

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 951 442**

51 Int. Cl.:

C12N 9/10 (2006.01)

C12P 21/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **18.03.2020 PCT/US2020/023415**

87 Fecha y número de publicación internacional: **24.09.2020 WO20191088**

96 Fecha de presentación y número de la solicitud europea: **18.03.2020 E 20718990 (3)**

97 Fecha y número de publicación de la concesión europea: **07.06.2023 EP 3942028**

54 Título: **Métodos de producción de bioconjugados de polisacáridos de *E. coli* O-antígeno, sus composiciones y métodos de utilización**

30 Prioridad:

18.03.2019 US 201962819762 P

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

20.10.2023

73 Titular/es:

JANSSEN PHARMACEUTICALS, INC. (50.0%)

1125 Trenton-Harbourton Road

Titusville, NJ 08560, US y

GLAXOSMITHKLINE BIOLOGICALS S.A. (50.0%)

72 Inventor/es:

GEURTSSEN, JEROEN;

BURGHOUT, PIETER, JAN;

WEERDENBURG, EVELINE, MARLEEN;

POOLMAN, JAN, THEUNIS;

FAE, KELLEN, CRISTHINA;

IBARRA YON, PATRICIA;

ABBANAT, DARREN, ROBERT;

KEMMLER, STEFAN, JOCHEN;

KOWARIK, MICHAEL, THOMAS;

MALLY, MANUELA;

GAMBILLARA, FONCK, VERONICA;

BRAUN, MARTIN, EDWARD y

CARRANZA SANDMEIER, MARIA, PAULA

74 Agente/Representante:

SÁEZ MAESO, Ana

ES 2 951 442 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

Métodos de producción de bioconjugados de polisacáridos de *E. coli* O-antígeno, sus composiciones y métodos de utilización

Antecedentes de la invención

5 Las cepas patogénicas extraintestinales de *Escherichia coli* (ExPEC) son normalmente habitantes inocuos del tracto gastrointestinal humano, junto con las cepas de *E. coli* comensales. Los aislados de ExPEC no pueden distinguirse fácilmente de los aislados comensales por serotipo, aunque muchos linajes clonales están dominados por ExPEC, como se define por serotipos de O-antígeno, cápsula y antígeno flagelar (abreviados como O:K:H, por ejemplo O25:K1:H4). En contraste a *E. coli* comensales, las cepas de ExPEC expresan una amplia gama de factores de virulencia que les permitan colonizar el tracto gastrointestinal, así como causar una amplia gama de infecciones extraintestinales, que se asocian con una carga importante en materia de costes sanitarios debido a hospitalizaciones y muertes. Los recién nacidos, los pacientes de edad avanzada y las personas inmunocomprometidas son especialmente susceptibles a la infección por ExPEC, incluyendo la enfermedad por ExPEC invasiva (IED).

10 Las cepas de ExPEC son la causa más común de infecciones del tracto urinario (UTI) y contribuyentes importantes a infecciones en el sitio quirúrgico y la meningitis neonatal. Las cepas también se asocian con infecciones abdominales y pélvicas y neumonía nosocomial, y ocasionalmente están implicadas en otras infecciones extraintestinales, tales como osteomielitis, celulitis e infecciones de las heridas. Todos estos sitios primarios de infección puede causar bacteriemia por ExPEC. ExPEC es la causa más común la mayor parte de bacteriemia de inicio en la comunidad y un importante patógeno causante de bacteriemia nosocomial y se encuentra en aproximadamente el 17% al 37% de los aislados sanguíneos clínicamente significativos. Los pacientes con un hemocultivo positivo para ExPEC normalmente sufren síndrome de sepsis, sepsis grave o shock séptico. Se ha observado un aumento de la resistencia de ExPEC contra antibióticos de primera línea, incluyendo las cefalosporinas, fluoroquinolonas y trimetoprima / sulfametoxazol. La aparición y rápida diseminación mundial de ExPEC secuencia de tipo 131 (ST131) se considera un factor principal en la mayor resistencia a los fármacos, incluyendo la resistencia a múltiples fármacos. Este clon se encuentra en el 12.5% al 30% de todos los aislados clínicos de ExPEC, exhibe mayormente serotipo O25b:H4 y muestra altos niveles de resistencia a las fluoroquinolonas, que a menudo se acompañan de resistencia a trimetoprima / sulfametoxazol y beta-lactamasas de amplio espectro que confieren resistencia a las cefalosporinas.

15 El O-antígeno comprende el componente inmunodominante del lipopolisacárido de pared celular (LPS) en bacterias Gram-negativas, incluyendo *E. coli*. Actualmente hay más de 180 O-antígenos de *E. coli* serológicamente únicos identificados, con la gran mayoría de los aislados de ExPEC clasificados dentro de menos de 20 serotipos de O-antígeno. Los O-antígenos de *E. coli* de longitud completa normalmente se componen de aproximadamente 10 a 25 unidades repetitivas de azúcar unidas a la estructura central de LPS altamente conservada, con cada componente sintetizado por separado por enzimas codificadas predominantemente en los grupos de genes *rfb* y *rfa*, respectivamente. Después de la polimerización del O-antígeno, la cadena principal de polisacáridos del O-antígeno puede ser modificada, típicamente a través de la adición de residuos acetilo o glucosa. Estas modificaciones aumentan eficazmente la diversidad serotípica mediante la creación de serotipos antigénicamente distintos que comparten un esqueleto de polisacáridos común, pero difieren en las ramas laterales. Los genes que codifican las enzimas modificadoras de O-antígeno típicamente residen fuera del grupo *rfb* en el cromosoma, y en algunos casos, estos genes se encuentran dentro de bacteriófagos lisogénicos.

20 Los aislados de ExPEC pertenecientes al serogrupo O4 se han identificado comúnmente en estudios de vigilancia contemporáneos de aislados sanguíneos de EE. UU. y la UE. Se determinó la estructura del polisacárido O4 como α -L-Rha (1 \rightarrow 2) α -L-Rha (1 \rightarrow 6) α -D-Glc (1 \rightarrow 3) α -L-FucNAc (1 \rightarrow 3) β -D-GlcNAc (1 \rightarrow a partir de una cepa de *E. coli* O4:K52 (Jann *et al.*, *Carbohydr. Res.* (1993) v. 248, pp. 241–250). Se determinó una forma singular de la estructura de polisacárido O4 para las cepas O4:K3, O4:K6 y O4: K12, en las que la estructura anterior se modificó mediante la adición de un α -D-Glc (1 \rightarrow 3) unido al residuo de ramosa del polisacárido (Jann *et al.*, 1993, *supra*). Esta forma del polisacárido se denomina en el presente documento a continuación como 'O4 glucosilada'. Las enzimas responsables de la modificación del O-antígeno dentro de las cepas de *E. coli* O4 no fueron identificadas.

25 Los esfuerzos hacia el desarrollo de una vacuna para prevenir las infecciones por ExPEC se han centrado en conjugados de polisacáridos O-antígenos. Una vacuna de conjugado de O-antígeno 12-valente se sintetizó a través de la extracción y la purificación de polisacárido de O-antígeno y la conjugación química a exotoxina A de *Pseudomonas aeruginosa* desintoxicada y se evaluó en cuanto a seguridad e inmunogenicidad en un estudio clínico de Fase 1 (Cross *et al.*, *J. Infect. Dis.* (1994) v.170, pp. 834–40). Esta vacuna candidata nunca fue autorizada para su uso clínico. Recientemente se ha desarrollado un sistema de bioconjugación en *E. coli*, en el que el antígeno de polisacárido y la proteína portadora son ambos sintetizados *in vivo* y posteriormente conjugados *in vivo* a través de las actividades de la oligosacaril transferasa PglB, una enzima de *Campylobacter jejuni*, expresada en *E. coli* (Wacker *et al.*, *Proc. Nat. Acad. Sci.* (2006) v. 103, pp. 7088–93). Este sistema de glicosilación de proteínas N-ligadas es capaz de transferir diversos polisacáridos a una proteína portadora, lo que permite métodos directos para purificar el conjugado.

La bioconjugación se ha utilizado con éxito para producir polisacáridos conjugados para una vacuna candidata de O-antígeno de *E. coli* cuatro-valente (Poolman y Wacker, *J. Infect. Dis.* (2016) v.213(1), pp. 6–13). Sin embargo, el desarrollo de una vacuna exitosa de ExPEC requiere cobertura de los serotipos predominantes, y la presencia de otras modificaciones de O-antígeno en subconjuntos de aislados de ExPEC presenta un desafío adicional en la cobertura de aislados que muestran LPS no modificados y modificados. Además, la eficiencia de la producción de los múltiples componentes para composiciones de vacunas más complejas que cubren múltiples serotipos se vuelve cada vez más importante, y por lo tanto sigue habiendo una necesidad de mejoras en la producción de bioconjugados individuales de O-antígenos específicos.

