
Met Lys Arg Ile Asn Ala Leu Thr Ile Ala Gly Thr Asp Pro Ser Gly
1 5 10 15

ggt gcg ggg att cag gcc gat ctt aaa acc ttc tcg gca ctt ggc gct 96
Gly Ala Gly Ile Gln Ala Asp Leu Lys Thr Phe Ser Ala Leu Gly Ala
20 25 30

tat ggt tgc tca gtt att act gca ctg gtg gcg caa aat acc cgt ggc 144
Tyr Gly Cys Ser Val Ile Thr Ala Leu Val Ala Gln Asn Thr Arg Gly
35 40 45

gta cag tcg gtg tat cgc att gag cct gat ttt gtc gcc gcc cag ctc 192
Val Gln Ser Val Tyr Arg Ile Glu Pro Asp Phe Val Ala Ala Gln Leu
50 55 60

gat tcg gtg ttc agc gat gtg cga atc gat acc act aaa atc ggt atg 240
Asp Ser Val Phe Ser Asp Val Arg Ile Asp Thr Thr Lys Ile Gly Met
65 70 75 80

15 ctg gcg gaa acc gat att gtt gaa gcg gtg gca gaa cgg ttg caa cgt 288

ES 2 806 732 T3

```

Leu Ala Glu Thr Asp Ile Val Glu Ala Val Ala Glu Arg Leu Gln Arg
      85                      90                      95

tat cag atc caa aac gtg gta ctc gac acc gtt atg ctg gca aaa agc      336
Tyr Gln Ile Gln Asn Val Val Leu Asp Thr Val Met Leu Ala Lys Ser
      100                      105                      110

ggc gac ccg ctg ctt tca cct tcg gcg gtt gct acg ctg cgc agt cga      384
Gly Asp Pro Leu Leu Ser Pro Ser Ala Val Ala Thr Leu Arg Ser Arg
      115                      120                      125

tta ttg cca cag gtt tca tta ata acg cca aac ttg ccc gaa gct gcc      432
Leu Leu Pro Gln Val Ser Leu Ile Thr Pro Asn Leu Pro Glu Ala Ala
      130                      135                      140

gcc ttg ctc gac gcg cca cac gcg cgc acc gaa cag gaa atg ctg gaa      480
Ala Leu Leu Asp Ala Pro His Ala Arg Thr Glu Gln Glu Met Leu Glu
      145                      150                      155                      160

caa ggg cga tcg ctg ttg gcg atg ggc tgt ggc gca gtg cta atg aaa      528
Gln Gly Arg Ser Leu Leu Ala Met Gly Cys Gly Ala Val Leu Met Lys
      165                      170                      175

ggt ggt cat ctg gat gat gag caa agc ccg gac tgg ctg ttt acc cgc      576
Gly Gly His Leu Asp Asp Glu Gln Ser Pro Asp Trp Leu Phe Thr Arg
      180                      185                      190

gag ggt gaa caa cgg ttt acc gca ccg cgc att atg acc aaa aac acc      624
Glu Gly Glu Gln Arg Phe Thr Ala Pro Arg Ile Met Thr Lys Asn Thr
      195                      200                      205

cac ggc act ggt tgt aca ctc tct gcg gcg ttg gct gca cta cgc ccg      672
His Gly Thr Gly Cys Thr Leu Ser Ala Ala Leu Ala Ala Leu Arg Pro
      210                      215                      220

cgc cat aca aac tgg gct gac acc gta cag gag gca aaa agc tgg ctt      720
Arg His Thr Asn Trp Ala Asp Thr Val Gln Glu Ala Lys Ser Trp Leu
      225                      230                      235                      240

tca tcg gcg tta gcc cag gcc gac acg ctg gaa gtt ggt cac ggt att      768
Ser Ser Ala Leu Ala Gln Ala Asp Thr Leu Glu Val Gly His Gly Ile
      245                      250                      255

ggt ccg gtt cac cac ttc cac gcc tgg tgg tga      801
Gly Pro Val His His Phe His Ala Trp Trp
      260                      265

```

<210> 100
 <211> 266
 5 <212> PRT
 <213> *Escherichia coli*

<400> 100

```

Met Lys Arg Ile Asn Ala Leu Thr Ile Ala Gly Thr Asp Pro Ser Gly
1          5          10          15

```

```

10 Gly Ala Gly Ile Gln Ala Asp Leu Lys Thr Phe Ser Ala Leu Gly Ala
    20          25          30

```

ES 2 806 732 T3

```

Tyr Gly Cys Ser Val Ile Thr Ala Leu Val Ala Gln Asn Thr Arg Gly
    35                                40                                45

Val Gln Ser Val Tyr Arg Ile Glu Pro Asp Phe Val Ala Ala Gln Leu
    50                                55                                60

Asp Ser Val Phe Ser Asp Val Arg Ile Asp Thr Thr Lys Ile Gly Met
    65                                70                                75                                80

Leu Ala Glu Thr Asp Ile Val Glu Ala Val Ala Glu Arg Leu Gln Arg
    85                                90                                95

Tyr Gln Ile Gln Asn Val Val Leu Asp Thr Val Met Leu Ala Lys Ser
    100                                105                                110

Gly Asp Pro Leu Leu Ser Pro Ser Ala Val Ala Thr Leu Arg Ser Arg
    115                                120                                125

Leu Leu Pro Gln Val Ser Leu Ile Thr Pro Asn Leu Pro Glu Ala Ala
    130                                135                                140

Ala Leu Leu Asp Ala Pro His Ala Arg Thr Glu Gln Glu Met Leu Glu
    145                                150                                155                                160

Gln Gly Arg Ser Leu Leu Ala Met Gly Cys Gly Ala Val Leu Met Lys
    165                                170                                175

Gly Gly His Leu Asp Asp Glu Gln Ser Pro Asp Trp Leu Phe Thr Arg
    180                                185                                190

Glu Gly Glu Gln Arg Phe Thr Ala Pro Arg Ile Met Thr Lys Asn Thr
    195                                200                                205

His Gly Thr Gly Cys Thr Leu Ser Ala Ala Leu Ala Ala Leu Arg Pro
    210                                215                                220

Arg His Thr Asn Trp Ala Asp Thr Val Gln Glu Ala Lys Ser Trp Leu
    225                                230                                235                                240

Ser Ser Ala Leu Ala Gln Ala Asp Thr Leu Glu Val Gly His Gly Ile
    245                                250                                255

Gly Pro Val His His Phe His Ala Trp Trp
    260                                265

```

<210> 101

<211> 984

5 <212> ADN

<213> *Escherichia coli*

<220>

<221> CDS

10 <222> (1)..(984)

<223> gen ThiB de *E. coli* que codifica para una proteína de unión periplásmica de transportadores ABC de tiamina

<400> 101

ES 2 806 732 T3

atg tta aaa aaa tgt ctg ccc ctg ctg ttg ctg tgc aca gcg ccc gtt Met Leu Lys Lys Cys Leu Pro Leu Leu Leu Leu Cys Thr Ala Pro Val 1 5 10 15	48
ttc gct aaa ccc gtt ctg act gtt tat acc tac gat tcc ttc gcc gcc Phe Ala Lys Pro Val Leu Thr Val Tyr Thr Tyr Asp Ser Phe Ala Ala 20 25 30	96
gac tgg ggg cct ggt ccg gtg gtt aaa aaa gcc ttt gaa gcc gac tgt Asp Trp Gly Pro Gly Pro Val Val Lys Lys Ala Phe Glu Ala Asp Cys 35 40 45	144
aat tgc gaa ctg aaa ctg gtg gcg ctg gaa gat ggc gtt tcg ctt ctc Asn Cys Glu Leu Lys Leu Val Ala Leu Glu Asp Gly Val Ser Leu Leu 50 55 60	192
aac cgt cta cgg atg gaa ggc aaa aac agt aaa gcc gat gtg gtg ctg Asn Arg Leu Arg Met Glu Gly Lys Asn Ser Lys Ala Asp Val Val Leu 65 70 75 80	240
ggg ctg gat aac aac ctg tta gac gcc gcc agt aaa acc gga ctg ttt Gly Leu Asp Asn Asn Leu Leu Asp Ala Ala Ser Lys Thr Gly Leu Phe 85 90 95	288
gcc aaa agc ggt gtg gca gcg gat gcc gtt aac gtt ccc ggc ggc tgg Ala Lys Ser Gly Val Ala Ala Asp Ala Val Asn Val Pro Gly Gly Trp 100 105 110	336
aat aat gac act ttc gta ccg ttt gat tat ggc tac ttc gcc ttc gtt Asn Asn Asp Thr Phe Val Pro Phe Asp Tyr Gly Tyr Phe Ala Phe Val 115 120 125	384
tat gac aag aac aaa ctg aaa aac ccg cca caa agc ctg aaa gaa ctg Tyr Asp Lys Asn Lys Leu Lys Asn Pro Pro Gln Ser Leu Lys Glu Leu 130 135 140	432
gtt gag agc gat caa aac tgg cgg gtg att tat cag gat ccg cgc acc Val Glu Ser Asp Gln Asn Trp Arg Val Ile Tyr Gln Asp Pro Arg Thr 145 150 155 160	480
agt aca ccg ggg ctg ggt ctg ttg cta tgg atg caa aaa gtc tat ggc Ser Thr Pro Gly Leu Gly Leu Leu Leu Trp Met Gln Lys Val Tyr Gly 165 170 175	528
gat gac gcc cca caa gcc tgg cag aaa ctg gcg aag aaa acg gtc acg Asp Asp Ala Pro Gln Ala Trp Gln Lys Leu Ala Lys Lys Thr Val Thr 180 185 190	576
gtc acc aaa ggc tgg agc gaa gcc tac ggc ctg ttt tta aaa ggt gaa Val Thr Lys Gly Trp Ser Glu Ala Tyr Gly Leu Phe Leu Lys Gly Glu 195 200 205	624
agc gat ctg gta ctg agt tac acc acc tct ccg gct tat cac att ctc	672

ES 2 806 732 T3

Ser	Asp	Leu	Val	Leu	Ser	Tyr	Thr	Thr	Ser	Pro	Ala	Tyr	His	Ile	Leu	
210						215					220					
gaa	gag	aag	aaa	gat	aac	tac	gcc	gcc	gcg	aac	ttc	agc	gaa	ggt	cac	720
Glu	Glu	Lys	Lys	Asp	Asn	Tyr	Ala	Ala	Ala	Asn	Phe	Ser	Glu	Gly	His	
225					230					235					240	
tat	ctg	caa	gtg	gaa	gtc	gcc	gcc	cgc	acc	gct	gcc	agc	aag	cag	ccg	768
Tyr	Leu	Gln	Val		Val	Ala	Ala	Arg	Thr	Ala	Ala	Ser	Lys	Gln	Pro	
				245					250						255	
gag	ctg	gcg	caa	aaa	ttc	ctc	cag	ttt	atg	gtt	tct	ccg	gct	ttc	cag	816
Glu	Leu	Ala	Gln	Lys	Phe	Leu	Gln	Phe	Met	Val	Ser	Pro	Ala	Phe	Gln	
			260					265						270		
aat	gcg	atc	cca	acc	ggc	aac	tgg	atg	tat	ccg	gtg	gca	aac	gtc	acg	864
Asn	Ala	Ile	Pro	Thr	Gly	Asn	Trp	Met	Tyr	Pro	Val	Ala	Asn	Val	Thr	
		275					280						285			
ctg	cct	gcc	ggt	ttt	gaa	aaa	ttg	acc	aaa	ccc	gca	acc	acg	ttg	gag	912
Leu	Pro	Ala	Gly	Phe	Glu	Lys	Leu	Thr	Lys	Pro	Ala	Thr	Thr	Leu	Glu	
	290					295					300					
ttc	acg	cca	gcc	gaa	gtg	gcg	gca	caa	cgt	cag	gca	tgg	att	agc	gaa	960
Phe	Thr	Pro	Ala	Glu	Val	Ala	Ala	Gln	Arg	Gln	Ala	Trp	Ile	Ser	Glu	
305					310					315					320	
tgg	caa	cgc	gcc	gtc	agc	cgt	taa									984
Trp	Gln	Arg	Ala	Val	Ser	Arg										
				325												

<210> 102
 <211> 327
 5 <212> PRT
 <213> *Escherichia coli*

<400> 102

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

10

ES 2 806 732 T3

Ala Lys Ser Gly Val Ala Ala Asp Ala Val Asn Val Pro Gly Gly Trp
100 105 110

Asn Asn Asp Thr Phe Val Pro Phe Asp Tyr Gly Tyr Phe Ala Phe Val
115 120 125

Tyr Asp Lys Asn Lys Leu Lys Asn Pro Pro Gln Ser Leu Lys Glu Leu
130 135 140

Val Glu Ser Asp Gln Asn Trp Arg Val Ile Tyr Gln Asp Pro Arg Thr
145 150 155 160

Ser Thr Pro Gly Leu Gly Leu Leu Leu Trp Met Gln Lys Val Tyr Gly
165 170 175

Asp Asp Ala Pro Gln Ala Trp Gln Lys Leu Ala Lys Lys Thr Val Thr
180 185 190

Val Thr Lys Gly Trp Ser Glu Ala Tyr Gly Leu Phe Leu Lys Gly Glu
195 200 205

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

Glu Glu Lys Lys Asp Asn Tyr Ala Ala Ala Asn Phe Ser Glu Gly His
225 230 235 240

Tyr Leu Gln Val Glu Val Ala Ala Arg Thr Ala Ala Ser Lys Gln Pro
245 250 255

Glu Leu Ala Gln Lys Phe Leu Gln Phe Met Val Ser Pro Ala Phe Gln
260 265 270

Asn Ala Ile Pro Thr Gly Asn Trp Met Tyr Pro Val Ala Asn Val Thr
275 280 285

Leu Pro Ala Gly Phe Glu Lys Leu Thr Lys Pro Ala Thr Thr Leu Glu
290 295 300

Phe Thr Pro Ala Glu Val Ala Ala Gln Arg Gln Ala Trp Ile Ser Glu
305 310 315 320

Trp Gln Arg Ala Val Ser Arg
325

<210> 103

<211> 1611

5 <212> ADN

<213> *Escherichia coli*

<220>

<221> CDS

10 <222> (1)..(1611)

<223> gen ThiP de *E. coli* que codifica para una permeasa de transportadores ABC de tiamina

<400> 103

ES 2 806 732 T3

atg gca acg cgc cgt cag ccg tta att ccc ggc tgg tta att cca ggt Met Ala Thr Arg Arg Gln Pro Leu Ile Pro Gly Trp Leu Ile Pro Gly 1 5 10 15	48
gta agc gcc acc acg ctg gtg gta gcg gtt gcg ctg gcg gcg ttt ctc Val Ser Ala Thr Thr Leu Val Val Ala Val Ala Leu Ala Ala Phe Leu 20 25 30	96
gcc ctg tgg tgg aac gcg ccg cag gat gac tgg gtg gca gtc tgg cag Ala Leu Trp Trp Asn Ala Pro Gln Asp Asp Trp Val Ala Val Trp Gln 35 40 45	144
gac agc tat ctg tgg cat gtg gtg cgc ttc tcc ttc tgg cag gcg ttt Asp Ser Tyr Leu Trp His Val Arg Phe Ser Phe Trp Gln Ala Phe 50 55 60	192
ctc tcg gca ctg ctc tct gtc ata ccc gcg ata ttc ctc gcc cgc gcg Leu Ser Ala Leu Leu Ser Val Ile Pro Ala Ile Phe Leu Ala Arg Ala 65 70 75 80	240
ctc tat cgc agg cgc ttt ccg ggt cgg ctg gcg ctg ttg cgt ctg tgt Leu Tyr Arg Arg Arg Phe Pro Gly Arg Leu Ala Leu Leu Arg Leu Cys 85 90 95	288
gca atg acc ttg atc ctc ccg gtg ttg gtc gct gtt ttc ggc att ctt Ala Met Thr Leu Ile Leu Pro Val Leu Val Ala Val Phe Gly Ile Leu 100 105 110	336
agc gtc tat ggt cgc cag ggc tgg ctg gca aca ctc tgc caa tcg ctc Ser Val Tyr Gly Arg Gln Gly Trp Leu Ala Thr Leu Cys Gln Ser Leu 115 120 125	384
ggg ctg gag tgg acc ttt tcg ccc tac ggc ctg caa ggt att ttg ctg Gly Leu Glu Trp Thr Phe Ser Pro Tyr Gly Leu Gln Gly Ile Leu Leu 130 135 140	432
gcc cat gtg ttt ttt aat ctg ccg atg gcg agc cgc tta tta ctc cag Ala His Val Phe Phe Asn Leu Pro Met Ala Ser Arg Leu Leu Leu Gln 145 150 155 160	480
gca ctg gaa aac atc ccc ggc gaa cag cgt caa ctt gcc gcc cag ctt Ala Leu Glu Asn Ile Pro Gly Glu Gln Arg Gln Leu Ala Ala Gln Leu 165 170 175	528
ggg atg cgt agc tgg cat ttt ttc cgc ttc gtc gaa tgg ccg tgg tta Gly Met Arg Ser Trp His Phe Phe Arg Phe Val Glu Trp Pro Trp Leu 180 185 190	576
cgg cga caa atc ccg ccg gtt gct gcg ctt atc ttt atg ctc tgt ttc Arg Arg Gln Ile Pro Pro Val Ala Ala Leu Ile Phe Met Leu Cys Phe 195 200 205	624
gcc agc ttc gcc acc gtg cta tcg ctg ggg ggc ggt ccg cag gcg acc	672

ES 2 806 732 T3

Ala 210	Ser	Phe	Ala	Thr	Val	Leu 215	Ser	Leu	Gly	Gly	Gly 220	Pro	Gln	Ala	Thr		
act	atc	gag	ctg	gca	atc	tat	cag	gcg	ctg	agt	tac	gac	tac	gat	cct		720
Thr	Ile	Glu	Leu	Ala	Ile	Tyr	Gln	Ala	Leu	Ser	Tyr	Asp	Tyr	Asp	Pro		
225					230					235					240		
gcc	cgc	gcg	gca	atg	ctg	gcg	ctg	ctc	cag	atg	gtg	tgc	tgc	ctc	ggg		768
Ala	Arg	Ala	Ala	Met	Leu	Ala	Leu	Leu	Gln	Met	Val	Cys	Cys	Leu	Gly		
				245					250					255			
ctg	gtg	ctg	ttg	agt	cag	cga	ttg	agt	aag	gcc	att	gcg	ccc	ggc	acc		816
Leu	Val	Leu	Leu	Ser	Gln	Arg	Leu	Ser	Lys	Ala	Ile	Ala	Pro	Gly	Thr		
			260					265					270				
acg	ctg	ctg	caa	ggc	tgg	cgc	gac	ccg	gac	gat	cgt	ctg	cat	agc	cgc		864
Thr	Leu	Leu	Gln	Gly	Trp	Arg	Asp	Pro	Asp	Asp	Arg	Leu	His	Ser	Arg		
			275				280					285					
att	tgc	gac	acg	gtg	tta	att	gtg	ctg	gcg	ctg	ctg	ctg	ttg	ctg	cca		912
Ile	Cys	Asp	Thr	Val	Leu	Ile	Val	Leu	Ala	Leu	Leu	Leu	Leu	Leu	Pro		
	290					295					300						
ccg	tta	ctg	gcg	gtg	atc	gtc	gat	ggg	gta	aat	cgc	cag	ttg	ccg	gaa		960
Pro	Leu	Leu	Ala	Val	Ile	Val	Asp	Gly	Val	Asn	Arg	Gln	Leu	Pro	Glu		
305					310					315					320		
gtg	ctg	gca	caa	ccg	gtg	ctg	tgg	cag	gcg	ctg	tgg	acc	tgc	ttg	cgt		1008
Val	Leu	Ala	Gln	Pro	Val	Leu	Trp	Gln	Ala	Leu	Trp	Thr	Ser	Leu	Arg		
				325					330					335			
att	gcg	ctg	gcg	gca	ggc	gta	ttg	tgc	gta	gtg	ctg	acc	atg	atg	ctg		1056
Ile	Ala	Leu	Ala	Ala	Gly	Val	Leu	Cys	Val	Val	Leu	Thr	Met	Met	Leu		
			340					345					350				
cta	tgg	agc	agt	cgc	gaa	ctg	cgg	gcg	cgg	cag	aaa	atg	ctg	gcg	ggc		1104
Leu	Trp	Ser	Ser	Arg	Glu	Leu	Arg	Ala	Arg	Gln	Lys	Met	Leu	Ala	Gly		
		355					360					365					
cag	gtg	ctg	gag	atg	agc	ggc	atg	ttg	atc	ctc	gcc	atg	ccg	ggg	att		1152
Gln	Val	Leu	Glu	Met	Ser	Gly	Met	Leu	Ile	Leu	Ala	Met	Pro	Gly	Ile		
	370					375					380						
gtg	ctg	gct	acc	ggc	ttc	ttt	tta	ctg	ctc	aac	aac	act	atc	ggc	ctg		1200
Val	Leu	Ala	Thr	Gly	Phe	Phe	Leu	Leu	Leu	Asn	Asn	Thr	Ile	Gly	Leu		
385					390					395					400		
cca	caa	tct	gct	gac	ggc	att	gtg	att	ttc	acc	aat	gcg	tta	atg	gcg		1248
Pro	Gln	Ser	Ala	Asp	Gly	Ile	Val	Ile	Phe	Thr	Asn	Ala	Leu	Met	Ala		
				405					410					415			
atc	cct	tat	gcg	ctg	aaa	gtg	ctg	gaa	aac	ccg	atg	cgc	gat	atc	acc		1296
Ile	Pro	Tyr	Ala	Leu	Lys	Val	Leu	Glu	Asn	Pro	Met	Arg	Asp	Ile	Thr		
			420					425					430				
gcc	cgc	tac	agc	atg	tta	tgt	cag	tgc	ctg	ggg	att	gaa	ggc	tgg	tca		1344
Ala	Arg	Tyr	Ser	Met	Leu	Cys	Gln	Ser	Leu	Gly	Ile	Glu	Gly	Trp	Ser		
			435				440					445					
cgc	tta	aaa	gtg	gtg	gag	ctg	cgc	gcc	ctg	aaa	cgt	cca	ctg	gcg	cag		1392
Arg	Leu	Lys	Val	Val	Glu	Leu	Arg	Ala	Leu	Lys	Arg	Pro	Leu	Ala	Gln		
	450					455					460						

ES 2 806 732 T3

gcg ctg gcc ttt gca tgc gtg ctg tcg att ggt gat ttt ggc gtg gtg 1440
Ala Leu Ala Phe Ala Cys Val Leu Ser Ile Gly Asp Phe Gly Val Val
465 470 475 480

gcg ttg ttc ggt aac gat gat ttc cgc acc ctg ccg ttt tat ctc tac 1488
Ala Leu Phe Gly Asn Asp Asp Phe Arg Thr Leu Pro Phe Tyr Leu Tyr
485 490 495

cag caa att ggc tcc tat cgc agc cag gac ggt gcg gtc acc gcg tta 1536
Gln Gln Ile Gly Ser Tyr Arg Ser Gln Asp Gly Ala Val Thr Ala Leu
500 505 510

att ctg ctg ctg ctc tgt ttt ctg ctg ttt acc gtg att gaa aaa cta 1584
Ile Leu Leu Leu Leu Cys Phe Leu Leu Phe Thr Val Ile Glu Lys Leu
515 520 525

ccg ggg cga aat gtt aaa act gac tga 1611
Pro Gly Arg Asn Val Lys Thr Asp
530 535

<210> 104
<211> 536
5 <212> PRT
<213> *Escherichia coli*

<400> 104

Met Ala Thr Arg Arg Gln Pro Leu Ile Pro Gly Trp Leu Ile Pro Gly
1 5 10 15

Val Ser Ala Thr Thr Leu Val Val Ala Val Ala Leu Ala Ala Phe Leu
20 25 30

Ala Leu Trp Trp Asn Ala Pro Gln Asp Asp Trp Val Ala Val Trp Gln
35 40 45

Asp Ser Tyr Leu Trp His Val Val Arg Phe Ser Phe Trp Gln Ala Phe
50 55 60

Leu Ser Ala Leu Leu Ser Val Ile Pro Ala Ile Phe Leu Ala Arg Ala
65 70 75 80

Leu Tyr Arg Arg Arg Phe Pro Gly Arg Leu Ala Leu Leu Arg Leu Cys
85 90 95

Ala Met Thr Leu Ile Leu Pro Val Leu Val Ala Val Phe Gly Ile Leu
100 105 110

Ser Val Tyr Gly Arg Gln Gly Trp Leu Ala Thr Leu Cys Gln Ser Leu
115 120 125

Gly Leu Glu Trp Thr Phe Ser Pro Tyr Gly Leu Gln Gly Ile Leu Leu
130 135 140

10

ES 2 806 732 T3

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

ES 2 806 732 T3

Pro Gln Ser Ala Asp Gly Ile Val Ile Phe Thr Asn Ala Leu Met Ala
405 410 415

Ile Pro Tyr Ala Leu Lys Val Leu Glu Asn Pro Met Arg Asp Ile Thr
420 425 430

Ala Arg Tyr Ser Met Leu Cys Gln Ser Leu Gly Ile Glu Gly Trp Ser
435 440 445

Arg Leu Lys Val Val Glu Leu Arg Ala Leu Lys Arg Pro Leu Ala Gln
450 455 460

Ala Leu Ala Phe Ala Cys Val Leu Ser Ile Gly Asp Phe Gly Val Val
465 470 475 480

Ala Leu Phe Gly Asn Asp Asp Phe Arg Thr Leu Pro Phe Tyr Leu Tyr
485 490 495

Gln Gln Ile Gly Ser Tyr Arg Ser Gln Asp Gly Ala Val Thr Ala Leu
500 505 510

Ile Leu Leu Leu Leu Cys Phe Leu Leu Phe Thr Val Ile Glu Lys Leu
515 520 525

Pro Gly Arg Asn Val Lys Thr Asp
530 535

<210> 105

<211> 699

5 <212> ADN

<213> *Escherichia coli*

<220>

<221> CDS

10 <222> (1)..(699)

<223> gen ThiQ de *E. coli* que codifica para una ATPasa de transportadores ABC de tiamina

<400> 105

atg tta aaa ctg act gat atc acc tgg ctt tac cac cat ttg ccg atg 48
Met Leu Lys Leu Thr Asp Ile Thr Trp Leu Tyr His His Leu Pro Met
1 5 10 15

cgt ttt agc tta acg gtg gaa cgc ggc gag cag gtg gcg atc ctc ggg 96
Arg Phe Ser Leu Thr Val Glu Arg Gly Glu Gln Val Ala Ile Leu Gly
20 25 30

cca agc ggc gcg ggt aaa agt acc ctg ctg aat ttg atc gcc ggt ttt 144
Pro Ser Gly Ala Gly Lys Ser Thr Leu Leu Asn Leu Ile Ala Gly Phe
35 40 45

15 ctg acg cca gcc agc ggt tcg ctg act atc gat ggc gta gat cac aca 192

ES 2 806 732 T3

Leu	Thr	Pro	Ala	Ser	Gly	Ser	Leu	Thr	Ile	Asp	Gly	Val	Asp	His	Thr		
50						55					60						
act	atg	ccg	ccg	tca	cgc	cgt	ccg	gtg	tcg	atg	ctg	ttt	cag	gag	aac	240	
Thr	Met	Pro	Pro	Ser	Arg	Arg	Pro	Val	Ser	Met	Leu	Phe	Gln	Glu	Asn		
65					70				75						80		
aac	ctg	ttc	agc	cac	ctg	acg	gtc	gca	cag	aac	atc	ggg	ctg	ggg	cta	288	
Asn	Leu	Phe	Ser	His	Leu	Thr	Val	Ala	Gln	Asn	Ile	Gly	Leu	Gly	Leu		
				85				90						95			
aat	ccg	gga	ttg	aaa	ctg	aac	gcg	gta	cag	cag	ggg	aaa	atg	cac	gct	336	
Asn	Pro	Gly	Leu	Lys	Leu	Asn	Ala	Val	Gln	Gln	Gly	Lys	Met	His	Ala		
			100				105						110				
atc	gcc	cgc	cag	atg	ggg	att	gat	aat	tta	atg	gcg	cgg	tta	ccg	ggc	384	
Ile	Ala	Arg	Gln	Met	Gly	Ile	Asp	Asn	Leu	Met	Ala	Arg	Leu	Pro	Gly		
		115					120					125					
gag	ctt	tcc	ggc	ggc	cag	cga	cag	cga	gtg	gcg	tta	gcg	cgt	tgt	ctg	432	
Glu	Leu	Ser	Gly	Gly	Gln	Arg	Gln	Arg	Val	Ala	Leu	Ala	Arg	Cys	Leu		
	130					135					140						
gta	cgc	gaa	cag	ccg	att	tta	ttg	ctc	gat	gaa	ccg	ttc	tct	gcg	ctc	480	
Val	Arg	Glu	Gln	Pro	Ile	Leu	Leu	Leu	Asp	Glu	Pro	Phe	Ser	Ala	Leu		
145					150				155						160		
gat	ccg	gcg	tta	cgt	cag	gag	atg	ttg	acg	ctg	gtg	agc	acg	agc	tgc	528	
Asp	Pro	Ala	Leu	Arg	Gln	Glu	Met	Leu	Thr	Leu	Val	Ser	Thr	Ser	Cys		
				165				170						175			
cag	cag	caa	aaa	atg	acg	cta	ttg	atg	gtg	tcg	cac	agc	gtg	gaa	gat	576	
Gln	Gln	Gln	Lys	Met	Thr	Leu	Leu	Met	Val	Ser	His	Ser	Val	Glu	Asp		
			180				185						190				
gcg	gcg	cgg	atc	gcc	acg	cgc	tcg	gta	gta	gtc	gcc	gac	ggg	cgc	atc	624	
Ala	Ala	Arg	Ile	Ala	Thr	Arg	Ser	Val	Val	Val	Ala	Asp	Gly	Arg	Ile		
		195					200					205					
gcc	tgg	cag	ggc	atg	acc	aat	gag	ttg	ttg	agc	ggc	aag	gcg	agt	gct	672	
Ala	Trp	Gln	Gly	Met	Thr	Asn	Glu	Leu	Leu	Ser	Gly	Lys	Ala	Ser	Ala		
	210					215					220						
tcg	gca	cta	ttg	ggg	att	acg	ggc	tag								699	
Ser	Ala	Leu	Leu	Gly	Ile	Thr	Gly										
225					230												

<210> 106

<211> 232

5 <212> PRT

<213> *Escherichia coli*

<400> 106

Met	Leu	Lys	Leu	Thr	Asp	Ile	Thr	Trp	Leu	Tyr	His	His	Leu	Pro	Met
1				5					10					15	

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

10

ES 2 806 732 T3

Pro Ser Gly Ala Gly Lys Ser Thr Leu Leu Asn Leu Ile Ala Gly Phe
35 40 45

Leu Thr Pro Ala Ser Gly Ser Leu Thr Ile Asp Gly Val Asp His Thr
50 55 60

Thr Met Pro Pro Ser Arg Arg Pro Val Ser Met Leu Phe Gln Glu Asn
65 70 75 80

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

Asn Pro Gly Leu Lys Leu Asn Ala Val Gln Gln Gly Lys Met His Ala
100 105 110

Ile Ala Arg Gln Met Gly Ile Asp Asn Leu Met Ala Arg Leu Pro Gly
115 120 125

Glu Leu Ser Gly Gly Gln Arg Gln Arg Val Ala Leu Ala Arg Cys Leu
130 135 140

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

Asp Pro Ala Leu Arg Gln Glu Met Leu Thr Leu Val Ser Thr Ser Cys
165 170 175

Gln Gln Gln Lys Met Thr Leu Leu Met Val Ser His Ser Val Glu Asp
180 185 190

Ala Ala Arg Ile Ala Thr Arg Ser Val Val Val Ala Asp Gly Arg Ile
195 200 205

Ala Trp Gln Gly Met Thr Asn Glu Leu Leu Ser Gly Lys Ala Ser Ala
210 215 220

Ser Ala Leu Leu Gly Ile Thr Gly
225 230

<210> 107

<211> 789

5 <212> ADN

<213> *Escherichia coli*

<220>

<221> CDS

10 <222> (1)..(789)