Breve síntesis de la invención

En vista de la creciente resistencia a los antibióticos entre aislados de ExPEC y la presencia de otras modificaciones de O-antígenos entre O-serotipos predominantes, existe una necesidad de mejores tratamientos profilácticos y terapéuticos para estas infecciones. La invención satisface esta necesidad definiendo la composición genética de los aislados clínicos contemporáneos, incluyendo la identificación de los genes que codifican enzimas modificadoras de O-antígenos, lo que permite la ingeniería de células huésped recombinantes capaces de sintetizar bioconjugados de los O-antígenos, incluyendo bioconjugados que comprenden modificaciones de O-antígenos seleccionadas. Además, en un aspecto de la invención, se proporcionan células huésped y métodos para mejorar la producción de bioconjugados de O-antígenos específicos mediante el uso de variantes de oligosacariltransferasa (OST), sobre la base de las ventajas de uso de ciertas variantes de OST para bioconjugados de determinados O-antígenos de *E. coli* de manera dependiente del serotipo impredecible. El uso de tales variantes de OST también pueden, en ciertos casos, afectar el patrón de glicosilación del bioconjugado, por ejemplo aumentando el número relativo de glicanos acoplados a la proteína portadora, en comparación con bioconjugados producidos utilizando variantes de tipo salvaje u otras variantes de la OST, y por lo tanto también se proporcionan nuevos bioconjugados producidos por tales métodos como un aspecto de la invención.

En un aspecto, se proporciona un método de preparación de un bioconjugado de un O_x antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora, comprendiendo el método:

(i) proporcionar una célula huésped recombinante que comprende:

una secuencia de nucleótidos de un grupo de genes *rfb* para el O_x-antígeno polisacárido;

una secuencia de nucleótidos que codifica la proteína portadora que comprende al menos un sitio de glicosilación que comprende una secuencia de consenso de glicosilación que tiene la SEQ ID NO: 1, preferiblemente que tiene la SEQ ID NO: 2; y

una secuencia de nucleótidos que codifica una oligosacaril transferasa PglB_y; y

(ii) cultivar la célula huésped recombinante en condiciones para la producción del bioconjugado,

en donde:

cuando el O_x- antígeno es O1A antígeno polisacárido, el PglB_y comprende las mutaciones de aminoácidos de N311V, K482R, D483H y A669V;

cuando el O_x-antígeno es O4 antígeno polisacárido glucosilado, el PglB_y comprende la mutación de aminoácido N311V o las mutaciones de aminoácidos Y77H y N311V, y la célula huésped recombinante comprende además una secuencia que codifica una glucosiltransferasa GtrS que tiene al menos 80% de identidad con SEQ ID NO: 4 y que es capaz de modificar un O4 antígeno polisacárido de *E. coli* mediante la adición de glucosa para producir el O4 antígeno polisacárido de *E. coli*, y secuencias de nucleótidos que codifican una translocasa GtrA y una glicosiltransferasa GtrB que tienen al menos 80% de identidad de secuencia con SEQ ID NOs: 7 y 8, respectivamente, en donde la translocasa es capaz de translocar glucosa unida a bactoprenol y la glicosiltransferasa es capaz de glucosilar bactoprenol;

cuando el O_x-antígeno es O6A antígeno polisacárido, el PglB_y comprende las mutaciones de aminoácidos de N311V, K482R, D483H y A669V;

cuando el O_x-antígeno es O8 antígeno polisacárido, el PglB_y no comprende mutaciones de aminoácidos en las posiciones 77, 80, 287, 289, 311, 482, 483 y 669;

cuando el O_x-antígeno es O15 antígeno polisacárido, el PglB_y comprende las mutaciones de aminoácidos de N311V, K482R, D483H y A669V;

cuando el O_x-antígeno es O16 antígeno polisacárido, el PglB_y comprende las mutaciones de aminoácidos de Y77H, S80R, Q287P, K289R y N311V;

cuando el Ox-antígeno es O18A antígeno polisacárido, el PglBy no comprende mutaciones de aminoácidos en las posiciones 77, 80, 287, 289, 311, 482, 483 y 669; y

cuando el Ox-antígeno es O75 antígeno polisacárido, el PglBy comprende la mutación de aminoácido de N311V,

5 en donde, en cada caso, las mutaciones de aminoácidos son en relación con la PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6,

en donde los O1A, O4 glucosilado, O6A, O8, O15, O16, O18A y O75 antígenos

polisacáridos tienen las estructuras de Fórmulas (O1A), (O4-Glc+), (O6A), (O8), (O15), (O16), (O18A) y (O75), respectivamente, como se muestra en la Tabla 1, y cada n es independientemente un número entero de 1 a 100, preferiblemente de 3 a 50, por ejemplo de 5 a 40, por ejemplo de 7 a 25, por ejemplo de 10 a 20.

10 En una forma de realización, el Ox-antígeno es O1A antígeno polisacárido, el PglBy comprende las mutaciones de aminoácidos de N311V, K482R, D483H y A669V en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6.

15 En una forma de realización, el Ox-antígeno es O4 antígeno polisacárido glucosilado, y el PglBy comprende la mutación de aminoácido N311V en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6. En una forma de realización, el Ox-antígeno es O4 antígeno polisacárido glucosilado, y el PglBy comprende las mutaciones de aminoácidos Y77H y N311V en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6. En formas de realización en donde el Ox-antígeno es O4 antígeno polisacárido glucosilado, la célula huésped recombinante comprende además, preferiblemente, una secuencia que codifica un GtrS que tiene al menos 80% de identidad con SEQ ID NO: 4, y secuencias de nucleótidos que codifican un GtrA y un GtrB que tienen al menos 80% de identidad de secuencia con SEQ ID NOs: 7 y 8, respectivamente.

20 En una forma de realización, el Ox-antígeno es O6A antígeno polisacárido, el PglBy comprende las secuencias de aminoácidos de N311V, K482R, D483H y A669V en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6.

25 En una forma de realización, el Ox-antígeno es O8 antígeno polisacárido, el PglBy no comprende mutaciones de aminoácidos en las posiciones 77, 80, 287, 289, 311, 482, 483 y 669 en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6.

En una forma de realización, el Ox-antígeno es O15 antígeno polisacárido, el PglBy comprende las mutaciones de aminoácidos de N311V, K482R, D483H y A669V en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6.

30 En una forma de realización, el Ox-antígeno es O16 antígeno polisacárido, el PglBy comprende las mutaciones de aminoácidos de Y77H, S80R, Q287P, K289R y N311V en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6.

35 En una forma de realización, el Ox-antígeno es O18A antígeno polisacárido, el PglBy no comprende mutaciones de aminoácidos en las posiciones 77, 80, 287, 289, 311, 482, 483 y 669 en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6, y preferiblemente comprende la secuencia de aminoácidos de SEQ ID NO: 6.

En una forma de realización, el Ox-antígeno es O75 antígeno polisacárido, el PglBy comprende la mutación de aminoácido de N311V en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácido de SEQ ID NO: 6.

40 En un aspecto particular, se proporciona un método de preparación de un bioconjugado de un Ox-antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora, comprendiendo el método:

(i) proporcionar una célula huésped recombinante que comprende:

una secuencia de nucleótidos de un grupo de genes *rfb* para el Ox-antígeno polisacárido;

45 una secuencia de nucleótidos que codifica la proteína portadora que comprende al menos un sitio de glicosilación que comprende una secuencia de consenso de glicosilación que tiene la SEQ ID NO: 1, preferiblemente que tiene la SEQ ID NO: 2; y

una secuencia de nucleótidos que codifica una oligosacaril transferasa PglBy; y

(ii) cultivar la célula huésped recombinante en condiciones para la producción del bioconjugado,

en donde el PglBy comprende la mutación de aminoácido N311V con relación al PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6,