<223> gen ThiM de *E. coli* que codifica para una hidroxietiltiazol cinasa

<400> 107

ES 2 806 732 T3

```

atg caa gtc gac ctg ctg ggt tca gcg caa tct gcg cac gcg tta cac      48
Met Gln Val Asp Leu Leu Gly Ser Ala Gln Ser Ala His Ala Leu His
1      5      10      15

ctt ttt cac caa cat tcc cct ctt gtg cac tgc atg acc aat gat gtg      96
Leu Phe His Gln His Ser Pro Leu Val His Cys Met Thr Asn Asp Val
      20      25      30

gtg caa acc ttt acc gcc aat acc ttg ctg gcg ctc ggt gca tcg cca      144
Val Gln Thr Phe Thr Ala Asn Thr Leu Leu Ala Leu Gly Ala Ser Pro
      35      40      45

gcg atg gtt atc gaa acc gaa gag gcc agt cag ttt gcg gct atc gcc      192
Ala Met Val Ile Glu Thr Glu Glu Ala Ser Gln Phe Ala Ala Ile Ala
      50      55      60

agt gcc ttg ttg att aac gtt ggc aca ctg acg cag cca cgc gct cag      240
Ser Ala Leu Leu Ile Asn Val Gly Thr Leu Thr Gln Pro Arg Ala Gln
      65      70      75      80

gcg atg cgt gct gcc gtt gag caa gca aaa agc tct caa aca ccc tgg      288
Ala Met Arg Ala Ala Val Glu Gln Ala Lys Ser Ser Gln Thr Pro Trp
      85      90      95

acg ctt gat cca gta gcg gtg ggt gcg ctc gat tat cgc cgc cat ttt      336
Thr Leu Asp Pro Val Ala Val Gly Ala Leu Asp Tyr Arg Arg His Phe
      100      105      110

tgt cat gaa ctt tta tct ttt aaa ccg gca gcg ata cgt ggt aat gct      384
Cys His Glu Leu Leu Ser Phe Lys Pro Ala Ala Ile Arg Gly Asn Ala
      115      120      125

tcg gaa atc atg gca tta gct ggc att gct aat ggc gga cgg gga gtg      432
Ser Glu Ile Met Ala Leu Ala Gly Ile Ala Asn Gly Gly Arg Gly Val
      130      135      140

gat acc act gac gcc gca gct aac gcg ata ccc gct gca caa aca ctg      480
Asp Thr Thr Asp Ala Ala Ala Asn Ala Ile Pro Ala Ala Gln Thr Leu
      145      150      155      160

gca cgg gaa act ggc gca atc gtc gtg gtc act ggc gag atg gat tat      528
Ala Arg Glu Thr Gly Ala Ile Val Val Thr Gly Glu Met Asp Tyr
      165      170      175

gtt acc gat gga cat cgt atc att ggt att cac ggt ggt gat ccg tta      576
Val Thr Asp Gly His Arg Ile Ile Gly Ile His Gly Gly Asp Pro Leu
      180      185      190

atg acc aaa gtg gta gga act ggc tgt gca tta tcg gcg gtt gtc gct      624
Met Thr Lys Val Val Gly Thr Gly Cys Ala Leu Ser Ala Val Val Ala
      195      200      205

gcc tgc tgt gcg tta cca ggc gat acg ctg gaa aat gtc gca tct gcc      672
Ala Cys Cys Ala Leu Pro Gly Asp Thr Leu Glu Asn Val Ala Ser Ala
      210      215      220

tgt cac tgg atg aaa caa gcc gga gaa cgc gca gtc gcc aga agc gag      720
Cys His Trp Met Lys Gln Ala Gly Glu Arg Ala Val Ala Arg Ser Glu
      225      230      235      240

ggg cca ggc agt ttt gtt cca cat ttc ctt gat gcg ctc tgg caa ttg      768
Gly Pro Gly Ser Phe Val Pro His Phe Leu Asp Ala Leu Trp Gln Leu
      245      250      255

acg cag gag gtg cag gca tga      789
Thr Gln Glu Val Gln Ala
      260

```

<210> 108
 <211> 262
 <212> PRT
 <213> *Escherichia coli*

ES 2 806 732 T3

<400> 108

```

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

Leu Phe His Gln His Ser Pro Leu Val His Cys Met Thr Asn Asp Val
20      25      30

Val Gln Thr Phe Thr Ala Asn Thr Leu Leu Ala Leu Gly Ala Ser Pro
35      40      45

Ala Met Val Ile Glu Thr Glu Glu Ala Ser Gln Phe Ala Ala Ile Ala
50      55      60

Ser Ala Leu Leu Ile Asn Val Gly Thr Leu Thr Gln Pro Arg Ala Gln
65      70      75      80

Ala Met Arg Ala Ala Val Glu Gln Ala Lys Ser Ser Gln Thr Pro Trp
85      90      95

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

Cys His Glu Leu Leu Ser Phe Lys Pro Ala Ala Ile Arg Gly Asn Ala
115     120     125

Ser Glu Ile Met Ala Leu Ala Gly Ile Ala Asn Gly Gly Arg Gly Val
130     135     140

Asp Thr Thr Asp Ala Ala Ala Asn Ala Ile Pro Ala Ala Gln Thr Leu
145     150     155     160

Ala Arg Glu Thr Gly Ala Ile Val Val Thr Gly Glu Met Asp Tyr
165     170     175

Val Thr Asp Gly His Arg Ile Ile Gly Ile His Gly Gly Asp Pro Leu
180     185     190

Met Thr Lys Val Val Gly Thr Gly Cys Ala Leu Ser Ala Val Val Ala
195     200     205

Ala Cys Cys Ala Leu Pro Gly Asp Thr Leu Glu Asn Val Ala Ser Ala
210     215     220

Cys His Trp Met Lys Gln Ala Gly Glu Arg Ala Val Ala Arg Ser Glu
225     230     235     240

Gly Pro Gly Ser Phe Val Pro His Phe Leu Asp Ala Leu Trp Gln Leu
245     250     255

Thr Gln Glu Val Gln Ala
260

```

5

<210> 109

<211> 978

<212> ADN
<213> *Escherichia coli*

<220>

5 <221> CDS

<222> (1)..(978)

<223> gen ThiL de *E. coli* que codifica para tiamina fosfato cinasa (nota: la mutación en los nucleótidos 133-135 (GGT por GAC) codifica para una sustitución G133D)

10 <400> 109

atg gca tgt ggc gag ttc tcc ctg att gcc cgt tat ttt gac cgt gta	48
Met Ala Cys Gly Glu Phe Ser Leu Ile Ala Arg Tyr Phe Asp Arg Val	
1 5 10 15	
aga agt tct cgt ctt gat gtc gaa ctg ggc atc ggc gac gat tgc gca	96
Arg Ser Ser Arg Leu Asp Val Glu Leu Gly Ile Gly Asp Asp Cys Ala	
20 25 30	
ctt ctc aat atc ccc gag aaa cag acc ctg gcg atc agc act gat acg	144
Leu Leu Asn Ile Pro Glu Lys Gln Thr Leu Ala Ile Ser Thr Asp Thr	
35 40 45	
ctg gtg gcg ggt aac cat ttc ctc cct gat atc gat cct gct gat ctg	192
Leu Val Ala Gly Asn His Phe Leu Pro Asp Ile Asp Pro Ala Asp Leu	
50 55 60	
gct tat aaa gca ctg gcg gtg aac cta agc gat ctg gca gcg atg ggg	240
Ala Tyr Lys Ala Leu Ala Val Asn Leu Ser Asp Leu Ala Ala Met Gly	
65 70 75 80	
gcc gat ccg gcc tgg ctg acg ctg gca tta acc tta ccg gac gta gac	288
Ala Asp Pro Ala Trp Leu Thr Leu Ala Leu Thr Leu Pro Asp Val Asp	
85 90 95	
gaa gcg tgg ctt gag tcc ttc agc gac agt ttg ttt gat ctt ctc aat	336
Glu Ala Trp Leu Glu Ser Phe Ser Asp Ser Leu Phe Asp Leu Leu Asn	
100 105 110	
tat tac gat atg caa ctc att ggc ggc gat acc acg cgt ggg cca tta	384
Tyr Tyr Asp Met Gln Leu Ile Gly Gly Asp Thr Thr Arg Gly Pro Leu	

ES 2 806 732 T3

115	120	125	
tca atg acg ttg ggt atc cac ggc ttt gtt ccg atg gga cga gcc tta			432
Ser Met Thr Leu Gly Ile His Gly Phe Val Pro Met Gly Arg Ala Leu			
130	135	140	
acg cgc tct ggg gcg aaa ccg ggt gac tgg atc tat gtg acc ggt aca			480
Thr Arg Ser Gly Ala Lys Pro Gly Asp Trp Ile Tyr Val Thr Gly Thr			
145	150	155	160
ccg ggc gat agc gcc gcc ggg ctg gcg att ttg caa aac cgt ttg cag			528
Pro Gly Asp Ser Ala Ala Gly Leu Ala Ile Leu Gln Asn Arg Leu Gln			
	165	170	175
gtt gcc gat gct aaa gat gcg gac tac ttg atc aaa cgt cat ctc cgt			576
Val Ala Asp Ala Lys Asp Ala Asp Tyr Leu Ile Lys Arg His Leu Arg			
	180	185	190
cca tcg ccg cgt att tta cag ggg cag gca ctg cgc gat ctg gca aat			624
Pro Ser Pro Arg Ile Leu Gln Gly Gln Ala Leu Arg Asp Leu Ala Asn			
	195	200	205
tca gcc atc gat ctc tct gac ggt ttg att tcc gat ctc ggg cat atc			672
Ser Ala Ile Asp Leu Ser Asp Gly Leu Ile Ser Asp Leu Gly His Ile			
	210	215	220
gtg aaa gcc agc gac tgc ggc gca cgt att gac ctg gca ttg ctg ccg			720
Val Lys Ala Ser Asp Cys Gly Ala Arg Ile Asp Leu Ala Leu Leu Pro			
	225	230	235
ttt tct gat gcg ctt tct cgc cat gtt gaa ccg gaa cag gcg ctg cgc			768
Phe Ser Asp Ala Leu Ser Arg His Val Glu Pro Glu Gln Ala Leu Arg			
	245	250	255
tgg gcg ctc tct ggc ggt gaa gat tac gag ttg tgt ttc act gtg ccg			816
Trp Ala Leu Ser Gly Gly Glu Asp Tyr Glu Leu Cys Phe Thr Val Pro			
	260	265	270
gaa ctg aac cgt ggc gcg ctg gat gtg gct ctc gga cac ctg ggc gta			864
Glu Leu Asn Arg Gly Ala Leu Asp Val Ala Leu Gly His Leu Gly Val			
	275	280	285
ccg ttt acc tgt atc ggg caa atg acc gcc gat atc gaa ggg ctt tgt			912
Pro Phe Thr Cys Ile Gly Gln Met Thr Ala Asp Ile Glu Gly Leu Cys			
	290	295	300
ttt att cgt gac ggc gaa cct gtt aca tta gac tgg aaa gga tat gac			960
Phe Ile Arg Asp Gly Glu Pro Val Thr Leu Asp Trp Lys Gly Tyr Asp			
	305	310	315
cat ttt gcc acg cca taa			978
His Phe Ala Thr Pro			
	325		

<210> 110
 <211> 325
 <212> PRT
 <213> *Escherichia coli*

<400> 110

Met Ala Cys Gly Glu Phe Ser Leu Ile Ala Arg Tyr Phe Asp Arg Val

ES 2 806 732 T3

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

Trp Ala Leu Ser Gly Gly Glu Asp Tyr Glu Leu Cys Phe Thr Val Pro
260 265 270

Glu Leu Asn Arg Gly Ala Leu Asp Val Ala Leu Gly His Leu Gly Val
275 280 285

Pro Phe Thr Cys Ile Gly Gln Met Thr Ala Asp Ile Glu Gly Leu Cys
290 295 300

Phe Ile Arg Asp Gly Glu Pro Val Thr Leu Asp Trp Lys Gly Tyr Asp
305 310 315 320

His Phe Ala Thr Pro
325

<210> 111

<211> 876

5 <212> ADN

<213> *Staphylococcus aureus*

<220>

<221> CDS

10 <222> (1)..(876)

<223> rgsA de Newmann de *S. aureus* que codifica para una fosfatasa

<400> 111

atg aag aca ggt cga ata gtg aaa tca att agt ggg gta tat caa gta 48
Met Lys Thr Gly Arg Ile Val Lys Ser Ile Ser Gly Val Tyr Gln Val
1 5 10 15

gac gtt aat ggc gaa cgt ttc aat aca aaa cca cga gga tta ttt aga 96
Asp Val Asn Gly Glu Arg Phe Asn Thr Lys Pro Arg Gly Leu Phe Arg
20 25 30

aag aaa aaa ttt tca ccg gta gtt ggt gat ata gtg gaa ttt gaa gtt 144
Lys Lys Lys Phe Ser Pro Val Val Gly Asp Ile Val Glu Phe Glu Val
35 40 45

caa aac att aac gaa ggt tat att cat caa gtg ttt gag cgg gaa aat 192
Gln Asn Ile Asn Glu Gly Tyr Ile His Gln Val Phe Glu Arg Glu Asn
50 55 60

gag ttg aaa aga cca cct gta agt aat ata gat aca cta gta att gta 240
Glu Leu Lys Arg Pro Pro Val Ser Asn Ile Asp Thr Leu Val Ile Val
65 70 75 80

atg agt gct gtc gag cca aat ttt tca acg caa tta tta gat cga ttt 288
Met Ser Ala Val Glu Pro Asn Phe Ser Thr Gln Leu Leu Asp Arg Phe
85 90 95

tta gtt att gca cat tcg tat cag tta aat gcg aga att ttg gtg act 336
Leu Val Ile Ala His Ser Tyr Gln Leu Asn Ala Arg Ile Leu Val Thr
100 105 110

aaa aaa gat aaa aca cca att gaa aag cag ttc gaa ata aat gag ttg 384
Lys Lys Asp Lys Thr Pro Ile Glu Lys Gln Phe Glu Ile Asn Glu Leu
115 120 125

15

ES 2 806 732 T3

ttg aaa ata tat gaa aat att ggc tat gag act gaa ttt att gga aat	432
Leu Lys Ile Tyr Glu Asn Ile Gly Tyr Glu Thr Glu Phe Ile Gly Asn	
130 135 140	
gat gat gat cga aaa aaa att gta gaa gct tgg cca gct gga ctt ata	480
Asp Asp Asp Arg Lys Lys Ile Val Glu Ala Trp Pro Ala Gly Leu Ile	
145 150 155 160	
gta ctt agt ggt caa tca ggt gtc ggt aag tcc act ttc tta aat cat	528
Val Leu Ser Gly Gln Ser Gly Val Gly Lys Ser Thr Phe Leu Asn His	
165 170 175	
tat cgt cca gaa ctt aat ctt gag aca aat gat ata tca aaa tca tta	576
Tyr Arg Pro Glu Leu Asn Leu Glu Thr Asn Asp Ile Ser Lys Ser Leu	
180 185 190	
aat cga gga aag cat act aca aga cat gtc gaa cta ttc gaa cgt caa	624
Asn Arg Gly Lys His Thr Thr Arg His Val Glu Leu Phe Glu Arg Gln	
195 200 205	
aac ggt tat att gca gac aca cct gga ttc agt gct tta gat ttt gat	672
Asn Gly Tyr Ile Ala Asp Thr Pro Gly Phe Ser Ala Leu Asp Phe Asp	
210 215 220	
cat ata gat aaa gat gaa ata aaa gat tat ttt ctt gaa tta aat cga	720
His Ile Asp Lys Asp Glu Ile Lys Asp Tyr Phe Leu Glu Leu Asn Arg	
225 230 235 240	
tat ggt gaa aca tgt aag ttt agg aat tgt aat cat atc aaa gaa cct	768
Tyr Gly Glu Thr Cys Lys Phe Arg Asn Cys Asn His Ile Lys Glu Pro	
245 250 255	
aat tgt aat gtt aag cat caa tta gag ata ggg aat att gcg caa ttt	816
Asn Cys Asn Val Lys His Gln Leu Glu Ile Gly Asn Ile Ala Gln Phe	
260 265 270	
aga tac gac cat tat tta caa cta ttt aat gaa att tca aat aga aag	864
Arg Tyr Asp His Tyr Leu Gln Leu Phe Asn Glu Ile Ser Asn Arg Lys	
275 280 285	
gtt aga tat taa	876
Val Arg Tyr	
290	

<210> 112

<211> 291

5 <212> PRT

<213> *Staphylococcus aureus*

<400> 112

Met Lys Thr Gly Arg Ile Val Lys Ser Ile Ser Gly Val Tyr Gln Val
1 5 10 15

Asp Val Asn Gly Glu Arg Phe Asn Thr Lys Pro Arg Gly Leu Phe Arg
20 25 30

Lys Lys Lys Phe Ser Pro Val Val Gly Asp Ile Val Glu Phe Glu Val
35 40 45

10

ES 2 806 732 T3

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

<210> 113

5 <211> 786

<212> ADN

<213> Especie *Pseudomonas*

<220>

10 <221> CDS

<222> (1)..(786)

<223> gen de *Pseudomonas* que codifica para una fosfatasa

ES 2 806 732 T3

<400> 113

atg tgc cag cag cat ccg cct gcc aac gac caa ctg gac gcg gtg ctg Met Cys Gln Gln His Pro Pro Ala Asn Asp Gln Leu Asp Ala Val Leu 1 5 10 15	48
tgg acc cag acc tcc atc gaa cac gag ctg att tat cgc cag gta ttc Trp Thr Gln Thr Ser Ile Glu His Glu Leu Ile Tyr Arg Gln Val Phe 20 25 30	96
gcc agc gcc acg cgc cag ctg gac gcc gcg ctg gcc gat ccg acc tgg Ala Ser Ala Thr Arg Gln Leu Asp Ala Ala Leu Ala Asp Pro Thr Trp 35 40 45	144
gac gcc ctg ccc ttg ccg ccg cgc aac ctc gcc ggg ctg ccg ccg gca Asp Ala Leu Pro Leu Pro Pro Arg Asn Leu Ala Gly Leu Pro Pro Ala 50 55 60	192
gtg gta gtg gat atc gat gaa acc gtg ctc gac aac gtg ccg ctc aac Val Val Val Asp Ile Asp Glu Thr Val Leu Asp Asn Val Pro Leu Asn 65 70 75 80	240
gcg cgg gac atc atc aac aac cag gtg tat tcc tat gac cgc tgg aac Ala Arg Asp Ile Ile Asn Asn Gln Val Tyr Ser Tyr Asp Arg Trp Asn 85 90 95	288
acc tgg gtc gac cag gcc aag gcc cag gcg ctg ccc ggc gcc gta gcg Thr Trp Val Asp Gln Ala Lys Ala Gln Ala Leu Pro Gly Ala Val Ala 100 105 110	336
ttc ctg caa gcg gcc gat aaa aaa ggc atc acc gtc tac tac atc acc Phe Leu Gln Ala Ala Asp Lys Lys Gly Ile Thr Val Tyr Tyr Ile Thr 115 120 125	384
aac cgt gaa cac agc cag gtc cag gcc acg gtc gac aac ctg cgc ctg Asn Arg Glu His Ser Gln Val Gln Ala Thr Val Asp Asn Leu Arg Leu 130 135 140	432
cgc ggt ttc ccg gtg cag cgc aac gaa cag gta ctg gcc gcc agc acc Arg Gly Phe Pro Val Gln Arg Asn Glu Gln Val Leu Ala Ala Ser Thr 145 150 155 160	480
ccg acc ggc cac tgc gaa cag gcc ggc tat ggc aag aac tgc cgc cgc Pro Thr Gly His Cys Glu Gln Ala Gly Tyr Gly Lys Asn Cys Arg Arg 165 170 175	528
cag tgg gtg gcc gcg cat gcc cgt gtg ctg atg ctc gcc ggc gac tcg Gln Trp Val Ala Ala His Ala Arg Val Leu Met Leu Ala Gly Asp Ser 180 185 190	576
ctg ggc gac ttc gtg cag gcc gag cac aac acc ctg gcc gac cag cgc Leu Gly Asp Phe Val Gln Ala Glu His Asn Thr Leu Ala Asp Gln Arg 195 200 205	624
aag gca gcc gaa ccg tac ctg gcc tgg ctc ggc caa cgc tgg ttc ctg Lys Ala Ala Glu Pro Tyr Leu Ala Trp Leu Gly Gln Arg Trp Phe Leu 210 215 220	672
ctg ccc aac cct acg tat ggc aac tgg tac agc gcg ccc tac ggc gac Leu Pro Asn Pro Thr Tyr Gly Asn Trp Tyr Ser Ala Pro Tyr Gly Asp 225 230 235 240	720
cag gaa aaa ctg ccc ttt gaa cgc aag cgc cag ctc aag caa cag gcg Gln Glu Lys Leu Pro Phe Glu Arg Lys Arg Gln Leu Lys Gln Gln Ala 245 250 255	768
ctg cct cta gaa aac tag Leu Pro Leu Glu Asn 260	786

5

<210> 114

ES 2 806 732 T3

<211> 261
 <212> PRT
 <213> *Pseudomonas species*

5 <400> 114

```

Met Cys Gln Gln His Pro Pro Ala Asn Asp Gln Leu Asp Ala Val Leu
1          5          10          15

Trp Thr Gln Thr Ser Ile Glu His Glu Leu Ile Tyr Arg Gln Val Phe
          20          25          30

Ala Ser Ala Thr Arg Gln Leu Asp Ala Ala Leu Ala Asp Pro Thr Trp
          35          40          45

Asp Ala Leu Pro Leu Pro Pro Arg Asn Leu Ala Gly Leu Pro Pro Ala
          50          55          60

Val Val Val Asp Ile Asp Glu Thr Val Leu Asp Asn Val Pro Leu Asn
65          70          75          80

Ala Arg Asp Ile Ile Asn Asn Gln Val Tyr Ser Tyr Asp Arg Trp Asn
          85          90          95

Thr Trp Val Asp Gln Ala Lys Ala Gln Ala Leu Pro Gly Ala Val Ala
          100          105          110

Phe Leu Gln Ala Ala Asp Lys Lys Gly Ile Thr Val Tyr Tyr Ile Thr
          115          120          125

Asn Arg Glu His Ser Gln Val Gln Ala Thr Val Asp Asn Leu Arg Leu
          130          135          140
Arg Gly Phe Pro Val Gln Arg Asn Glu Gln Val Leu Ala Ala Ser Thr
145          150          155          160

Pro Thr Gly His Cys Glu Gln Ala Gly Tyr Gly Lys Asn Cys Arg Arg
          165          170          175

Gln Trp Val Ala Ala His Ala Arg Val Leu Met Leu Ala Gly Asp Ser
          180          185          190

Leu Gly Asp Phe Val Gln Ala Glu His Asn Thr Leu Ala Asp Gln Arg
          195          200          205

Lys Ala Ala Glu Pro Tyr Leu Ala Trp Leu Gly Gln Arg Trp Phe Leu
          210          215          220

Leu Pro Asn Pro Thr Tyr Gly Asn Trp Tyr Ser Ala Pro Tyr Gly Asp
225          230          235          240

Gln Glu Lys Leu Pro Phe Glu Arg Lys Arg Gln Leu Lys Gln Gln Ala
          245          250          255

Leu Pro Leu Glu Asn
          260
  
```

<210> 115
 <211> 1404
 <212> ADN
 <213> *Saccharomyces cerevisiae*

5

<220>
 <221> CDS
 <222> (1)..(1404)
 <223> gen PHHO3 de *S. cerevisiae* que codifica para una fosfatasa (YBR092C)

10

<400> 115

atg ttc aaa tcc gtg gtg tat agc gtc ctg gcg gcc gct ttg gtc aac	48
Met Phe Lys Ser Val Val Tyr Ser Val Leu Ala Ala Ala Leu Val Asn	
1 5 10 15	
gct ggt act att ccg ctg gga gaa ctt gcg gat gtg gcg aaa att ggc	96
Ala Gly Thr Ile Pro Leu Gly Glu Leu Ala Asp Val Ala Lys Ile Gly	
20 25 30	
acc caa gag gac atc ttc ccg ttc ctc ggc ggc gcc gga ccg tat ttt	144
Thr Gln Glu Asp Ile Phe Pro Phe Leu Gly Gly Ala Gly Pro Tyr Phe	
35 40 45	
agt ttt ccg ggt gac tac ggg att tcc cgc gac tta cca gaa ggc tgt	192
Ser Phe Pro Gly Asp Tyr Gly Ile Ser Arg Asp Leu Pro Glu Gly Cys	
50 55 60	
gag atg aaa cag tta caa atg tta gct cgt cat ggc gaa cgc tac cca	240
Glu Met Lys Gln Leu Gln Met Leu Ala Arg His Gly Glu Arg Tyr Pro	

ES 2 806 732 T3

65	70	75	80	
acc tac tca aag ggg gcg act atc atg aaa acc tgg tat aaa ctc agt Thr Tyr Ser Lys Gly Ala Thr Ile Met Lys Thr Trp Tyr Lys Leu Ser 85 90 95				288
aac tat acc cgc cag ttt aac ggc agc ctt agc ttc ctt aac gat gac Asn Tyr Thr Arg Gln Phe Asn Gly Ser Leu Ser Phe Leu Asn Asp Asp 100 105 110				336
tac gag ttt ttt atc cgt gat gat gat gac ctc gaa atg gaa aca acg Tyr Glu Phe Phe Ile Arg Asp Asp Asp Leu Glu Met Glu Thr Thr 115 120 125				384
ttt gca aac agc gac aat gta ttg aac ccg tat acg ggc gaa atg gat Phe Ala Asn Ser Asp Asn Val Leu Asn Pro Tyr Thr Gly Glu Met Asp 130 135 140				432
gca aaa cgt cat gca cgc gaa ttc ctt gcg caa tat ggc tat atg ttc Ala Lys Arg His Ala Arg Glu Phe Leu Ala Tyr Gly Tyr Met Phe 145 150 155 160				480
gaa aat cag act tcg ttt cca atc ttt gcg gcg tcc tct gaa cgc gtt Glu Asn Gln Thr Ser Phe Pro Ile Phe Ala Ala Ser Ser Glu Arg Val 165 170 175				528
cat gac aca gct caa tac ttt att gat ggt ctg ggg gat cag ttc aac His Asp Thr Ala Gln Tyr Phe Ile Asp Gly Leu Gly Asp Gln Phe Asn 180 185 190				576
atc agt ctg caa acc gtt agc gaa gcg atg agc gcg ggt gcg aat act Ile Ser Leu Gln Thr Val Ser Glu Ala Met Ser Ala Gly Ala Asn Thr 195 200 205				624
ctg tcc gct gga aac gct tgt cct ggc tgg gac gag gat gcg aat gat Leu Ser Ala Gly Asn Ala Cys Pro Gly Trp Asp Glu Asp Ala Asn Asp 210 215 220				672
gac att tta gat aaa tac gat acc acc tac ctc gat gat att gct aag Asp Ile Leu Asp Lys Tyr Asp Thr Thr Tyr Leu Asp Asp Ile Ala Lys 225 230 235 240				720
cgt tta aac aaa gaa aac aag ggg ctg aat ctg acg tca aaa gac gcc Arg Leu Asn Lys Glu Asn Lys Gly Leu Asn Leu Thr Ser Lys Asp Ala 245 250 255				768
aac acc ctg ttc gcc tgg tgt gcg tac gaa ctg aac gca cgc ggc tac Asn Thr Leu Phe Ala Trp Cys Ala Tyr Glu Leu Asn Ala Arg Gly Tyr 260 265 270				816
agc gac gta tgc gat atc ttt acc gaa gac gaa ctt gtc cgc tat agt Ser Asp Val Cys Asp Ile Phe Thr Glu Asp Glu Leu Val Arg Tyr Ser 275 280 285				864
tac ggt caa gac ctg gtt agc ttt tat cag gat ggc cca ggg tac gat Tyr Gly Gln Asp Leu Val Ser Phe Thr Tyr Gln Asp Gly Pro Gly Tyr Asp 290 295 300				912
atg att cgc agc gtg ggt gcc aat ctg ttc aac gcc acc ctt aag ttg Met Ile Arg Ser Val Gly Ala Asn Leu Phe Asn Ala Thr Leu Lys Leu 305 310 315 320				960
ctg aaa cag tcg gag act cag gat ctg aaa gtg tgg ctc tca ttc acc				1008

ES 2 806 732 T3

Leu	Lys	Gln	Ser	Glu	Thr	Gln	Asp	Leu	Lys	Val	Trp	Leu	Ser	Phe	Thr		
				325					330					335			
cat	gac	acc	gat	att	ctg	aat	tat	ctg	acc	act	gca	ggc	atc	att	gat		1056
His	Asp	Thr	Asp	Ile	Leu	Asn	Tyr	Leu	Thr	Thr	Ala	Gly	Ile	Ile	Asp		
			340					345					350				
gac	aag	aac	aac	ctg	acc	gct	gaa	tat	gta	ccg	ttt	atg	ggg	aac	acc		1104
Asp	Lys	Asn	Asn	Leu	Thr	Ala	Glu	Tyr	Val	Pro	Phe	Met	Gly	Asn	Thr		
			355				360					365					
ttt	cac	aaa	agc	tgg	tac	gtg	cct	caa	ggg	gcg	cgt	gtg	tat	acg	gaa		1152
Phe	His	Lys	Ser	Trp	Tyr	Val	Pro	Gln	Gly	Ala	Arg	Val	Tyr	Thr	Glu		
			370			375					380						
aaa	ttt	cag	tgc	tcg	aac	gat	acc	tat	gtg	cgc	tac	gtc	att	aat	gat		1200
Lys	Phe	Gln	Cys	Ser	Asn	Asp	Thr	Tyr	Val	Arg	Tyr	Val	Ile	Asn	Asp		
385						390				395					400		
gcc	gtc	gtc	ccc	att	gag	acg	tgc	tct	acc	ggc	cca	ggg	ttt	agt	tgt		1248
Ala	Val	Val	Pro	Ile	Glu	Thr	Cys	Ser	Thr	Gly	Pro	Gly	Phe	Ser	Cys		
				405					410					415			
gaa	atc	aac	gac	ttc	tat	gac	tac	gcc	gag	aaa	cgc	gtt	gcc	ggg	acc		1296
Glu	Ile	Asn	Asp	Phe	Tyr	Asp	Tyr	Ala	Glu	Lys	Arg	Val	Ala	Gly	Thr		
			420					425					430				
gat	ttt	ctt	aaa	gtg	tgc	aac	gtt	tct	agc	gtt	tct	aat	gtg	act	gaa		1344
Asp	Phe	Leu	Lys	Val	Cys	Asn	Val	Ser	Ser	Val	Ser	Asn	Val	Thr	Glu		
			435				440					445					
ctg	acc	ttt	tac	tgg	gat	tgg	aac	acg	acg	cac	tac	aat	gac	aca	ctg		1392
Leu	Thr	Phe	Tyr	Trp	Asp	Trp	Asn	Thr	Thr	His	Tyr	Asn	Asp	Thr	Leu		
			450			455					460						
tta	aaa	cag	taa														1404
Leu	Lys	Gln															
465																	

<210> 116

<211> 467

5 <212> PRT

<213> *Saccharomyces cerevisiae*

<400> 116

Met	Phe	Lys	Ser	Val	Val	Tyr	Ser	Val	Leu	Ala	Ala	Ala	Leu	Val	Asn		
1				5					10					15			

Ala	Gly	Thr	Ile	Pro	Leu	Gly	Glu	Leu	Ala	Asp	Val	Ala	Lys	Ile	Gly		
			20					25					30				

Thr	Gln	Glu	Asp	Ile	Phe	Pro	Phe	Leu	Gly	Gly	Ala	Gly	Pro	Tyr	Phe		
		35					40					45					

Ser	Phe	Pro	Gly	Asp	Tyr	Gly	Ile	Ser	Arg	Asp	Leu	Pro	Glu	Gly	Cys		
	50					55					60						

10

ES 2 806 732 T3

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

ES 2 806 732 T3

Leu Lys Gln Ser Glu Thr Gln Asp Leu Lys Val Trp Leu Ser Phe Thr
325 330 335

His Asp Thr Asp Ile Leu Asn Tyr Leu Thr Thr Ala Gly Ile Ile Asp
340 345 350

Asp Lys Asn Asn Leu Thr Ala Glu Tyr Val Pro Phe Met Gly Asn Thr
355 360 365

Phe His Lys Ser Trp Tyr Val Pro Gln Gly Ala Arg Val Tyr Thr Glu
370 375 380

Lys Phe Gln Cys Ser Asn Asp Thr Tyr Val Arg Tyr Val Ile Asn Asp
385 390 395 400

Ala Val Val Pro Ile Glu Thr Cys Ser Thr Gly Pro Gly Phe Ser Cys
405 410 415

Glu Ile Asn Asp Phe Tyr Asp Tyr Ala Glu Lys Arg Val Ala Gly Thr
420 425 430

Asp Phe Leu Lys Val Cys Asn Val Ser Ser Val Ser Asn Val Thr Glu
435 440 445

Leu Thr Phe Tyr Trp Asp Trp Asn Thr Thr His Tyr Asn Asp Thr Leu
450 455 460

Leu Lys Gln
465

<210> 117-rg

<211> 462

5 <212> ADN

<213> *Escherichia coli*

<220>

<221> CDS

10 <222> (1)..(462)