- en donde el O_x-antígeno es O1A antígeno polisacárido, O4 antígeno polisacárido glucosilado, O6A antígeno polisacárido, O15 antígeno polisacárido, O16 antígeno polisacárido u O75 antígeno polisacárido, y cuando el O_x-antígeno es O4 antígeno polisacárido glucosilado, la célula huésped recombinante comprende además una secuencia que codifica una glucosiltransferasa GtrS que tiene al menos 80% de identidad con SEQ ID NO: 4 y que es capaz de modificar un O4 antígeno polisacárido de *E. coli* mediante la adición de glucosa para producir el O4 antígeno polisacárido glucosilado de *E. coli*, y secuencias de nucleótidos que codifican una translocasa GtrA y una glucosiltransferasa GtrB que tiene al menos 80% de identidad de secuencia con SEQ ID NOs: 7 y 8, respectivamente, en donde la translocasa es capaz de translocar glucosa unida a bactoprenol y la glucosiltransferasa es capaz de glucosilar bactoprenol, y
- en donde los O1A, O4 glucosilado, O6A, O15, O16 y O75 antígenos polisacáridos tienen las estructuras de Fórmulas (O1A), (O4-Glc+), (O6A), (O15), (O16) y (O75), respectivamente, como se muestra en la Tabla 1, y cada n es independientemente un número entero de 1 a 100, preferiblemente de 3 a 50, por ejemplo de 5 a 40, por ejemplo de 7 a 25, por ejemplo de 10 a 20.
- En ciertas formas de realización, el método comprende además aislar el bioconjugado de la célula huésped recombinante.
- En ciertas formas de realización, la proteína portadora se selecciona del grupo que consiste en Exotoxina A desintoxicada de *P. aeruginosa* (EPA), flagelina de *E. coli* (FliC), CRM197, proteína de unión a maltosa (MBP), toxoide diftérico, toxoide tetánico, hemolisina A desintoxicada de *S. aureus*, factor de aglutinación A, factor de aglutinación B, enterotoxina termolábil de *E. coli*, variantes desintoxicadas de enterotoxina termolábil de *E. coli*, toxina del cólera subunidad B (CTB), toxina del cólera, variantes desintoxicadas de la toxina del cólera, proteína Sat de *E. coli*, el dominio pasajero de la proteína Sat de *E. coli*, neumolisina de *Streptococcus pneumoniae*, hemocianina de lapa californiana (KLH), PcrV de *P. aeruginosa*, proteína de membrana externa de *Neisseria meningitidis* (OMPC) y proteína D de *Haemophilus influenzae* no tipificable.
- En ciertas formas de realización, la proteína portadora es exotoxina A desintoxicada de *Pseudomonas aeruginosa* (EPA). Preferiblemente, la proteína portadora EPA comprende 1–10, preferiblemente 2–4, más preferiblemente 4 sitios de glicosilación. En ciertas formas de realización, cada sitio de glicosilación comprende una secuencia de consenso de glicosilación que tiene la SEQ ID NO: 2. En una forma de realización particular, la proteína portadora EPA comprende la SEQ ID NO: 3.
- En ciertas formas de realización, la célula huésped recombinante es una célula de *E. coli*, por ejemplo, una cepa K-12 de *E. coli*, como la cepa W3110.
- En otro aspecto, se proporciona un bioconjugado producido por un método de preparación de un bioconjugado de un O_x antígeno polisacárido unido covalentemente a una proteína portadora tal como se describe aquí.
- En otro aspecto, se proporciona una composición que comprende tal bioconjugado. En algunas formas de realización, una composición comprende al menos 2, preferiblemente al menos 3, más preferiblemente al menos 5, aún más preferiblemente al menos 7 de tales bioconjugados.
- En ciertas formas de realización, una composición de acuerdo con la invención comprende un bioconjugado de O4 antígeno polisacárido glucosilado de *E. coli* unido covalentemente a una proteína portadora, en donde el O4 antígeno polisacárido glucosilado tiene la estructura de Fórmula (O4-Glc+) tal como se muestra en la Tabla 1, y n es un número entero de 1 a 100, preferiblemente de 3 a 50, por ejemplo de 5 a 40, por ejemplo de 7 a 25, por ejemplo de 10 a 20.
- En ciertas formas de realización, una composición de acuerdo con la invención comprende además al menos un bioconjugado de O25B antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora, en donde el O25B antígeno polisacárido tiene la estructura de Fórmula (O25B) tal como se muestra en la Tabla 1, y n es un número entero de 1 a 100, preferiblemente de 3 a 50, por ejemplo de 5 a 40, por ejemplo de 7 a 25, por ejemplo de 10 a 20.
- En ciertas formas de realización, una composición de acuerdo con la invención comprende además al menos un bioconjugado de O2 antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora, en donde el O2 antígeno polisacárido tiene la estructura de Fórmula (O2) tal como se muestra en la Tabla 1, y n es un número entero de 1 a 100, preferiblemente de 3 a 50, por ejemplo de 5 a 40, por ejemplo de 7 a 25, por ejemplo de 10 a 20.
- En ciertas formas de realización, una composición de la invención comprende: (i) bioconjugado de O1A antígeno polisacárido de *E. coli* acoplado covalentemente a una proteína portadora, (ii) bioconjugado de O2 antígeno polisacárido de *E. coli* acoplado covalentemente a una proteína portadora, (iii) bioconjugado de O4 antígeno polisacárido glucosilado de *E. coli* acoplado covalentemente a una proteína portadora, (iv) bioconjugado de O6A antígeno polisacárido de *E. coli* acoplado covalentemente a una proteína portadora, (v) bioconjugado de O8 antígeno polisacárido de *E. coli* acoplado covalentemente a una proteína portadora, (vi) bioconjugado de O15 antígeno polisacárido de *E. coli* acoplado covalentemente a una proteína portadora, (vii) bioconjugado de O16 antígeno polisacárido de *E. coli* acoplado covalentemente a una proteína portadora, (viii) bioconjugado de O25B antígeno polisacárido de *E. coli* acoplado covalentemente a una proteína portadora, y (ix) bioconjugado de O75 antígeno polisacárido de *E. coli* acoplado covalentemente a una proteína portadora, en donde los O1A, O2, O4 glucosilado,

- O6A, O8, O15, O16, O25B y O75 antígenos polisacáridos tienen las estructuras de Fórmulas (O1A), (O2), (O4–Glc+), (O6A), (O8), (O15), (O16), (O25B) y (O75), respectivamente, tal como se muestra en la Tabla 1, y cada n es independientemente un número entero de 1 a 100, preferiblemente de 3 a 50, por ejemplo de 5 a 40, por ejemplo de 7 a 25, por ejemplo de 10 a 20. En ciertas formas de realización, dicha composición comprende además: (x) bioconjugado de O18A antígeno polisacárido de *E. coli* acoplado covalentemente a una proteína portadora, en donde el O18A antígeno polisacárido tiene la estructura de Fórmula (O18A) tal como se muestra en la Tabla 1, y n es un número entero de 1 a 100, preferiblemente de 3 a 50, por ejemplo de 5 a 40, por ejemplo de 7 a 25, por ejemplo de 10 a 20. En ciertas formas de realización, una composición de la invención es una composición inmunogénica.
- En otros aspectos, se proporciona un método de vacunación de un sujeto contra *E. coli* patogénica extraintestinal (ExPEC), que comprende administrar al sujeto dicho bioconjugado o composición tal como se describe en el presente documento. En otros aspectos, se proporciona dicho bioconjugado o composición tal como se describe en el presente documento para su uso en la vacunación contra *E. coli* patogénica extraintestinal (ExPEC).
- En otros aspectos, se proporcionan células huésped recombinantes para la preparación de un bioconjugado de un O_x antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora, donde la célula huésped recombinante comprende:
- (a) una secuencia de nucleótidos de un grupo de genes *rfb* para el O_x–antígeno polisacárido;
 - (b) una secuencia de nucleótidos que codifica la proteína portadora que comprende al menos un sitio de glicosilación que comprende una secuencia de consenso de glicosilación que tiene la SEQ ID NO: 1, preferiblemente que tiene la SEQ ID NO: 2; y
 - (c) una secuencia de nucleótidos que codifica una oligosacaril transferasa PglB_y,
- en donde:
- cuando el O_x–antígeno es O1A antígeno polisacárido, el PglB_y comprende las mutaciones de aminoácidos de N311V, K482R, D483H y A669V;
 - cuando el O_x–antígeno es O4 antígeno polisacárido glucosilado, el PglB_y comprende la mutación de aminoácido N311V o las mutaciones de aminoácidos Y77H y N311V, y la célula huésped recombinante comprende además una secuencia que codifica una glucosiltransferasa GtrS que tienen al menos 80% de identidad con SEQ ID NO: 4 y que es capaz de modificar un O4 antígeno polisacárido de *E. coli* mediante la adición de glucosa para producir el O4 antígeno polisacárido glucosilado de *E. coli*, y secuencias de nucleótidos que codifican una translocasa GtrA y una glicosiltransferasa GtrB que tienen al menos 80% de identidad de secuencia con SEQ ID NOs: 7 y 8, respectivamente, en donde la translocasa es capaz de translocar glucosa unida a bactoprenol y la glicosiltransferasa es capaz de glucosilar bactoprenol;
 - cuando el O_x–antígeno es O6A antígeno polisacárido, el PglB_y comprende las mutaciones de aminoácidos de N311V, K482R, D483H y A669V;
 - cuando el O_x–antígeno es O6A antígeno polisacárido, el PglB_y comprende las mutaciones de aminoácidos de N311V, K482R, D483H y A669V;
 - cuando el O_x–antígeno es O15 antígeno polisacárido, el PglB_y comprende las mutaciones de aminoácidos de N311V, K482R, D483H y A669V;
 - cuando el O_x–antígeno es O16 antígeno polisacárido, el PglB_y comprende las mutaciones de aminoácidos de Y77H, S80R, Q287P, K289R y N311V;
 - cuando el O_x–antígeno es O18A antígeno polisacárido, el PglB_y no comprende mutaciones de aminoácidos en las posiciones 77, 80, 287, 289, 311, 482, 483 y 669; y
 - cuando el O_x–antígeno es O75 antígeno polisacárido, el PglB_y comprende la mutación de aminoácido de N311V,
- en donde, en cada caso, las mutaciones de aminoácidos son en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6, y
- en donde los O1A, O4 glucosilado, O6A, O8, O15, O16, O18A y O75 antígenos polisacáridos tienen las estructuras de Fórmulas (O1A), (O4–Glc+), (O6A), (O8), (O15), (O16), (O18A) y (O75), respectivamente, tal como se muestra en la Tabla 1, y cada n es independientemente un número entero de 1 a 100, preferiblemente de 3 a 50, por ejemplo de 5 a 40, por ejemplo de 7 a 25, por ejemplo de 10 a 20.