<223> gen nudJ de *E. coli*

<400> 117

atg ttt aaa ccg cac gtt acc gtt gct tgc gtg gtg cac gca gaa ggc 48
Met Phe Lys Pro His Val Thr Val Ala Cys Val Val His Ala Glu Gly
1 5 10 15

aaa ttt tta gtc gtt gaa gag acg att aat ggt aaa gcg tta tgg aac 96
Lys Phe Leu Val Val Glu Glu Thr Ile Asn Gly Lys Ala Leu Trp Asn
20 25 30

caa cct gcc ggg cat ctg gaa gcc gat gaa acc tta gtg gaa gcc gcc 144
15 Gln Pro Ala Gly His Leu Glu Ala Asp Glu Thr Leu Val Glu Ala Ala

ES 2 806 732 T3

	35	40	45	
	gcc cgt gaa ctg tgg gaa gaa acc ggc atc agc gcg cag ccg caa cac			192
	Ala Arg Glu Leu Trp Glu Glu Thr Gly Ile Ser Ala Gln Pro Gln His			
	50	55	60	
	ttt att cgt atg cat cag tgg att gcg cca gat aaa acg ccg ttt ttg			240
	Phe Ile Arg Met His Gln Trp Ile Ala Pro Asp Lys Thr Pro Phe Leu			
	65	70	75	80
	cgt ttc ctc ttt gcc att gag ctt gag caa ata tgc ccg acc cag cct			288
	Arg Phe Leu Phe Ala Ile Glu Leu Glu Gln Ile Cys Pro Thr Gln Pro			
		85	90	95
	cat gac agc gat atc gac tgc tgc cgt tgg gtc agc gcc gaa gaa att			336
	His Asp Ser Asp Ile Asp Cys Cys Arg Trp Val Ser Ala Glu Glu Ile			
		100	105	110
	tta cag gcg tca aat ctt cgt tcg ccg ctg gtg gcg gaa agt att cgt			384
	Leu Gln Ala Ser Asn Leu Arg Ser Pro Leu Val Ala Glu Ser Ile Arg			
		115	120	125
	tgt tat caa agc ggg caa cgt tat ccg ctg gag atg att ggc gat ttt			432
	Cys Tyr Gln Ser Gly Gln Arg Tyr Pro Leu Glu Met Ile Gly Asp Phe			
		130	135	140
	aac tgg cct ttt aca aag ggt gtc atc taa			462
	Asn Trp Pro Phe Thr Lys Gly Val Ile			
	145	150		
	<210> 118			
	<211> 153			
5	<212> PRT			
	<213> <i>Escherichia coli</i>			
	<400> 118			
	Met Phe Lys Pro His Val Thr Val Ala Cys Val Val His Ala Glu Gly			
	1	5	10	15
	Lys Phe Leu Val Val Glu Glu Thr Ile Asn Gly Lys Ala Leu Trp Asn			
		20	25	30
	Gln Pro Ala Gly His Leu Glu Ala Asp Glu Thr Leu Val Glu Ala Ala			
		35	40	45
	Ala Arg Glu Leu Trp Glu Glu Thr Gly Ile Ser Ala Gln Pro Gln His			
		50	55	60
	Phe Ile Arg Met His Gln Trp Ile Ala Pro Asp Lys Thr Pro Phe Leu			
		65	70	75
	Arg Phe Leu Phe Ala Ile Glu Leu Glu Gln Ile Cys Pro Thr Gln Pro			
		85	90	95
10	His Asp Ser Asp Ile Asp Cys Cys Arg Trp Val Ser Ala Glu Glu Ile			

ES 2 806 732 T3

	100	105	110
	Leu Gln Ala Ser Asn Leu Arg Ser Pro Leu Val Ala Glu Ser Ile Arg		
	115	120	125
	Cys Tyr Gln Ser Gly Gln Arg Tyr Pro Leu Glu Met Ile Gly Asp Phe		
	130	135	140
	Asn Trp Pro Phe Thr Lys Gly Val Ile		
	145	150	
	<210> 119		
	<211> 47		
5	<212> ADN		
	<213> <i>Escherichia coli</i>		
	<220>		
	<221> unión de cebador		
10	<222> (1)..(47)		
	<223> cebador oBS194		
	<400> 119		
15	gtcctactca ggagagcgtt caccgacaac tcaggagagc gttcacc		47
	<210> 120		
	<211> 48		
	<212> ADN		
20	<213> <i>Escherichia coli</i>		
	<220>		
	<221> unión de cebador		
	<222> (1)..(48)		
25	<223> cebador oBS195		
	<400> 120		
30	ctttcgtctt cacctcgagg gaaatcaaaa taggcgtatc acgaggcc		48
	<210> 121		
	<211> 26		
	<212> ADN		
	<213> <i>Escherichia coli</i>		
35	<220>		
	<221> unión de cebador		
	<222> (1)..(26)		
	<223> cebador oBS196		
40	<400> 121		
	gatttcctc gaggtgaaga cgaaag		26
45	<210> 122		
	<211> 20		
	<212> ADN		
	<213> <i>Escherichia coli</i>		
50	<220>		
	<221> unión de cebador		
	<222> (1)..(20)		
	<223> cebador oBS197		
55	<400> 122		

	tgctcggtgaa cgctctctg	20
5	<210> 123 <211> 26 <212> ADN <213> <i>Escherichia coli</i> <220> <221> unión de cebador <222> (1)..(26) <223> cebador oBS198	
15	<400> 123 attattactc gtgtgtgtgc agaaag	26
20	<210> 124 <211> 41 <212> ADN <213> <i>Escherichia coli</i> <220> <221> unión de cebador <222> (1)..(41) <223> cebador oBS199	
30	<400> 124 ctagtattac ctgcgtatta gtgacgtaat aggaggtgaag c	41
35	<210> 125 <211> 51 <212> ADN <213> <i>Escherichia coli</i> <220> <221> unión de cebador <222> (1)..(51) <223> cebador oBS200	
45	<400> 125 ctattacgct actaatagcg aggtaatact agatgtgccca gcagcatccg c	51
50	<210> 126 <211> 54 <212> ADN <213> <i>Escherichia coli</i> <220> <221> unión de cebador <222> (1)..(54) <223> cebador oBS201	
55	<400> 126 cggaggcctt tctgacaaca cagcagtaat aatctagttt tctagaggca gcgc	54
60	<210> 127 <211> 53 <212> ADN <213> <i>Escherichia coli</i>	
65	<220> <221> unión de cebador	

	<222> (1)..(53)	
	<223> cebador oBS202	
	<400> 127	
5	ctattacgtc actaatagcg aggtaatact agatgaagac aggtcgaata gtg	53
	<210> 128	
	<211> 56	
10	<212> ADN	
	<213> <i>Escherichia coli</i>	
	<220>	
	<221> unión de cebador	
15	<222> (1)..(56)	
	<223> cebador oBS203	
	<400> 128	
20	cggaggcctt tctgacaaca cagagtaat aatttaatat ctaacccttc tatttg	56
	<210> 129	
	<211> 55	
	<212> ADN	
25	<213> <i>Escherichia coli</i>	
	<220>	
	<221> unión de cebador	
	<222> (1)..(55)	
30	<223> cebador oBS210	
	<400> 129	
35	cctattacgt cactaatagc gaggaataac tagatgttta aaccgcacgt taccg	55
	<210> 130	
	<211> 50	
	<212> ADN	
	<213> <i>Escherichia coli</i>	
40	<220>	
	<221> unión de cebador	
	<222> (1)..(50)	
	<223> cebador oBS211	
45	<400> 130	
	ctttctgaca acacacgagt aataatttag atgacaccct ttgtaaaagg	50
50	<210> 131	
	<211> 37	
	<212> ADN	
	<213> <i>Escherichia coli</i>	
55	<220>	
	<221> unión de cebador	
	<222> (1)..(37)	
	<223> cebador oBS220	
60	<400> 131	
	gtacctataa tgtgtggatg tcccaccgct tacctcc	37
	<210> 132	
65	<211> 47	
	<212> ADN	

	<213> <i>Escherichia coli</i>	
	<220>	
	<221> unión de cebador	
5	<222> (1)..(47)	
	<223> cebador oBS221	
	<400> 132	
10	catcaccatc atcaccactg aattattact cgtgtgtgtg cagaaag	47
	<210> 133	
	<211> 22	
	<212> ADN	
15	<213> <i>Escherichia coli</i>	
	<220>	
	<221> unión de cebador	
	<222> (1)..(22)	
20	<223> cebador oGEN184	
	<400> 133	
	aaacctcttt atgtgcagt cg	22
25	<210> 134	
	<211> 22	
	<212> ADN	
	<213> <i>Escherichia coli</i>	
30	<220>	
	<221> unión de cebador	
	<222> (1)..(22)	
	<223> cebador oGEN182	
35	<400> 134	
	aaattcgcg gttccactaa ga	22
40	<210> 135	
	<211> 68	
	<212> ADN	
	<213> <i>Escherichia coli</i>	
45	<220>	
	<221> unión de cebador	
	<222> (1)..(68)	
	<223> cebador oGEN227	
50	<400> 135	
	ccgcttacct cctattacgt cactaatagc taaggaggta aatatgtctg caacaaaact	60
	gacccgcc	68
	<210> 136	
55	<211> 52	
	<212> ADN	
	<213> <i>Escherichia coli</i>	
	<220>	
60	<221> unión de cebador	
	<222> (1)..(52)	
	<223> cebador oGEN264	
	<400> 136	

	cggaggcctt tctgacaaca cagagtaat aattcatagt ctttgcgagg cg	52
5	<210> 137 <211> 33 <212> ADN <213> <i>Escherichia coli</i> <220> <221> unión de cebador <222> (1)..(33) <223> cebador oGEN265	
15	<400> 137 attattactc gtgtgtgtgc agaaaggcct ccg	33
20	<210> 138 <211> 31 <212> ADN <213> <i>Escherichia coli</i> <220> <221> unión de cebador <222> (1)..(31) <223> cebador oGEN266	
30	<400> 138 attagtgcgcg taataggagg taagcgggtgg g	31
35	<210> 139 <211> 52 <212> ADN <213> <i>Escherichia coli</i> <220> <221> unión de cebador <222> (1)..(52) <223> cebador oBS421	
45	<400> 139 cggaggcctt tctgacaaca cagagtaat aattcaccac caggcgtgga ag	52
50	<210> 140 <211> 60 <212> ADN <213> <i>Escherichia coli</i> <220> <221> unión de cebador <222> (1)..(60) <223> cebador oBS422	
55	<400> 140 gcttacctcc tattacgtca ctaatagcta aggaggtaaa tatgcaagtc gacctgctgg	60
60	<210> 141 <211> 90 <212> ADN <213> <i>Escherichia coli</i>	
65	<220> <221> unión de cebador	

	<222> (1)..(90)		
	<223> cebador oBS445		
	<400> 141		
5	cggaacaaag ccgtggatgt ccaacgtcat tgataatggc ccacgcgtgg tatcgccgcc	60	
	aatgagttgc atatcgtaat aattgagaag	90	
	<210> 142		
	<211> 90		
10	<212> ADN		
	<213> <i>Escherichia coli</i>		
	<220>		
	<221> unión de cebador		
15	<222> (1)..(90)		
	<223> cebador oBS446		
	<400> 142		
20	ccagagcgcg ttaaggctcg tccatcgga acaaagccgt ggatgtccaa cgtcattgat	60	
	aatggccac gcgtggtatc gccgccaatg	90	
	<210> 143		
	<211> 20		
25	<212> ADN		
	<213> <i>Escherichia coli</i>		
	<220>		
	<221> unión de cebador		
30	<222> (1)..(20)		
	<223> cebador oBS456		
	<400> 143		
35	catgtggcga gttctccctg	20	
	<210> 144		
	<211> 20		
40	<212> ADN		
	<213> <i>Escherichia coli</i>		
	<220>		
	<221> unión de cebador		
45	<222> (1)..(20)		
	<223> cebador oBS457		
	<400> 144		
50	caggtaaacg gtacgccag	20	
	<210> 145		
	<211> 20		
	<212> ADN		
55	<213> <i>Escherichia coli</i>		
	<220>		
	<221> unión de cebador		
	<222> (1)..(20)		
60	<223> cebador oBS521		
	<400> 145		
	gcaactgtcg atggcgaagc	20	

	<210> 146	
	<211> 19	
	<212> ADN	
5	<213> <i>Escherichia coli</i>	
	<220>	
	<221> unión de cebador	
	<222> (1)..(19)	
10	<223> cebador oBS522	
	<400> 146	
	cctgatcaac cgccaccac	19
15	<210> 147	
	<211> 47	
	<212> ADN	
	<213> <i>Escherichia coli</i>	
20	<220>	
	<221> promotor	
	<222> (1)..(47)	
	<223> promotor apFAB46	
25	<400> 147	
	aaaaagagta ttgacttcgc atctttttgt acctataata gattcat	47
30	<210> 148	
	<211> 37	
	<212> ADN	
	<213> <i>Escherichia coli</i>	
35	<220>	
	<221> promotor	
	<222> (1)..(37)	
	<223> promotor apFAB70	
40	<400> 148	
	ttgacatgc atctttttgt acctataatg tgggat	37
45	<210> 149	
	<211> 37	
	<212> ADN	
	<213> <i>Escherichia coli</i>	
50	<220>	
	<221> promotor	
	<222> (1)..(37)	
	<223> promotor apFAB71	
55	<400> 149	
	ttgacatgc atctttttgt acctataata gattcat	37
60	<210> 150	
	<211> 29	
	<212> ADN	
	<213> <i>Escherichia coli</i>	
65	<220>	
	<221> promotor	
	<222> (1)..(29)	
	<223> promotor pBAD ara	

	<400> 150	
	ctgacgcttt ttatcgcaac tctctactg	29
5	<210> 151 <211> 74 <212> ADN <213> <i>Escherichia coli</i>	
10	<220> <221> promotor <222> (1)..(74) <223> promotor lac con sitio de operador lacO	
15	<400> 151	
	agctgtttcc tgtgtgaaat tgttatccgc tcacaattcc acacaacata cgagccggaa	60
	gcataaaagtg taaa	74
20	<210> 152 <211> 91 <212> ADN <213> <i>Escherichia coli</i>	
25	<220> <221> terminador <222> (1)..(91) <223> terminador apFAB378	
30	<400> 152	
	gagttggttag ctcttgatcc ggcaaacaaa ccaccgttgg tagcgggtggt ttttttgttt	60
	gcaagcagca gattacgcgc agaaaaaaag g	91
35	<210> 153 <211> 91 <212> ADN <213> <i>Escherichia coli</i>	
40	<220> <221> terminador <222> (1)..(91) <223> terminador apFAB377	
	<400> 153	
	atgaccatct acattactga gctaataaca ggcttgcctgg taatcgcagg cctttttatt	60
45	tgggggagag ggaagtcattg aaaaaactaa c	91
50	<210> 154 <211> 90 <212> ADN <213> <i>Escherichia coli</i>	
55	<220> <221> terminador <222> (1)..(90) <223> terminador apFAB381	
	<400> 154	

	accctcaaga gaaaatgtaa ccaactcact ggctcacctt cacgggtggg cctttcttcg	60
	ttccgggcat taaccctcac taacaggaga	90
	<210> 155	
	<211> 31	
5	<212> ADN	
	<213> Sintético	
	<220>	
	<221> unión de cebador	
10	<222> (1)..(31)	
	<223> cebador oBS234	
	<400> 155	
15	aggcctutct gacaacacac gagtaataat t	31
	<210> 156	
	<211> 26	
	<212> ADN	
20	<213> Sintético	
	<220>	
	<221> unión de cebador	
	<222> (1)..(26)	
25	<223> cebador oBS1720	
	<400> 156	
	cgtcactaat agcgaggttaa tactag	26
30	<210> 157	
	<211> 42	
	<212> ADN	
	<213> Sintético	
35	<220>	
	<221> unión de cebador	
	<222> (1)..(42)	
	<223> cebador oBS1721	
40	<400> 157	
	ctagtattac ctcgctatta gtgacgataa taggaggttaa gc	42

REIVINDICACIONES

1. Bacteria modificada genéticamente para la producción de tiamina no fosforilada; en la que dicha bacteria se caracteriza por tener transgenes que codifican para:

a. un polipéptido que tiene actividad tiamina monofosfato fosfatasa (E.C.3.1.3.-), en la que la secuencia de aminoácidos de dicho polipéptido tiene al menos el 80% de identidad de secuencia con respecto a una secuencia seleccionada del grupo que consiste en SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72 y 74;

b. un polipéptido que tiene actividad 4-amino-5-hidroximetil-2-metilpirimidina fosfato sintasa (HMP-P) (E.C.4.1.99.17);

c. un polipéptido que tiene actividad tiamina fosfato sintasa (E.C.2.5.1.3);

d. un polipéptido que tiene actividad adenilil transferasa de ThiS (E.C.2.7.7.73);

e. un polipéptido que tiene actividad portadora de azufre de ThiS;

f. un polipéptido que tiene actividad tiazol sintasa (E.C.2.8.1.10);

g. un polipéptido que tiene actividad 2-iminoacetato sintasa (4.1.99.19) o que tiene actividad glicina oxidasa (EC.1.4.3.19); y

h. un polipéptido que tiene actividad fosfohidroximetilpirimidina cinasa (E.C.2.7.4.7),

en la que el género de la bacteria es *Escherichia*.