En ciertas formas de realización, se proporcionan tales células huésped en donde el O_x-antígeno es O1A antígeno polisacárido, y el PglB_y comprende las mutaciones de aminoácidos de N311V, K482R, D483H y A669V en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6.

5 En ciertas formas de realización, se proporcionan células huésped recombinantes de la invención en donde el O_x-antígeno es O4 antígeno polisacárido glucosilado, y el PglB_y comprende la mutación de aminoácido N311V en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6. En ciertas formas de realización, se proporcionan células huésped recombinantes de la invención en donde el O_x-antígeno es O4 antígeno polisacárido glucosilado, y el PglB_y comprende las mutaciones de aminoácidos Y77H y N311V en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6. En ciertas formas de realización en donde el O_x-antígeno es O4 antígeno polisacárido glucosilado, la célula huésped recombinante comprende además una secuencia que codifica un GtrS que tiene la secuencia de aminoácidos de SEQ ID NO: 4, y secuencias de nucleótidos que codifican un GtrA y un GtrB que tienen las secuencias de aminoácidos de SEQ ID NOs: 7 y 8, respectivamente.

10 En ciertas formas de realización, se proporcionan células huésped recombinantes de la invención en donde el O_x-antígeno es O6A antígeno polisacárido, y el PglB_y comprende las mutaciones de aminoácidos de N311V, K482R, D483H y A669V en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6.

15 En ciertas formas de realización, se proporcionan células huésped recombinantes de la invención en donde el O_x-antígeno es O8 antígeno polisacárido, el PglB_y no comprende mutaciones de aminoácidos en las posiciones 77, 80, 287, 289, 311, 482, 483 y 669 en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6.

20 En ciertas formas de realización, se proporcionan células huésped recombinantes de la invención en donde el O_x-antígeno es O15 antígeno polisacárido, el PglB_y comprende las mutaciones de aminoácidos de N311V, K482R, D483H y A669V en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6.

25 En ciertas formas de realización, se proporcionan células huésped recombinantes de la invención, en donde el O_x-antígeno es O16 antígeno polisacárido, el PglB_y comprende las mutaciones de aminoácidos de Y77H, S80R, Q287P, K289R y N311V en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6.

30 En ciertas formas de realización, se proporcionan células huésped recombinantes de la invención, en donde el O_x-antígeno es O18A antígeno polisacárido, el PglB_y no comprende mutaciones de aminoácidos en las posiciones 77, 80, 287, 289, 311, 482, 483 y 669 en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6.

35 En ciertas formas de realización, se proporcionan células huésped recombinantes de la invención, en donde el O_x-antígeno es O75 antígeno polisacárido, el PglB_y comprende la mutación de aminoácido de N311V en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácido de SEQ ID NO: 6.

40 En ciertas formas de realización, se proporcionan células huésped recombinantes de la invención, en donde la proteína portadora se selecciona del grupo que consiste en Exotoxina A desintoxicada de *P. aeruginosa* (EPA), flagelina de *E. coli* (FliC), CRM197, proteína de unión a maltosa (MBP), toxoide diftérico, toxoide tetánico, hemolisina A desintoxicada de *S. aureus*, factor de aglutinación A, factor de aglutinación B, enterotoxina termolábil de *E. coli*, variantes desintoxicadas de enterotoxina termolábil de *E. coli*, toxina del cólera subunidad B (CTB), toxina del cólera, variantes desintoxicadas de la toxina del cólera, proteína Sat de *E. coli*, el dominio pasajero de la proteína Sat de *E. coli*, neumolisina de *Streptococcus pneumoniae*, hemocianina de lapa californiana (KLH), PcrV de *P. aeruginosa*, proteína de membrana externa de *Neisseria meningitidis* (OMPC) y proteína D de *Haemophilus influenzae* no tipificable.

45 En ciertas formas de realización, se proporcionan células huésped recombinantes de la invención, en donde la proteína portadora es exotoxina A desintoxicada de *Pseudomonas aeruginosa* (EPA). En ciertas formas de realización de la misma, la proteína portadora EPA comprende 1–10, preferiblemente 2–4, más preferiblemente 4, de los sitios de glicosilación. En ciertas formas de realización, cada sitio de glicosilación comprende una secuencia de consenso de glicosilación que tiene la SEQ ID NO: 2. En ciertas formas de realización, la proteína portadora EPA comprende la SEQ ID NO: 3.

En ciertas formas de realización, se proporcionan células huésped recombinantes de la invención, en donde la célula huésped recombinante es una célula de *E. coli*, por ejemplo, una cepa K-12 de *E. coli*, tal como la cepa W3110.

50 En ciertas formas de realización para las células huésped y métodos para la preparación de un bioconjugado de un O4 antígeno polisacárido glucosilado de *E. coli* unido covalentemente a una proteína portadora de acuerdo con la invención, el grupo de genes *rfb* para el O4 antígeno polisacárido de *E. coli* comprende una secuencia que codifica las enzimas que crean el O4 antígeno polisacárido de *E. coli* (Fórmula (O4-Glc-) en la Tabla 1) y es al menos 80%, por ejemplo al menos 90%, por ejemplo al menos 95%, por ejemplo al menos 98% idéntica a la SEQ ID NO: 9. En ciertas formas de realización, el grupo de genes *rfb* comprende la SEQ ID NO: 9.

En ciertas formas de realización para las células huésped y métodos para la preparación de un bioconjugado de un O4 antígeno polisacárido glucosilado de *E. coli* unido covalentemente a una proteína portadora de acuerdo con la invención, la glucosil transferasa que es capaz de modificar el O4 antígeno polisacárido de *E. coli* para producir el O4 antígeno polisacárido glucosilado de *E. coli* tiene una secuencia de aminoácidos que tiene al menos 90%, preferiblemente al menos 95%, preferiblemente al menos 98% de identidad de secuencia con la SEQ ID NO: 4. En ciertas formas de realización, la glucosil transferasa comprende la SEQ ID NO: 4.

En ciertas formas de realización para las células huésped y los métodos para la preparación de un bioconjugado de un O4 antígeno polisacárido glucosilado de *E. coli* unido covalentemente a una proteína portadora de acuerdo con la invención, la translocasa es capaz de translocar glucosa unida a bactoprenol y tiene al menos 90%, preferiblemente al menos 95%, preferiblemente al menos 98% de identidad de secuencia con la SEQ ID NO: 7. En ciertas formas de realización, la translocasa comprende la SEQ ID NO: 7.

En ciertas formas de realización para las células huésped y métodos para la preparación de un bioconjugado de un O4 antígeno polisacárido glucosilado de *E. coli* unido covalentemente a una proteína portadora de acuerdo con la invención, la glicosiltransferasa es capaz de glucosilar bactoprenol y tiene al menos 90%, preferiblemente al menos 95%, preferiblemente al menos 98% de identidad de secuencia con la SEQ ID NO: 8. En ciertas formas de realización, la glicosiltransferasa comprende la SEQ ID NO: 8.

Breve descripción de las figuras

La síntesis anterior, así como la siguiente descripción detallada de la invención, se comprenderá mejor cuando se lea conjuntamente con los dibujos adjuntos. Se debe entender que la invención no se limita a las formas de realización precisas mostradas en los dibujos.

En los dibujos:

La **FIG. 1** muestra títulos de IgG por ELISA contra O4 LPS no modificado (GLC-) o modificado con glucosa (GLC+) en el suero de dos conejos inmunizados con bioconjugado de O4 polisacárido modificado con Glc como se describe en el Ejemplo 4; los títulos de ELISA se determinaron por cuadruplicado;

La **FIG. 2** muestra títulos de IgG en ELISA de células enteras con mezcla de los sueros de conejos inmunizados con un bioconjugado O4 modificado con Glc contra aislados O4 de *E. coli* con estado *gtrS* caracterizado como se describe en el Ejemplo 4; los siguientes aislados fueron *gtrS*-negativos: A2625, stGVXN4988, OC24784, OC24787 y OC24788; los siguientes aislados fueron *gtrS*-positivos: Y1382, E551, OC24334, stGVXN4983, stGVXN4994 y OC24794; también se incluyó la cepa de control negativo OC9487 (ATCC 35383; serotipo O75);

La **FIG. 3** muestra transferencias Western de LPS extraído de aislados O4 *gtrS*-positivos y negativos sondeados con mezcla de los sueros de conejos inmunizados con O4 polisacárido modificado;

Las **FIGS. 4A y 4B** muestran respuestas de anticuerpos inducidas por bioconjugados O4 glucosilados (O4-Glc+) EPA; la **FIG. 4A** muestra los niveles séricos de anticuerpos medidos por ELISA el Día 0, 14 y 42 después de la inmunización; se muestran los títulos individuales (\log_{10} EC50 título) y GMT \pm IC del 95%; la línea de puntos grises indica el umbral por encima del cual las curvas de dilución de las muestras tienen un ajuste de 4PL; la **FIG. 4B** muestra los resultados del ensayo de opsonofagocítico (OPK) para determinar la funcionalidad de los anticuerpos en muestras de suero obtenidas el Día 42 después de la inmunización con bioconjugado O4 glucosilado (O4-Glc+)-EPA (4.0 μ g); prueba de suma de rangos de Wilcoxon y corrección de Bonferroni; * $P \leq 0.05$, *** $P \leq 0.0001$;