2. Bacteria modificada genéticamente según la reivindicación 1, caracterizada además por un gen *thiL* endógeno modificado genéticamente que puede expresar actividad tiamina fosfato cinasa (EC.2.7.4.16) reducida en comparación con el gen *thiL* endógeno parental.

3. Bacteria modificada genéticamente según la reivindicación 2, en la que el gen *thiL* endógeno modificado genéticamente codifica para un polipéptido que tiene actividad tiamina fosfato cinasa (EC.2.7.4.16) reducida en comparación con el polipéptido codificado por el gen *thiL* endógeno parental de tipo natural.

4. Microorganismo modificado genéticamente según una cualquiera de las reivindicaciones 1 a 3, en el que dicha bacteria se caracteriza por la inactivación o delección de uno o más genes que codifican para una proteína seleccionada del grupo: proteína de unión periplásmica de transportadores ABC de tiamina, permeasa de transportadores ABC de tiamina y ATPasa de transportadores ABC de tiamina.

5. Microorganismo modificado genéticamente según una cualquiera de las reivindicaciones 1 a 4, en el que dicha bacteria se caracteriza por un transgén adicional que codifica para un polipéptido que tiene actividad hidroxietiltiazol cinasa (E.C.2.7.1.50).

6. Bacteria modificada genéticamente según una cualquiera de las reivindicaciones 1 a 5, en la que la secuencia de aminoácidos del polipéptido que tiene actividad tiamina monofosfato fosfatasa (E.C.3.1.3.-) tiene al menos el 80% de identidad de secuencia con respecto a una secuencia seleccionada del grupo que consiste en SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28 y 30.

7. Bacteria modificada genéticamente según una cualquiera de las reivindicaciones 1 a 6, en la que la secuencia de aminoácidos de:

a. el polipéptido que tiene actividad 4-amino-5-hidroximetil-2-metilpirimidina fosfato sintasa (HMP-P) (E.C.4.1.99.17) tiene al menos el 80% de identidad de secuencia con respecto a una secuencia seleccionada del grupo de SEQ ID NO: 76, 78, 80 y 82;

b. el polipéptido que tiene actividad tiamina fosfato sintasa (E.C.2.5.1.3) tiene al menos el 80% de identidad de secuencia con respecto a SEQ ID NO: 84;

c. el polipéptido que tiene actividad adenilil transferasa de ThiS (E.C.2.7.7.73) tiene al menos el 80% de identidad de secuencia con respecto a SEQ ID NO: 86;

d. el polipéptido que tiene actividad portadora de azufre de ThiS tiene al menos el 80% de identidad de secuencia con respecto a SEQ ID NO: 88;

e. el polipéptido que tiene actividad tiazol sintasa (E.C.2.8.1.10) tiene al menos el 80% de identidad de secuencia con respecto a SEQ ID NO: 90;

f. el polipéptido que tiene actividad 2-iminoacetato sintasa (E.C.4.1.99.19) tiene al menos el 80% de identidad de secuencia con respecto a SEQ ID NO: 92 y el polipéptido que tiene actividad glicina oxidasa (EC 1.4.3.19) tiene al menos el 80% de identidad de secuencia con respecto a una secuencia seleccionada del grupo de SEQ ID NO: 94, 96 y 98; y

g. el polipéptido que tiene actividad fosfohidroximetilpirimidina cinasa (E.C.2.7.4.7) tiene al menos el 80% de identidad de secuencia con respecto a SEQ ID NO: 100.

8. Bacteria modificada genéticamente según una cualquiera de las reivindicaciones 5 a 7, en la que la secuencia de aminoácidos del polipéptido que tiene actividad hidroxietiltiazol cinasa (E.C.2.7.1.50) tiene al menos el 80% de identidad de secuencia con respecto a SEQ ID NO: 108.

9. Bacteria modificada genéticamente según una cualquiera de las reivindicaciones 2 a 8, en la que la secuencia de aminoácidos del polipéptido que tiene actividad tiamina fosfato cinasa (EC.2.7.4.16) tiene al menos el 80% de identidad de secuencia con respecto a SEQ ID NO: 110 con la condición de que la secuencia tenga una sustitución G133D de residuos de aminoácido.

10. Bacteria modificada genéticamente según una cualquiera de las reivindicaciones 1 a 9, en la que la secuencia de aminoácidos del polipéptido que tiene actividad tiamina monofosfato fosfatasa (E.C.3.1.3.-) tiene al menos el 80% de identidad de secuencia con respecto a una secuencia seleccionada del grupo de SEQ ID NO: 2, 14, 26, 32, 34, 40, 68 y 70.

11. Método para producir tiamina no fosforilada que comprende las etapas de:

a. introducir una bacteria modificada genéticamente según una cualquiera de las reivindicaciones 1-10 en un medio de crecimiento para producir un cultivo;

b. cultivar el cultivo; y

c. recuperar la tiamina no fosforilada producida por dicho cultivo, y purificar opcionalmente la tiamina no fosforilada recuperada.

12. Uso de un transgén que codifica para un polipéptido que tiene actividad tiamina monofosfato fosfatasa (E.C.3.1.3.-) para potenciar la producción de tiamina no fosforilada y la exportación en una bacteria *Escherichia*.

13. Uso de un transgén que codifica para un polipéptido que tiene actividad tiamina monofosfato fosfatasa (E.C.3.1.3.-) según la reivindicación 12, en el que la secuencia de aminoácidos de dicho polipéptido tiene al menos el 80% de identidad de secuencia con respecto a una secuencia seleccionada de entre:

a. el grupo que consiste en SEQ ID NO: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28 y 30;

b. el grupo que consiste en SEQ ID NO: 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64 y 66; y

c. el grupo que consiste en SEQ ID NO: 68, 70, 72 y 74.

14. Uso de un transgén que codifica para un polipéptido que tiene actividad tiamina monofosfato fosfatasa (E.C.3.1.3.-) según la reivindicación 12 ó 13, en el que la bacteria *Escherichia* es una bacteria modificada genéticamente caracterizada por tener transgenes que codifican para:

a. un polipéptido que tiene actividad tiamina monofosfato fosfatasa (E.C.3.1.3.-);

b. un polipéptido que tiene actividad 4-amino-5-hidroximetil-2-metilpirimidina fosfato sintasa (HMP-P) (E.C.4.1.99.17);

c. un polipéptido que tiene actividad tiamina fosfato sintasa (E.C.2.5.1.3);

d. un polipéptido que tiene actividad adenilil transferasa de ThiS (E.C.2.7.7.73);

e. un polipéptido que tiene actividad portadora de azufre de ThiS;

f. un polipéptido que tiene actividad tiazol sintasa (E.C.2.8.1.10);

g. un polipéptido que tiene actividad 2-iminoacetato sintasa (E.C.4.1.99.19) o que tiene actividad glicina oxidasa (EC

1.4.3.19); y

h. un polipéptido que tiene actividad fosfohidroximetilpirimidina cinasa (E.C.2.7.4.7).

- 5 15. Uso de una bacteria modificada genéticamente según una cualquiera de las reivindicaciones 1-10, para la producción de tiamina no fosforilada.

Figura 1

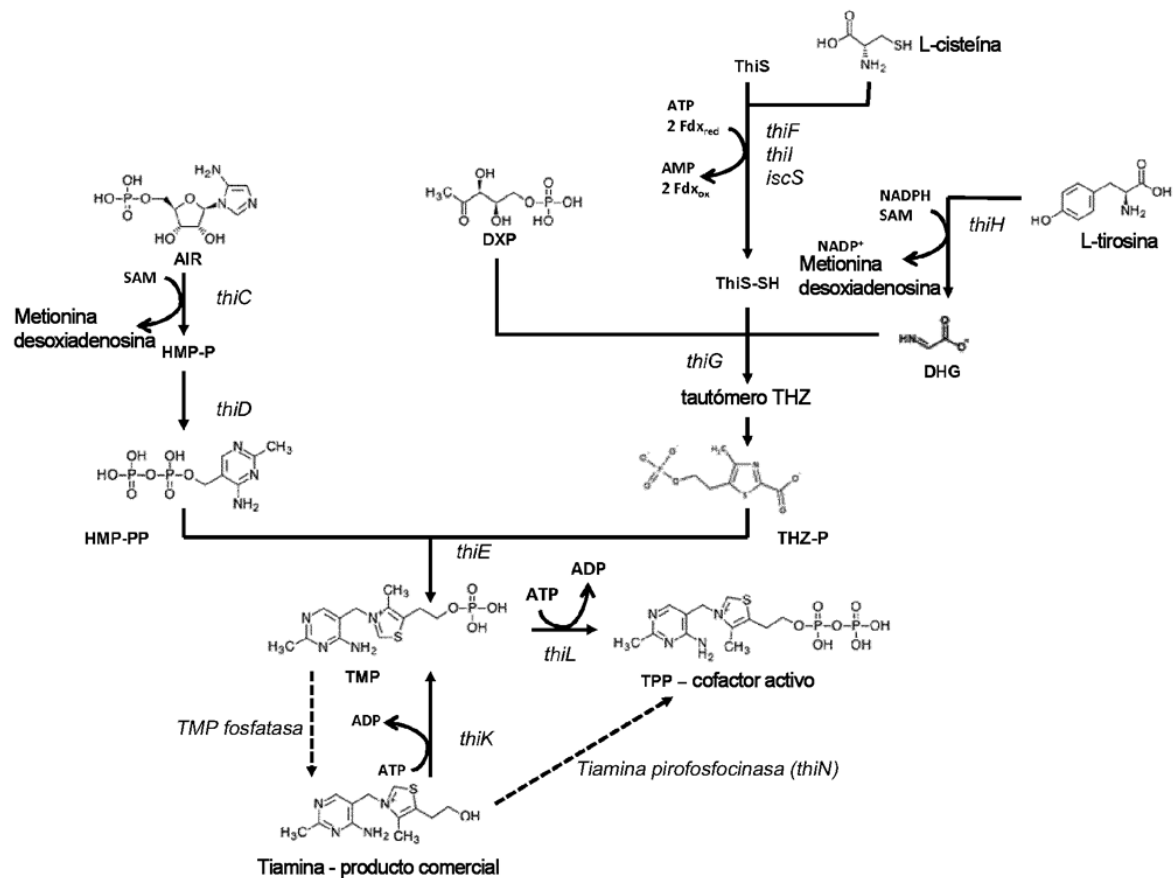


Figura 2

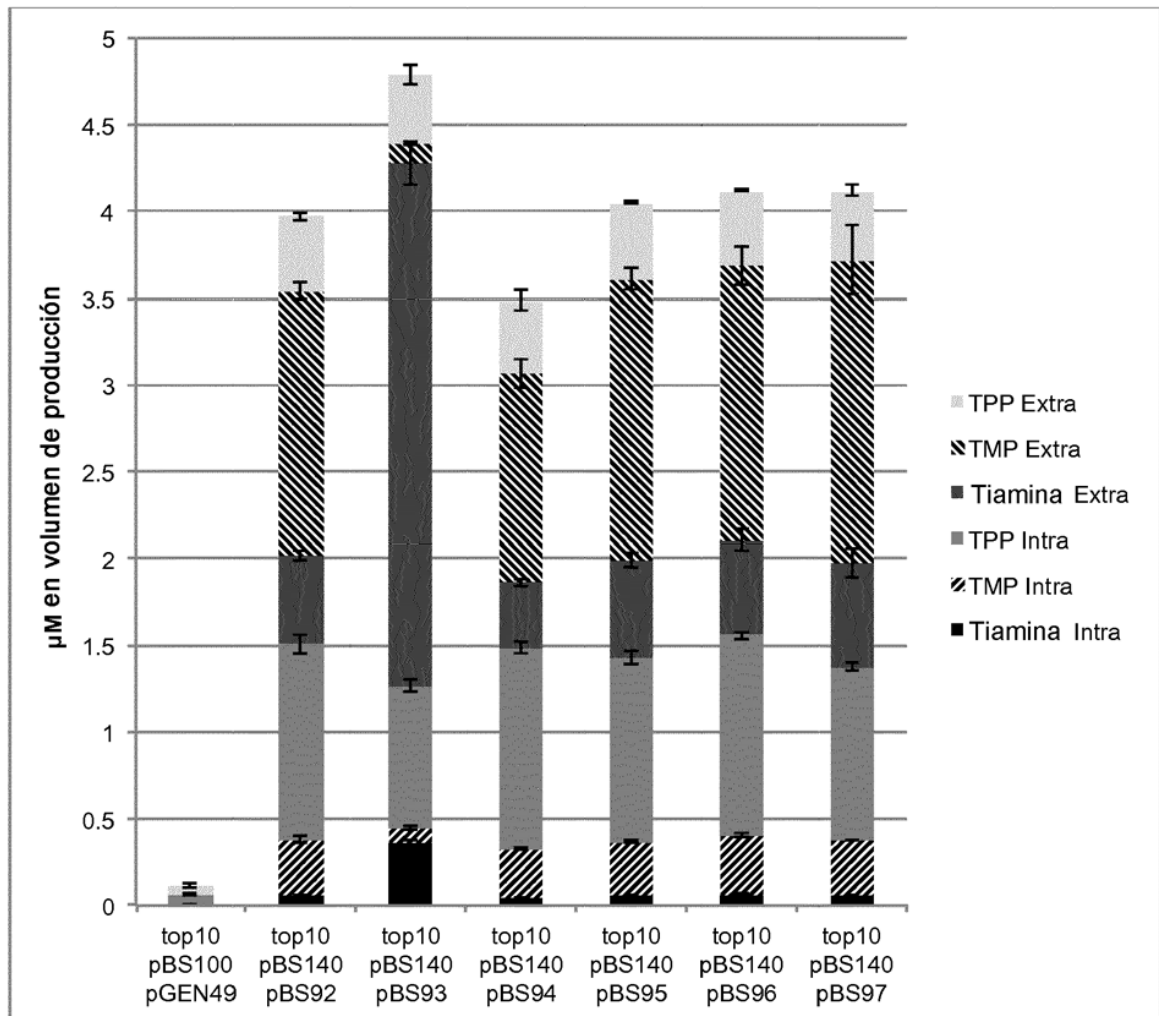


Figura 3

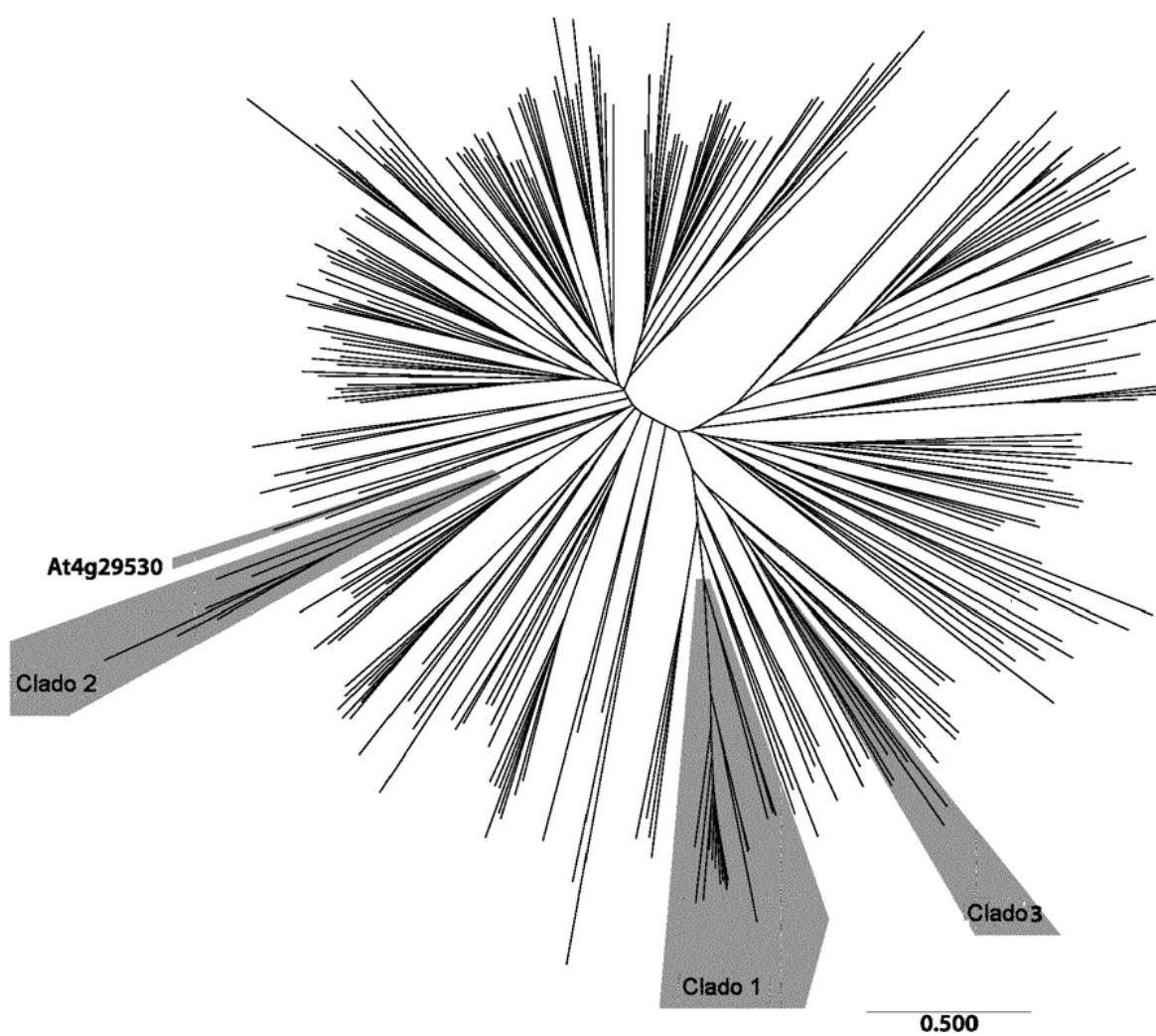


Figura 4A

Clado 1

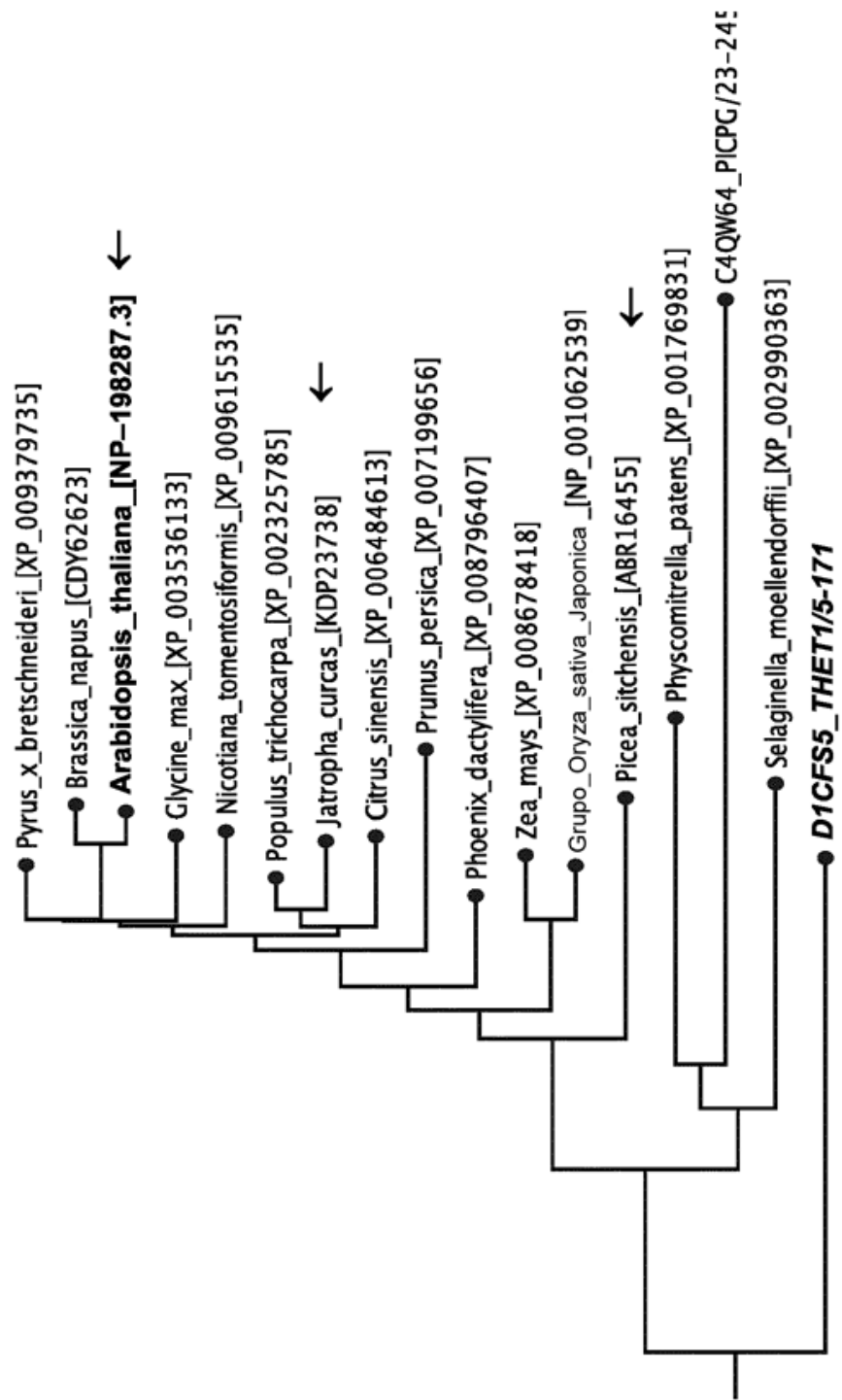


Figura 4B

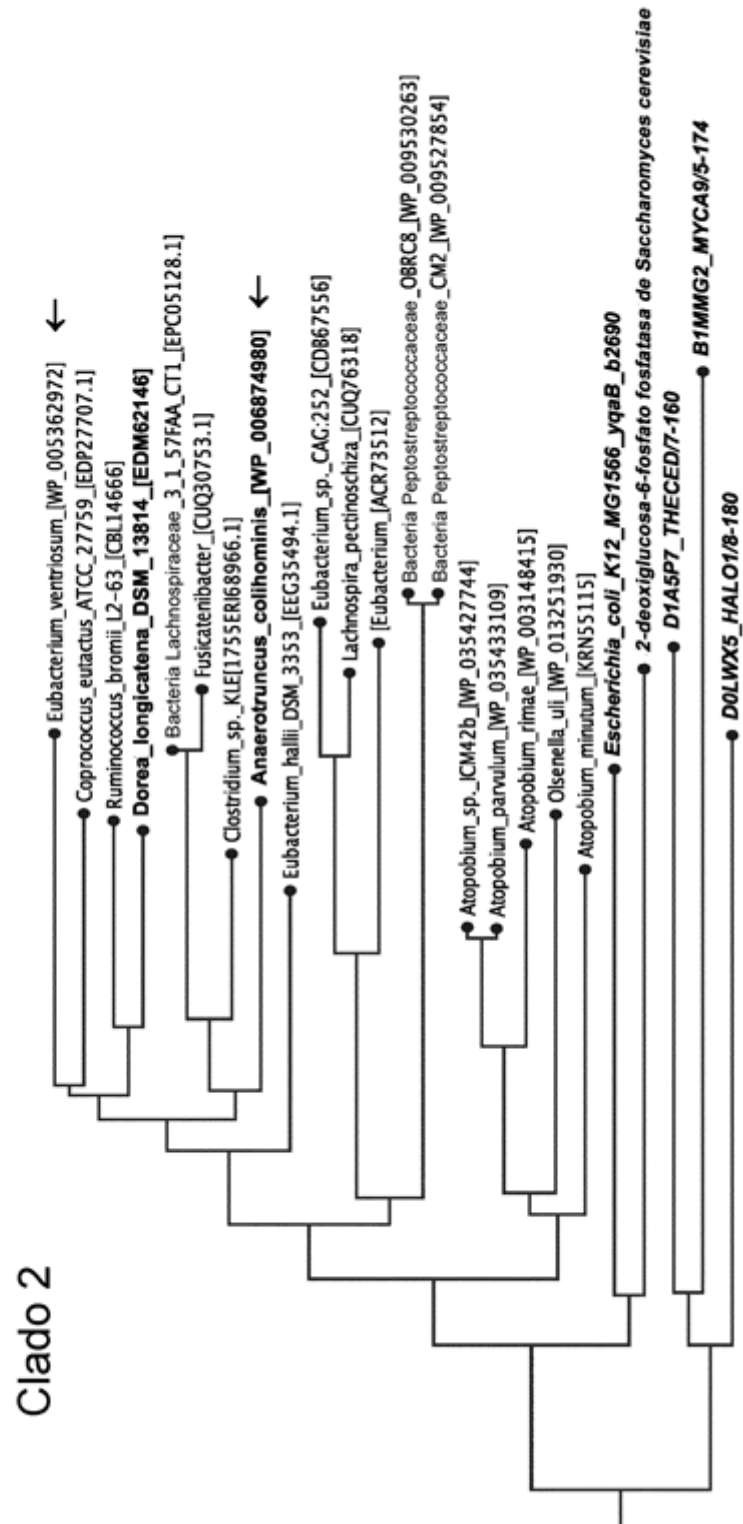


Figura 4C

Clado 3

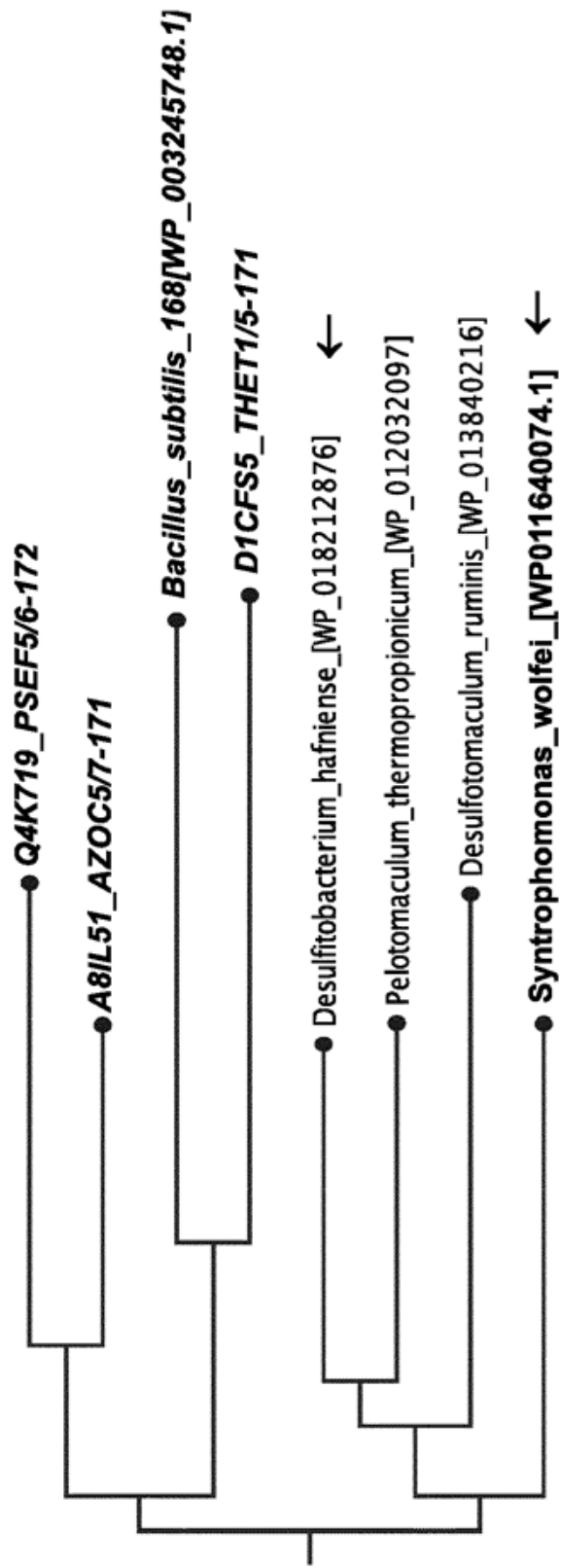


Figura 5

1. *A. thaliana*
2. *B. napus*
3. *C. sinensis*
4. *G. max*
5. *J. curcas*
6. *N. tomentosiformis*
7. *O. sativa*
8. *P. bretschneideri*
9. *P. persica*
10. *P. trichocarpa*
11. *Z. mays*

```

      1      10      20      30      40      50
1  MRFLFPTRLINNSSLGLLRSPHTTAPIRSLWFRTKSPVFRSATTP----IMTAVA
2  MRILNNSLALLRSPRAAAPIRSLLFSGKKSSVSR-----AAAFSS
3  MRFLFTNPIKTPLLSSILFHCPNSPRLGLLDSVRVNSPSSLTT-----QRSLS
4  MRMRFRLRSPIIKTSLNLSPPIISFRPHWARRTFTSSRLS
5  MAIPPKLASSSSSM
6  MRFSLLSPLVLNPVIRFSNSNALFGLRFQLYPRYSRYLRSPVTMA----SAKPKP
7  MRGLLRRVYLRLLPFPFPAATSLYYWSRTRPAAAGP----NHPIPRR
8  MRILFPPNPIKTPTLFNSLRLRFNSLRSHC--ANSMVPP
9  MA
10 MRLLLFTSPNPIKTSSSLYFLNSLRSNLTKRTLPTRRSFIPARMAIPPRSIASAPSC
11 MLVLRRLRLRLPLPRPLLVSFSSTSPSSSPSTSSSSSCWSSTG--ESRRAMAS

      60      70      80      90      100     110
1  FSSSLSIPPTSE-EALPGKLWI-KFNRECLFSIYSPFAVCLAAGNLKIDT-FRQYIA
2  AMSIPPPSISTSEEALAGRLWI-KFNRECLFSMYSPPFAVSLAAGNLKIET-FRQYIA
3  MAAIPPKSPSPEEGLARRLWI-KFKRESVFAMYSPFTVCLASGNLKLET-FRHYIA
4  MAAIHNSNSNSETGLARRFWI-KFTRESIFAMYTFFAIALASGNLHIDS-FHHYIA
5  AASPTSAGGTNE-EGLASKFWI-KFRRESVLAMYTFFVVSFAAGNLKIES-FRHYIA
6  AAAVNKFPVEEECVGIARKCWI-KFKRES TFALYTPFVVSLASGTLNLDL-FRHYIA
7  MSTSSATAAVVAEGSAARRFWIAAASREAAFAAYTPFLVSLAAGALRLDS-FRQYIA
8  PKSAMASAVVGNEVGLARRFWI-KFKRESIFAMYTPTLCLAGNLKIET-FRDYIA
9  ALARHSIVRLNHEGGLARRLWF-KFRDDSVFSLYTPFFVGLASATLHSETTFRHFIS
10 TTTSGRSNINIE-EGLASKFWI-KFRRESVFAMYTFFVISLASGTLKIDS-FRHYIS
11 SPSPDSAAVVAE-GSAARRFWIAASTREAAFAAYTPFLLSLAAGNLRLNV-FRHYIA

      120     130     140     150     160     170
1  QDVHFLKAFAHAYELAADCADD DDDKLAI S DLRKS VM EELKMHDSFVQDWDLD-INK
2  QDVHFLKAFAHAYELAAECADD DDDKLAI S DLRKS VMDELKMHNSFVQDWDLD-ISK
3  QDFHFLKAFHQAYELAEBCADD DDAKLSI SELRKGVLEELKMHDSFVKEWGTD-LAK
4  QDVHFLRAFAQAYELAEBCADD DDAKLGI CELRKAVLEELKMHNLVQERELD-LAK
5  QDFHFLKAFAHAYELAEBCADD DDAKLAI AALRKGVLEELKLEKSFVQEWGMD-PSK
6  QDVHFLKSFAQAYEAAEECTDD DDAKVGI SELRKNVIEELKMHDAVLKEWGID-LVK
7  QDAYFLHAFARAYEMAEECADD DDDKATI VVLRKAILRELNLHASVLQEWGVD-PNK
8  QDVHFLKAFAHAYELAEDCADD DDAKPVI SELRRAVLQELKMHDSFVKEWGLQG-AK
9  QDLHFLKAFVLAYELAEDCADD EDDKNGLRDLRKRAVGRLQMHDTFVREWGFEFPNE
10 QDSHFLKSFAHAFELAEECADD DEAKLAI SELRKGVLEELKMHNSFVQEWGID-PGK
11 QDAHFLHAFARAYEMAEDCADD DDDMATI AALRKAILQELNLHSSVLKEWGVD-PTK