La **FIG. 5** muestra el efecto de potenciación del bioconjugado O4 glucosilado (O4 Glc+)-EPA en ratas Sprague Dawley inmunizadas con 3 dosis diferentes como se describe en el Ejemplo 4; los niveles séricos de anticuerpos se midieron por ELISA el Día 0, 14 y 42 después de la inmunización; se muestran los títulos individuales (\log_{10} EC50 título) para cada animal; las líneas entre los puntos de datos conectan los títulos de IgG para cada animal en el tiempo; la línea de puntos grises indica el umbral por encima del cual las curvas de dilución de las muestras tienen un ajuste de 4PL; el análisis estadístico se realizó con una prueba de rango firmado de Wilcoxon y corrección de Bonferroni para comparaciones múltiples (día 14 versus día 0, $P = 0.012$ para 4.0 μ g/dosis; día 42 versus día 0, $P = 0.006$ para todas las dosis; día 42 versus día 14, $P = 0.006$ para todas las dosis);

La **FIG. 6** muestra la funcionalidad de los anticuerpos inducidos por el bioconjugado O4-Glc+-EPA; las ratas Sprague Dawley fueron inmunizadas por vía intramuscular 3 veces con tampón de formulación o bioconjugado O4(Glc+)-EPA a razón de 4.00 μ g/dosis; la funcionalidad de los anticuerpos se determinó mediante el ensayo de matanza opsonofagocítica (OPKA) usando cepas de *E. coli* O4(Glc+) y O4(Glc-); se muestran los títulos opsónicos individuales (OI) y GMT \pm IC del 95%;

La **FIG. 7** muestra la lectura de electroforesis capilar de pantalla de PgIB que visualiza la producción de bioconjugado O4-Glc+ para cada cepa ensayada en una imagen tipo transferencia, utilizando anticuerpos monoclonales para detectar bioconjugado O4-Glc+ en la fracción periplásmica. Producto monoglicosilado de aproximadamente 180 kDa,

producto diglicosilado de aproximadamente 320 kDa y producto triglicosilado de aproximadamente 450 kDa. A) Primera ronda de selección. Wt PglB en línea 3, N311V–PglB en líneas 2 y 4, cepa de control vacía en línea 1 y otras variantes de PglB en líneas 5 y 6. B) Segunda ronda de selección. N311V PglB en línea 3, N311V+Y77H PglB en línea 9, cepa de control vacía en líneas 1 y 2, otras variantes de PglB en las líneas restantes.

5 La **FIG. 8** muestra respuestas de anticuerpos inducidos por la vacuna ExPEC10V en conejos blancos de Nueva Zelanda. Los animales recibieron 3 inmunizaciones intramusculares con ExPEC10V o solución salina administrada con 2 semanas de diferencia. La vacuna ExPEC10V se administró en 3 concentraciones diferentes (grupo 1: dosis alta, grupo 2: dosis media, y grupo 3: dosis baja, Tabla 11) y un grupo de control recibió solo solución salina (grupo 4, 0.9% (p/v) de solución de cloruro de sodio). Los niveles de anticuerpos se midieron por ELISA el día 0 (antes de la vacunación) y los días 14, 27 y 42 (después de la vacunación). Se muestran los títulos individuales (título EC50) y los títulos de media geométrica (GMT) \pm IC del 95%. Prueba de suma de rangos de Wilcoxon con corrección de Bonferroni para comparaciones múltiples. Comparaciones de animales vacunados con ExPEC10V (grupos 1, 2 y 3) versus control de solución salina (grupo 4). * $p \leq 0.05$, ** $p \leq 0.01$; *** $p \leq 0.001$; **** $p \leq 0.0001$. LOD: límite de detección.

10 La **FIG. 9** muestra respuestas de anticuerpos inducidas por ExPEC10V. Los conejos blancos de Nueva Zelanda recibieron 3 inmunizaciones intramusculares con ExPEC10V (105.6 μ g de polisacárido total) o 0.9% p/v solución de cloruro de sodio (control). Los títulos de IgG se determinaron por ELISA el día 1 (antes de la inmunización, $n = 20$ /grupo), el día 31 (después de la inmunización, $n = 20$ /grupo) y el día 50 (después de la inmunización, $n = 10$ /grupo). Los gráficos muestran los títulos individuales y la media geométrica \pm intervalo de confianza del 95% para cada grupo. Las diferencias en los títulos de IgG entre el grupo de ExPEC10V y el grupo de control se analizaron usando un modelo Tobit con una prueba de razón de probabilidad. Los valores de $p \leq 0.05$ se consideraron significativos. * $P \leq 0.05$, **** $P \leq 0.0001$.

15 La **FIG. 10** muestra el diseño general de estudio para un ensayo clínico de fase 1/2a con la vacuna ExPEC10V en seres humanos. La **FIG. 10A** muestra el diseño general de estudio para la Cohorte 1, y la **FIG. 10B** muestra el diseño general de estudio para la Cohorte 2. Véase el Ejemplo 11 para obtener los detalles.

25 Descripción detallada de la invención

Diversas publicaciones, artículos y patentes se citan o describen en los antecedentes y en toda la memoria descriptiva. La discusión de los documentos, actos, materiales, dispositivos, artículos o similares que se incluyen en la presente memoria descriptiva es a los fines de proporcionar contexto a la invención. Tal discusión no es una admisión de que cualquiera o todos estos asuntos forman parte del arte previo con respecto a cualquier invención divulgada o reivindicada.

30 A menos que se defina lo contrario, todos los términos técnicos y científicos usados en este documento tienen el mismo significado que entiende comúnmente un experto en la técnica a la que pertenece esta invención. De lo contrario, ciertos términos usados en este documento tienen los significados que se exponen en la memoria descriptiva.

35 Debe observarse que, tal como se usan aquí y en las reivindicaciones adjuntas, las formas singulares “un”, “una”, “el” y “la” incluyen la referencia plural, a menos que el contexto claramente dicte otra cosa.

A menos que se indique lo contrario, el término “al menos” que precede una serie de elementos ha de entenderse en referencia a todos los elementos de la serie.

40 Los expertos en la técnica reconocerán o serán capaces de determinar, sin más que la experimentación de rutina, muchos equivalentes a las formas de realización específicas de la invención aquí descrita. Tales equivalentes están comprendidos por la invención.

45 A lo largo de esta memoria descriptiva y las reivindicaciones que siguen, a menos que el contexto requiera otra cosa, se entenderá que la palabra “comprenden”, y variaciones tales como “comprende” y “que comprende(n)”, implican la inclusión de un número entero o etapa o grupo de números enteros o etapas indicados, pero no la exclusión de cualquier otro número entero o etapa o grupo de números enteros o etapas. Cuando se utiliza aquí, el término “que comprende(n)” puede estar sustituido por el término “que contiene(n)” o “que incluye(n)”, o a veces cuando se usa en el presente documento por el término “que tiene(n)”.

50 Cuando se usa en el presente documento, “que consiste(n) en” excluye cualquier elemento, etapa o ingrediente no especificado en el elemento de la reivindicación. Cuando se usa en el presente documento, “que consiste(n) esencialmente en” no excluye materiales o etapas que no afectan materialmente las características básicas y nuevas de la reivindicación. Cualquiera de los términos antes mencionados de “que comprende(n)”, “que contiene(n)”, “que incluye(n)” y “que tiene(n)”, cuando se usan aquí en el contexto de un aspecto o forma de realización de la invención, se puede sustituir por el término “que consiste(n) en” o “que consiste(n) esencialmente en” para variar los alcances de la divulgación.

Tal como se utiliza aquí, el término conjuntivo “y/o” entre múltiples elementos mencionados se entiende como que abarca ambas opciones individuales y combinadas. Por ejemplo, cuando dos elementos están unidos por “y/o”, una primera opción se refiere a la aplicabilidad del primer elemento sin el segundo. Una segunda opción se refiere a la aplicabilidad del segundo elemento sin el primero. Una tercera opción se refiere a la aplicabilidad de los elementos primero y segundo juntos. Cualquiera de estas opciones se entiende que caen dentro del significado y, por lo tanto, satisfacen el requisito del término “y/o”, como se usa en el presente documento. La aplicabilidad simultánea de más de una de las opciones también se entiende que cae dentro del significado y, por lo tanto, satisface el requisito del término “y/o”.

La identificación de una modificación estructural de O–antígeno, es decir, la ramificación de glucosa, dentro del serotipo de O4 de *E. coli* (Jann *et al.*, 1993) presenta un desafío para el descubrimiento y el desarrollo de una vacuna de glicoconjugado que apunta a aislados bacterianos dentro de este serotipo. Se desconoce la proporción de aislados de O4 contemporáneos clínicos que expresan las formas no modificadas (que no tienen una rama lateral de glucosa) y modificadas (que tienen una rama lateral de glucosa) del O4 O–antígeno. La obtención de información sobre esta característica es fundamental para seleccionar la estructura antigénica relevante. Además, el grado en el que los anticuerpos inducidos por la vacuna generados contra una forma del O4 polisacárido reaccionarán en forma cruzada con otra forma no se ha determinado. La purificación del O–antígeno libre de lípido A y la posterior conjugación química a una proteína portadora es un proceso largo y laborioso. Además, los procesos de purificación, desintoxicación de lípido A y conjugación química pueden causar pérdida de epítomos, heterogeneidad de antígeno e inmunogenicidad reducida del polisacárido conjugado. La síntesis de glicoconjugados por bioconjugación puede superar estas limitaciones de la purificación y la conjugación química clásicas, pero la síntesis *in vivo* de O4 O–antígeno con ramificación de glucosa requiere la actividad de una enzima de ramificación de polisacárido, que se encuentra fuera del grupo de genes *rfb*. Hasta la fecha, la enzima modificadora de O–antígeno responsable de la ramificación de glucosa en cepas de O4 de *E. coli* no ha sido identificada. La clonación del grupo de genes O4 *rfb* en la bioconjugación de cepas de *E. coli* que expresan PglB no será suficiente para sintetizar el glicoconjugado O4 con ramificación de glucosa, sino más bien solamente produciría bioconjugados O4 sin ramificación de glucosa (la estructura del glicano de la misma se muestra en la Fórmula (O4) de la Tabla 1). Tal como se usan en el presente documento, los términos O–antígeno “O4 glucosilado”, “O4 con ramificación de glucosa”, “O4 Glc+” y “Glc+ O4” se refieren a O4 O–antígeno con una rama lateral de glucosa, y la estructura del mismo se muestra en la fórmula (O4–Glc+) de la Tabla 1.

Se divulga en el presente documento el gen que codifica la enzima modificadora de O–antígeno responsable de la ramificación de glucosa del O4 antígeno polisacárido de *E. coli*. También se divulgan en el presente documento células huésped, por ejemplo, células huésped modificadas recombinantemente por ingeniería genética que comprenden enzimas codificadoras de ácido nucleico capaces de producir bioconjugados que comprenden el O4 antígeno polisacárido glucosilado unido covalentemente a una proteína portadora *in vivo*. Tales células huésped se pueden utilizar para generar bioconjugados que comprenden el O4 antígeno glucosilado unido a una proteína portadora, que puede ser utilizado en, por ejemplo, la formulación de composiciones terapéuticas y/o profilácticas (por ejemplo, vacunas). Además se proporcionan en el presente documento composiciones que comprenden bioconjugados del O4 antígeno polisacárido glucosilado, solo o en combinación con otros antígenos de *E. coli* (por ejemplo, O1, O2, O6, O8, O15, O16, O18, O25 y/o O75 antígenos polisacáridos y subserotipos de los mismos). Las composiciones se pueden usar en tratamiento profiláctico y/o métodos terapéuticos, por ejemplo, en la vacunación de huéspedes contra la infección por *E. coli*, y son útiles en la generación de anticuerpos, que se pueden utilizar, por ejemplo, en métodos terapéuticos tales como la inmunización de sujetos.

Tal como se utilizan aquí, los términos “O–antígeno”, “O–antígeno polisacárido”, “O–antígeno sacárido” y “OPS” se refieren al O–antígeno de bacterias Gram–negativas. Típicamente, un O–antígeno es un polímero de unidades de polisacárido repetidoras inmunogénicas. En una forma de realización particular, los términos “O–antígeno”, “O–antígeno polisacárido” y “OPS” se refieren al O–antígeno de *Escherichia coli*. Diferentes serotipos de *E. coli* expresan diferentes O–antígenos. En *E. coli*, los productos génicos implicados en la biogénesis de O–antígenos son codificados por el grupo de genes *rfb*.

Tal como se usan en el presente documento, “grupo de *rfb*” y “grupo de genes *rfb*” se refieren a un grupo de genes que codifica maquinaria enzimática capaz de sintetizar una estructura de cadena principal de O–antígeno. El término grupo *rfb* puede aplicarse a cualquier grupo biosintético de O–antígenos, y preferiblemente se refiere a un grupo de genes del género *Escherichia*, en particular *E. coli*.

Tal como se utiliza aquí, el término “O1A” se refiere al antígeno O1A de *E. coli* (un subserotipo de *E. coli* serotipo O1). El término “O2” se refiere al antígeno O2 de *E. coli* (*E. coli* serotipo O2). El término “O6A” se refiere al antígeno O6A de *E. coli* (un subserotipo de *E. coli* serotipo O6). El término “O8” se refiere al antígeno O8 de *E. coli* (*E. coli* serotipo O8). El término “O15” se refiere al antígeno O15 de *E. coli* (*E. coli* serotipo O15). El término “O16” se refiere al antígeno O16 de *E. coli* (*E. coli* serotipo O16). El término “O18A” se refiere al antígeno O18A de *E. coli* (un subserotipo de *E. coli* serotipo O18). El término “O25B” se refiere al antígeno O25B de *E. coli* (un subserotipo de *E. coli* serotipo O25). El término “O75” se refiere al antígeno O75 de *E. coli* (*E. coli* serotipo O75).

Las estructuras de O-antígenos polisacáridos de *E. coli* mencionadas en toda esta solicitud se muestran a continuación en la Tabla 1. Se muestra una única unidad repetidora para cada O-antígeno polisacárido de *E. coli*.

Tabla 1: Estructuras de O-antígenos polisacáridos de *E. coli*

O-antígeno polisacárido de <i>E. coli</i>	Estructura de unidad repetidora ¹
O4 antígeno polisacárido no glucosilado (O4-Glc-)	$[\rightarrow 2)\text{-}\alpha\text{-L-Rhap}\text{-}(1\rightarrow 6)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow 3)\text{-}\alpha\text{-L-FucpNac}\text{-}(1\rightarrow 3)\text{-}\beta\text{-D-GlcpNac}\text{-}(1\rightarrow)]_n$
O4 antígeno polisacárido glucosilado (O4-Glc+)	$\begin{array}{c} \alpha\text{-D-Glcp} \\ 1 \\ \downarrow \\ 3 \\ [\rightarrow 2)\text{-}\alpha\text{-L-Rhap}\text{-}(1\rightarrow 6)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow 3)\text{-}\alpha\text{-L-FucpNac}\text{-}(1\rightarrow 3)\text{-}\beta\text{-D-GlcpNac}\text{-}(1\rightarrow)]_n \end{array}$
O1A antígeno polisacárido (O1A)	$\begin{array}{c} 2 \\ \uparrow \\ 1 \\ \beta\text{-D-ManpNac} \\ [\rightarrow 3)\text{-}\alpha\text{-L-Rhap}\text{-}(1\rightarrow 3)\text{-}\alpha\text{-L-Rhap}\text{-}(1\rightarrow 3)\text{-}\beta\text{-L-Rhap}\text{-}(1\rightarrow 4)\text{-}\beta\text{-D-GlcpNac}\text{-}(1\rightarrow)]_n \end{array}$
O2 antígeno polisacárido (O2)	$\begin{array}{c} 2 \\ \uparrow \\ 1 \\ \alpha\text{-D-Fucp3Nac} \\ [\rightarrow 3)\text{-}\alpha\text{-L-Rhap}\text{-}(1\rightarrow 2)\text{-}\alpha\text{-L-Rhap}\text{-}(1\rightarrow 3)\text{-}\beta\text{-L-Rhap}\text{-}(1\rightarrow 4)\text{-}\beta\text{-D-GlcpNac}\text{-}(1\rightarrow)]_n \end{array}$
O6A antígeno polisacárido (O6)	$\begin{array}{c} 2 \\ \uparrow \\ 1 \\ \beta\text{-D-Glcp} \\ [\rightarrow 4)\text{-}\alpha\text{-D-GalpNac}\text{-}(1\rightarrow 3)\text{-}\beta\text{-D-Manp}\text{-}(1\rightarrow 4)\text{-}\beta\text{-D-Manp}\text{-}(1\rightarrow 3)\text{-}\alpha\text{-D-GlcpNac}\text{-}(1\rightarrow)]_n \end{array}$
O8 antígeno polisacárido (O8)	$\alpha\text{-D-Manp3Me}\text{-}(1\rightarrow [3])\text{-}\beta\text{-D-Manp}\text{-}(1\rightarrow 2)\text{-}\alpha\text{-D-Manp}\text{-}(1\rightarrow 2)\text{-}\alpha\text{-D-Manp}\text{-}(1\rightarrow)]_n$
O15 antígeno polisacárido (O15)	$[\rightarrow 2)\text{-}\beta\text{-D-Galp}\text{-}(1\rightarrow 3)\text{-}\alpha\text{-L-FucpNac}\text{-}(1\rightarrow 3)\text{-}\beta\text{-D-GlcpNac}\text{-}(1\rightarrow)]_n$