```


Figura 5 continuación

	180	190	200	210	220	
1	EVSVNSATLRYTEFLLATASGKVEGCKAP-GMLDTPFEKTKVAAAYTLGAVTPCMRLY					
2	EVSVNSATLRYTEFLLATSSGKVEGLKAP-GMLDTPFEKTKVAAAYTLGAVTPCMRLY					
3	MATVNSATVKYTEFLLATASGKVEGVKGP-GKLATPFEKTKVAAAYTLGAMSPCMRLY					
4	EHGINSATVKYTEFLLATASGKIEGLKGP-GKLATPFEKTKIAAYTLGAMTPCMRLY					
5	EVTINSATAKYTDFLLATASGKVEGVKGP-GKLATPFERTKVAAYTLGTMTPCMRLY					
6	ESSLNPATAKYTDFLSATASGKVEGVKAA--KLATPFERTKLAAYTLGAMTPCMRLY					
7	EIPPIPATTKYTDFFLLATSTGKVDGKGKS-DKMVTPEKTKIAAYTVGAMTPCMRLY					
8	ETPINSAAVKYTDFFLLATASGKVEGVKGP-GKLATPFERTKVAAYTLGAMTPCMRLY					
9	DISKDIATTKYTDFFLLATASGKIEGERSVLDKIATPFEKTKVAAAYTLAALAPCMRLY					
10	EGTINSATVKYTDFFLLATASGKVEGVKGL-GKLATPFERTKVAAYTLGAMTPCMRLY					
11	EIPPSAATTKYTDFFLLATAAGKVDGKGKS-DKMVTPEKTKIAAYTVGAMTPCMRLY					
	230	240	250	260	270	280
1	AFLGKEFGSLLDLS-DVNHPYKKWIDNYS SDAFQASAKQTEDLLEKLSVSMTGEELD					
2	AFLGKEFGALLDSS-EANHPYKKWIENYS SDAFQASAKQTEDLLEKLSVCMTGEELD					
3	AFLGKEFHGLLNAN-EGNHPYKKWIDNYS SESFQASALQNE DLLDKLSVSLTGEELD					
4	AVMGKKFQELDSN-ESTHPYNKWIDNYS SDGFQATTLQTEDLLEKLSVSLTGEELD					
5	AFLAKELQALIDAE-AGIHPYQKWIDNYS SESFQASALQTEDLLEKLSVPLTGEELD					
6	AYIGKELQVFLGE--KIHPYKKWIDSYASESFQASALQTEDLLEKLSVPLTGEELD					
7	AYLGKELAVFLKQD--ENHPYKKWIETYASSDFENNALQIEELLDKLSVSLTGEELD					
8	AFLGKEFKALLDPS-EGSHPYLKWIDSYSSKSFQASAVQIEELLDKLSVSLTGEELD					
9	AFISTEIQGIINPDQDSTHIYKSWIENYS SQVFEEIALQNE DMLDKLSVSLTGEELD					
10	SFLGKELQAVLDPE-EDGHPYKKWIDSYSSSFQASALQTEDLLEKLSVSLTGEELD					
11	AYLGKELMVFLKQD--ENHPYKKWINTYASSDFEDTTLQIEELLDKLSVSLTGEELD					
	290	300	310	320	330	340
1	IIEKLYQQAMKLEVEFFPHAQPL-AQPTIVPLLKNHS--KDDLVIIFSDFDLTCTVVD					
2	IIEKLYQQAMKLEVEFFPHAQPF-AQPTIVPLLKNHS--KDELMIFSDFDLTCTVVD					
3	IIEKLYHQAMKLEVEFFCAQPL-AQPTVVPLIKGHNPAGDRLIIFSDFDLTCTIVD					
4	VIEKLYYQAMKLEIEFFSAQPL-FQPTIVPLTKGHKPVEDHLIIFSDFDLTCTVVD					
5	IIEKLYHQAMKLEIEFFNAQPL-DQPTVVPLTKENPLEDRLVIFSDFDLTCTVVD					
6	IIEKLYHQAMKLEIDFFLTQPL-VQKAVIPLSKDHNPAEHRLTIFSDFDLTCTVVD					
7	IIGKLYQQAMRLEVEFFSAQPV-DQPVVAPLSRYCGP-KDKLLIFCDFDLTCTVVD					
8	IIEKLYHQAMKLEIEFFSAQSL-VQPTVVPLIREHNPAEDRLMIFSDFDLTCTVVD					
9	IIEKLYHQAMKLQVDFIAAQPI SDQQSVVPLSRVHDFSKRHILTILCDFDLACTAFDS					
10	IIEKLYHQAMKLEIEFFLAQPI-AQTTLAPLTKGHNPEDRLVIFSDFDLTCTVVD					
11	IIGKLYQQAMKLEVEFFSSQLI-DQPVVAPLSRYCDP-KYKLLIFSDFDLTCTIVD					
	350	360	370	380	390	
1	SAILAEIAIVTAPKDEQSRSGQ-QIHRMLSSDLKNTWNLLSKQYTEHYEECIESILN					
2	SAILAEIAIVTAPKDDQGQ---QINRMLSADLKNWLSLLSKQYTEHYEECIESILN					
3	SAILAEIAIVTAPKSDQNQPEN-QLGRMSSGELRNTWGLLSKQYTEHYEQCIESFMP					
4	SAILAEIAIVTAPKSDQNQPEN-QIVRMLSSDLRNTWGLLSKQYTEHYEQCIESIMP					
5	SAILAEIAILTASKSDQSQSDN-QNARMSSTELRNTWVLLSGQYTEHYEQCIESILP					
6	SAILAEIAIITAPRSDQNRPEN-QIARMLSADLRNTWGDLSKQYTEHYEQCIEKMLL					
7	SAILAEIAILSHQKASQGGADS-SLDRTKSADLRNSWNMLS NQYMEHYEQCIASLLP					
8	SAILAEIAIVTAPKSDQHQPEN-QIARMSADLRNTWGLLSRQYTEHYEQCIESIVP					
9	AAILAEIAIITAPKADMDGSDQTQLARMP SADLRSTWDVLS TQYTEQFEQCVESIVA					
10	SAILAEIAILTAPKSDVVQPET-QIARMSADLRNTWGLLSGQYTEHYEQCIESIMP					
11	SAILAEIAILSFQKANQSGIDN-NLDRAKSGDLRSWNMLS KQYMEHYEKCIMERLLP					

Figura 5 continuación

```

400      410      420      430      440      450
1  KKKA-----DKFDYEGLCCKA
2  KEKA-----DKFDYEGLCCKA
3  SEKV-----NPNYETLHKA
4  PDRL-----NFDYKELSHMA
5  SEKMEFNFEALCKALEQLSDFERRANARVIKSGVLKGLNLEDIKRAVEFNFEALCKA
6  TEKA-----EKFDYERLHKT
7  PEEA-----SLDYDQLYKG
8  TEKA-----VFDYENLLKA
9  SERV-----EFDYERLCSA
10 SAKV-----E-FNYEALCKA
11 PEES-----KSLDYDKLYKG

460      470      480      490      500      510
1  LEQLSDFEKEANNRVIESGVLKGLNLEDIKRAGERLILQDGCINVFQKILKT-ENLN
2  LEQLSEFEKKANDRVIESGVLKGLNLDIIRKAGERLILQDGCINVFQKILKT-QDVN
3  LEQLSHFEKRANSRVIESGVLKGLNLEDIKRAGERLSLQDGCCTFFQKVVKN-ENLN
4  LEQLSKFENTANNRVIESGVLKGLNLEDIKRAGERLILQDGCNFFQKIVKN-ENLN
5  LEQLSDFERRANARVIKSGVLKGLNLEDIKRAGERLILQDGCSTFFQKISK-ENLN
6  LEELSDFEKRANTRVIESGVLKGLNLEDIKRAGERLILQDGCNFFQKIVKN-ENLN
7  LEVLSQFEKLANSRVVDSGVLKGLNLDIIRKAGERLILQDGCCKIFFQKIGKTRENLN
8  LEKLSDFEKRANRVTKSEVLKGLNLEDIKRAGERLILQDGCINFFQKIAKS-ENLN
9  LEQLAEFERKANERVVQSGVLKGLNLEDIKRAGERLILQDGCSTFFQKIVKN-KNLK
10 LEQLSDFERRANSRVVDSGVLKGLNLEDIKRAGERLILQDGCIGFFQKIVKN-ENLN
11 LEVLAEFEKLANSRVVDSGVLKGLNLEDIKRAGERLILQDGCCKNFFQKIVKTRENLN

520      530      540      550      560      570
1  AELHVLSCYWCGLIRAAFSAGG-VDAVEVHANEFTFEESIISTGEIERKVESPINKA
2  AKLHVLSCYWCGLIRAAFSARG-VDAVEVHANEFTFEESIISTGEIERKVESPIDKA
3  ANVHVLSCYWCGLIRASFSSAG-LNALNVHANEFSPKESIISTGEIERKVESPIDKV
4  ANVHVLSCYWCGLIRSTFSSAD-LNELNVHANEFTEGVSSTGEIVKVESPIDKV
5  ANIHVLSCYWCGLIRSAFSSGG-LDVNLHANEFDFVESIISTGEIIMKVETPTDKA
6  ADIHVLSCYWCGLIRSSFSSGG-IDALNVHANEFMFQESLSTGEIVKVESPIDKV
7  LDVHVLSCYWCGLIRSAFSSVGCGLDGLNIHSNEFAFEGSVSTGHINRQMESPLDKA
8  ANVHVLSCYWCGLIRSAFSSGG-LNELDVHANEFTFEESIISTGDIVKVESPIDKV
9  TDVHVLSCYWCGLIRSAFSSGD-LNVLNVHSELVYQESVTTGEIVKVMESPMKEL
10 TNDVHVLSCYWCGLIRSAFSSGG-LDALNVHANEFIFEESIISTGEIISTGEIISTGEI
11 LDVHVLSCYWCGLIRSAFSSAGCLDGLNIHSNEFAFEDSVSTGEIDRKMQSPLDKV

580      590      600      610      620
1  QQFKSILQNRKNENNKKSFLSVYIGDSVGDLLCLLEADIGIVVSSSSSLRRVGS HFG
2  QQFKSILQNRKK-DEEKSLTVYIGDSVGDLLCLLEADIGIVVASSSSSLRRVGS HFG
3  QAFNNTLEKY---GTDKKNLSVYIGDSVGDLLCLLEADIGIVIGSSSSSLRRVGS QFG
4  BAFRNILKNC---NDDKKKLTVYIGDSVGDLLCLLEADVIGIVIGSSSSSLRSVGT QFG
5  QAFNNILMNY S---PDKKNLTVYIGDSVGDLLCLLEADIGIVIGSSSSSLRRVGS QFG
6  QAFSKIRMNC---GNDQKNLTLYIGDSVGDLLCLLEADVIGIVLGTSSSLRTVGN HFG
7  EKFKSISDV---GSTGTLLSVYIGDSVGDLLCLLEADIGIVVGSSTTLRRVGK QFG
8  KSPKDLKNC---SNDKKNLTVYIGDSVGDLLCLLEADIGIVIGSSSSSLRRVGT QFG
9  QVFNDVLIDRR---GEGNKLTVYIGDSVGDLLCLLEADIGIVVGSSSSLRRLGD HFG
10 -----N-----LTVYIGDSVGDLLCLLEADIGIVVGSASLSRVGS QYG
11 EKFKSIRSDV---DSTVPFLSVYIGDSVGDLLCLLEADIGIVIGSTTSLRRVGK QFG

630      640      650      660      670      680
1  VSFVPLFSGIVQKQKQHTRESSSSAW-KGLSGTLYTVSSWAEIHSFALGWE
2  VSFVPLFSGIVQKQKQ-----EETW-KGLSGTLYTVSSWAEIHSFALGWE
3  VTFIPLYPGLVKKQKQKEYTEGS-SSNW-KEKSGILYTVSSWAEVHAFILGW
4  ISFVPLYSGLVKKQKQKEYVEGS-TSDW-KGLSGILYTVSSWAEVHAFILGC
5  VTFIPLYPGLVKKQKQKEYTEGS-SWNW-KGQSGVLYTVSSWAEIHSFVLGW
6  VSFVPLFPGLVQKQKQKCTGVDS SSCW-KGLSGVLYTASSWAEIHSFVLGS
7  VSFVPLFTGLVEKQRRIEKEE--SSIFKARSILYTVSSWSEVQAFILGNDFS
8  VSFVPLFPGLVKKQKQKCTDGR-SPSW-KGLTGILYTVNSWAEIHSFALGC
9  VSFVPLFSGLVKKQKQKELADQDCASNWWKPLSGVLYTVSSWAEIHSFALGT
10 VSFVPLFPGLVKKQKQKESDGS--PNW-KGLSGILYTVSSWSEIHSFALGW
11 VSFVPLFPGLVEKQKQKQKELADQDCASNWWKPLSGVLYTVSSWSEIHSFVLGSDFS

```

Figura 6

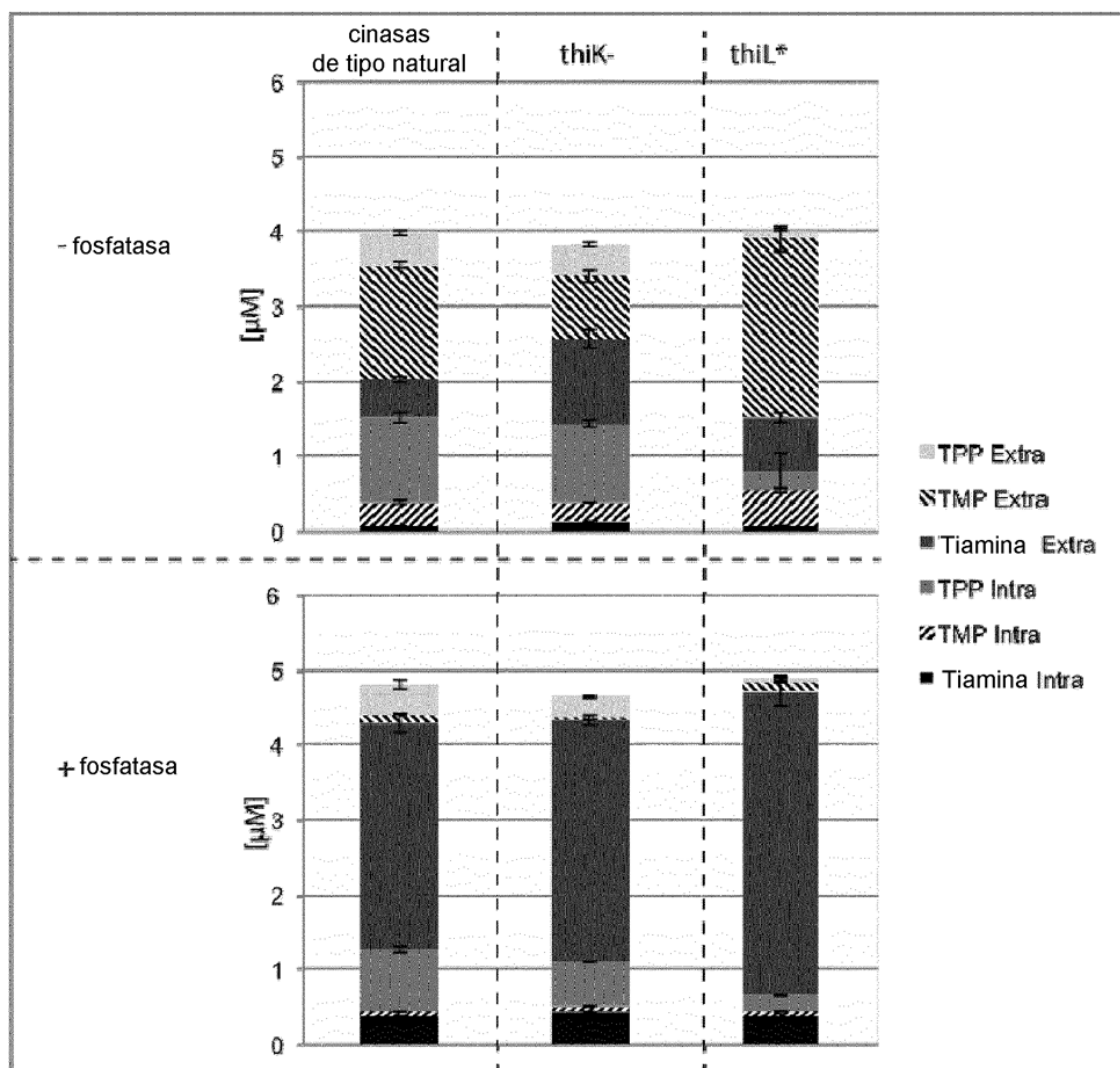


Figura 7