O16 antígeno polisacárido (O16)	$[\rightarrow 2)\text{-}\beta\text{-D-Galp}\text{-}(1\rightarrow 6)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow 3)\text{-}\alpha\text{-L-Rhap}\text{-}(1\rightarrow 3)\text{-}\alpha\text{-D-GlcpNAc}\text{-}(1\rightarrow)]_n$ $\begin{matrix} 2 \\ \uparrow \\ \text{Ac} \end{matrix}$
O18A antígeno polisacárido (O18A)	$[\rightarrow 2)\text{-}\alpha\text{-L-Rhap}\text{-}(1\rightarrow 6)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow 4)\text{-}\alpha\text{-D-Galp}\text{-}(1\rightarrow 3)\text{-}\alpha\text{-D-GlcpNAc}\text{-}(1\rightarrow)]_n$ $\begin{matrix} 3 \\ \uparrow \\ 1 \\ \beta\text{-D-GlcpNAc} \end{matrix}$
O25B antígeno polisacárido (O25B)	$\beta\text{-D-Glcp}$ $\begin{matrix} 1 \\ \downarrow \\ 6 \end{matrix}$ $[\rightarrow 4)\text{-}\alpha\text{-D-Glcp}\text{-}(1\rightarrow 3)\text{-}\alpha\text{-L-Rhap}\text{-}(1\rightarrow 3)\text{-}\beta\text{-D-GlcpNAc}\text{-}(1\rightarrow)]_n$ $\begin{matrix} 3 \\ \uparrow \\ 1 \\ \alpha\text{-L-Rhap} \end{matrix}$ $\begin{matrix} 2 \\ \uparrow \\ \text{Ac} \end{matrix}$
O75 antígeno polisacárido (O75)	$\beta\text{-D-Manp}$ $\begin{matrix} 1 \\ \downarrow \\ 4 \end{matrix}$ $[\rightarrow 3)\text{-}\alpha\text{-D-Galp}\text{-}(1\rightarrow 4)\text{-}\alpha\text{-L-Rhap}\text{-}(1\rightarrow 3)\text{-}\beta\text{-D-GlcpNAc}\text{-}(1\rightarrow)]_n$
<p>[†]Cada n es independientemente un número entero de 1 a 100, tal como 1-50, 1-40, 1-30, 1-20 y 1-10, 3-50, 3-40, por ejemplo al menos 5, tal como 5-40, por ejemplo 7-30, por ejemplo 7 a 25, por ejemplo 10 a 20, pero en algunos casos puede ser 1-2.</p>	

Todos los monosacáridos descritos en este documento tienen su significado común conocido en la técnica. Los monosacáridos pueden tener la configuración D o L. Si no se especifica D o L, se entiende que el azúcar tiene la configuración D. Los monosacáridos se denominan típicamente por abreviaturas comúnmente conocidas y usadas en la técnica. Por ejemplo, Glc se refiere a glucosa; D-Glc se refiere a D-glucosa; y L-Glc se refiere a L-glucosa. Otras abreviaturas comunes para monosacáridos incluyen: Rha, ramnosa; GlcNAc, N-acetilglucosamina; GalNAc, N-acetilgalactosamina; Fuc, fucosa; Man, manosa; Man3Me, 3-O-metil-manosa; Gal, galactosa; FucNAc, N-acetilfucosamina; y Rib, ribosa. El sufijo "f" se refiere a furanosa, y el sufijo "p" se refiere a piranosa.

Los términos "RU", "unidad de repetición" y "unidad repetidora", tal como se utilizan con respecto a un O-antígeno, se refieren a la unidad de repetición biológica (BRU) de un O-antígeno como se sintetiza *in vivo* por la maquinaria celular (por ejemplo, glicosiltransferasas). El número de RU de un O-antígeno puede variar según el serotipo, y en algunas formas de realización de la invención varía típicamente de aproximadamente 1-100 RU, preferiblemente de aproximadamente 1 a 50 RU, tal como 1-50 RU, 1-40 RU, 1-30 RU, 1-20 RU y 1-10 RU, y más preferiblemente al menos 3 RU, al menos 4 RU, al menos 5 RU, tal como 3-50 RU, preferiblemente 5-40 RU, por ejemplo 7-25 RU, por ejemplo 10-20 RU. Sin embargo, en algunos casos, el número de RU de un O-antígeno puede ser 1-2. La estructura de cada O-antígeno que se describe específicamente en este documento se muestra como que contiene una RU con la variable "n" que designa el número de RU. En cada O-antígeno polisacárido en un bioconjugado de la invención, n es independientemente un número entero de 1-100, tal como 1-50, 1-40, 1-30, 1-20, 1-10, preferiblemente al menos 3, más preferiblemente al menos 5, tal como 3-50, preferiblemente 5-40 (por ejemplo, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39 o 40), pero en algunos casos puede ser 1-2. En algunas formas de realización, n es independientemente un número entero de aproximadamente 7-25, por ejemplo aproximadamente 10-20. Los valores pueden variar entre O-antígenos polisacáridos individuales en una composición, y se proporcionan aquí como valores medios, es decir, si un bioconjugado se describe aquí como que tiene un n que es independientemente un número entero de 5-40, la composición contiene una mayoría de O-antígenos polisacáridos con 5-40 unidades de repetición, pero también puede contener algunos O-antígenos polisacáridos que tienen menos de 5 unidades de repetición o más de 40 unidades de repetición.

El término "glicoconjugado" se refiere a un conjugado de antígeno de azúcar o sacárido (por ejemplo, oligo- y polisacárido)-proteína unido a otra especie química, incluyendo, pero sin limitaciones, proteínas, péptidos, lípidos, etc. Los glicoconjugados se pueden preparar químicamente, por ejemplo, por enlace químico (sintético) de la proteína y el antígeno de azúcar o sacárido. El término "glicoconjugado" también incluye bioconjugados.

El término "bioconjugado" se refiere a un conjugado entre una proteína (por ejemplo, una proteína portadora) y un antígeno de azúcar o sacárido (por ejemplo, oligo- y poli-sacárido) preparado en un fondo de célula huésped, preferiblemente una célula huésped bacteriana, por ejemplo una célula huésped de *E. coli*, en donde la maquinaria de la célula huésped une el antígeno a la proteína (por ejemplo, con un N-enlace). Preferiblemente, el término "bioconjugado" se refiere a un conjugado entre una proteína (por ejemplo, una proteína portadora) y un O-antígeno, preferiblemente un O-antígeno de *E. coli* (por ejemplo, O1A, O2, O4 glucosilado, O6A, O8, O15, O16, O18A, O25B, O75, etc.) preparado en un fondo de célula huésped, en donde la maquinaria de la célula huésped une el antígeno a la proteína (por ejemplo, con un N-enlace). Debido a que los bioconjugados son preparados en células huésped por la maquinaria de la célula huésped, el antígeno y la proteína se unen covalentemente a través de un enlace glicosídico o se unen en un bioconjugado. Los bioconjugados pueden prepararse en células huésped recombinantes modificadas por ingeniería genética para expresar la maquinaria celular necesaria para sintetizar el O-antígeno y/o enlazar el O-antígeno a la proteína diana. Los bioconjugados, como se describe en este documento, tienen propiedades ventajosas respecto de los glicoconjugados preparados químicamente, donde los glicanos se purifican a partir de las paredes celulares bacterianas y posteriormente se acoplan químicamente a una proteína portadora, por ejemplo, los bioconjugados requieren menos productos químicos en la fabricación y son más consistente en términos del producto final generado, y contienen menos o ningún glicano (es decir, no se unen a una proteína portadora). Así, en formas de realización típicas, los bioconjugados se prefieren por sobre los glicoconjugados producidos químicamente.

El término "aproximadamente", cuando se usa junto con un número, se refiere a cualquier número dentro de ± 1 , ± 5 o $\pm 10\%$ del número de referencia.

El término "porcentaje (%) de identidad de secuencia" o "% de identidad" describe el número de coincidencias ("aciertos") de aminoácidos idénticos de dos o más secuencias de aminoácidos alineadas en comparación con el número de residuos de aminoácidos que componen la longitud total de las secuencias de aminoácidos. En otros términos, utilizando una alineación, para dos o más secuencias el porcentaje de residuos de aminoácidos que son los mismos (por ejemplo, 90%, 95%, 97% o 98% de identidad) se puede determinar cuando las secuencias se comparan y se alinean para una correspondencia máxima, medida usando un algoritmo de comparación de secuencias, tal como se conoce en la técnica, o cuando se alinean manualmente y se inspeccionan visualmente. Las secuencias que se comparan para determinar la identidad de secuencia pueden por lo tanto diferir por sustituciones, adiciones o eliminaciones de aminoácidos. Los programas adecuados para alinear secuencias de proteínas son conocidos para la persona experta. El porcentaje de identidad de secuencia de secuencias de proteínas se puede determinar, por

ejemplo, con programas tales como CLUSTALW, Clustal Omega, FASTA o BLAST, por ejemplo, usando el algoritmo NCBI BLAST (Altschul S.F., *et al.* (1997), *Nucleic Acids Res.* 25:3389–3402).

- 5 Por ejemplo, para secuencias de aminoácidos, la identidad y/o similitud de secuencia se puede determinar mediante el uso de técnicas estándar conocidas en la técnica, que incluyen, pero sin limitaciones, el algoritmo de identidad de secuencias local de Smith y Waterman, 1981, *Adv. Appl. Math.* 2:482, el algoritmo de alineación de identidad de secuencias de Needleman y Wunsch, 1970, *J. Mol. Biol.* 48:443, el método de búsqueda de similitud de Pearson y Lipman, 1988, *Proc. Nat. Acad. Sci. U.S.A.* 85:2444, implementaciones informáticas de estos algoritmos (GAP, BESTFIT, FASTA y TFASTA en el Paquete de Software de Wisconsin Genetics, Genetics Computer Group, 575 Science Drive, Madison, Wis.), el programa de secuencias Best Fit descrito por Devereux *et al.*, 1984, *Nucl. Acid Res.* 12:387–395, preferentemente con la configuración predeterminada, o por inspección. En ciertas formas de realización, el porcentaje de identidad se calcula mediante FastDB en base a los siguientes parámetros: penalización por desajustes de 1; penalización por huecos de 1; penalización por tamaño de hueco de 0.33; y penalización por ingreso de 30, "Current Methods in Sequence Comparison and Analysis," *Macromolecule Sequencing and Synthesis, Selected Methods and Applications*, pp 127–149 (1988), Alan R. Liss, Inc.
- 10
- 15 Otro ejemplo de un algoritmo útil es el algoritmo BLAST, descrito en: Altschul *et al.*, 1990, *J. Mol. Biol.* 215:403–410; Altschul *et al.*, 1997, *Nucleic Acids Res.* 25:3389–3402; y Karin *et al.*, 1993, *Proc. Natl. Acad. Sci. U.S.A.* 90:5873–5787. Un programa BLAST particularmente útil es el programa WU-BLAST-2, que fue obtenido por Altschul *et al.*, 1996, *Methods in Enzymology* 266:460–480. WU-BLAST-2 usa varios parámetros de búsqueda, la mayoría de los cuales se establecen en los valores por defecto.
- 20 Un algoritmo útil adicional es gapped BLAST según lo informado por Altschul *et al.*, 1993, *Nucl. Acids Res.* 25:3389–3402.

25 El término "enfermedad por *Escherichia coli* patogénica extraintestinal (ExPEC) invasiva (IED)" se define aquí como una enfermedad aguda consistente con infección bacteriana sistémica, que se confirma microbiológicamente ya sea por el aislamiento y la identificación de *E. coli* en sangre u otros sitios corporales normalmente estériles, o por el aislamiento y la identificación de *E. coli* en la orina en un paciente con presencia de signos y síntomas de enfermedad invasiva (síndrome de respuesta inflamatoria sistémica (SIRS), sepsis o shock séptico) y sin otra fuente de identificación de infección.

Bioconjugados de O4 antígenos polisacáridos glucosilados de *E. coli*

30 En un aspecto, se proporciona en este documento un bioconjugado de un O4 antígeno polisacárido glucosilado de *E. coli* unido covalentemente a una proteína portadora. Tal como se utiliza aquí, el término "O4" se refiere al O4 antígeno de *E. coli* (*E. coli* serotipo O4). Se sabe que la modificación estructural de O-antígeno existe dentro de *E. coli* serotipo O4. En particular, algunos O4 serotipos expresan un O-antígeno modificado que tiene una unidad de glucosa ramificada. Tal como se usa en este documento, "O4 antígeno glucosilado", "O4 antígeno polisacárido glucosilado", "O4-Glc+ antígeno polisacárido" y "O4-Glc+ antígeno" se refieren a un O4 antígeno (por ejemplo, O4 antígeno de *E. coli*) que tiene una rama de glucosa, en el que la D-glucosa está unida a L-ramnosa en la unidad de repetición L-Rha → D-Glc → L-FucNAc → D-GlcNAc. En una forma de realización particular, un O4 antígeno polisacárido glucosilado de *E. coli* comprende la estructura de fórmula (O4-Glc+), como se muestra en la Tabla 1, en donde n es un número entero de 1 a 100. En formas de realización preferidas, n es un número entero de 3 a 50, por ejemplo de 5 a 40, por ejemplo de 7 a 25, por ejemplo de 10 a 20.

40 Las cepas de O4 de *E. coli*, independiente del estado de ramificación de glucosa, llevan un grupo de genes *rfb* sustancialmente idéntico que codifica los genes responsables de la producción del O4 antígeno polisacárido. Sin embargo, la síntesis *in vivo* del antígeno O4 modificado que tiene ramificación de glucosa requiere la actividad de una enzima de ramificación de polisacárido, que se encuentra fuera del grupo de genes *rfb*. Al leer y entender de los inventores, la identidad de la enzima de ramificación del polisacárido responsable de la modificación de glucosa del O4 antígeno ha permanecido desconocida hasta la fecha. Aquí, los inventores descubrieron la secuencia de la enzima de ramificación del polisacárido responsable de la modificación de glucosa del O4 antígeno. La identificación de esta enzima permite la producción de bioconjugados del O4 antígeno polisacárido modificado que tiene ramificación de glucosa. La forma modificada por glucosa del O4 antígeno polisacárido está presente en los serotipos predominantes y, por lo tanto, puede usarse para proporcionar una respuesta inmune mejorada, por ejemplo, para uso profiláctico o terapéutico.

55 En particular, se proporciona en este documento la secuencia de un gen *gtrS* que codifica una enzima glucosiltransferasa específica para *E. coli* serotipo O4 que glucosila el O4 antígeno. En general, los genes *gtrA*, *gtrB* y *gtrS* codifican las enzimas responsables de la glucosilación de O-antígeno. Mientras que los genes *gtrA* y *gtrB* en diferentes serotipos son altamente homólogos e intercambiables, el gen *gtrS* codifica una glucosil transferasa de O-antígeno de serotipo específico. El gen *gtrS* de *E. coli* serotipo O4 codifica la enzima GtrS que modifica el O4 antígeno mediante la introducción de ramificación de glucosa. La caracterización de aislados de *E. coli* clínicos contemporáneos del serotipo O4 reveló la presencia de *gtrS* en el 78% de los aislados ensayados, lo que indica que O4 antígeno

polisacárido de *E. coli* modificado con la adición de un residuo de glucosa es predominante en los actuales aislados infecciosos.

En una forma de realización, en este documento se proporciona un ácido nucleico de un gen *gtrS* de *E. coli* serotipo O4 que codifica una glucosiltransferasa GtrS que comprende la secuencia de aminoácidos de SEQ ID NO: 4. En otra forma de realización, un ácido nucleico *gtrS* codifica una proteína GtrS de *E. coli* serotipo O4 que es aproximadamente 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% o 100% idéntica a la secuencia de aminoácidos de SEQ ID NO: 4, preferiblemente 98%, 99% o 100% idéntica a la secuencia de aminoácidos de SEQ ID NO: 4. Una proteína GtrS que es al menos 80% idéntica a la secuencia de aminoácidos de SEQ ID NO: 4 es capaz de glucosilar específicamente el O4 antígeno polisacárido de *E. coli* para obtener un O4 antígeno glucosilado que tiene la estructura de Fórmula (O4–Glc+) como se muestra en la Tabla 1. Un experto en la técnica podrá hacer formas mutadas de la proteína GtrS de SEQ ID NO: 4 que tiene al menos 80% de identidad de secuencia con SEQ ID NO: 4, y evaluar tales secuencias en cuanto a la actividad de glucosilación del O4 antígeno de *E. coli* en vista de la presente divulgación. Las células huésped recombinantes que comprenden la secuencia de ácido nucleico que codifica la glucosiltransferasa gen *gtrS* de *E. coli* serotipo O4, y el uso de las células huésped recombinantes en la producción de los O4 antígenos polisacáridos modificados con glucosa y bioconjugados de los mismos se describen en mayor detalle a continuación.

Las secuencias para proteínas codificadas por *gtrA* y *gtrB*, que funcionan como glucosa translocasa ligada a bactoprenol (GtrA, voltea la glucosa ligada a bactoprenol sobre la membrana interna al periplasma) y bactoprenol glucosil transferasa (GtrB, une la glucosa al bactoprenol), respectivamente, pueden comprender secuencias de aminoácidos que son al menos aproximadamente 80% idénticas a SEQ ID NOs: 7 y 8, respectivamente. En ciertas formas de realización, las secuencias de ácido nucleico que codifican proteínas GtrA y GtrB que son al menos aproximadamente 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% o 100% idénticas a SEQ ID NOs: 7 y 8, respectivamente, y que tienen actividad de glucosa translocasa unida a bactoprenol y bactoprenol glucosil transferasa, respectivamente, también están presentes en las células huésped de la invención, que comprenden además un locus *rfb* O4-específico, la secuencia codificante de GtrS específica de O4 que se describió antes, una oligosacáril transferasa como se describe en la presente y una secuencia que codifica una proteína portadora que tiene una o más secuencias de consenso de glicosilación como se describe en la presente, para producir bioconjugados de *E. coli* serotipo O4 glucosilado (que comprenden la estructura de glicano de Fórmula (O4–Glc+) de la Tabla 1).

Los bioconjugados de un O4 antígeno polisacárido glucosilado de *E. coli* proporcionado en el presente documento están unidos covalentemente a una proteína portadora, preferentemente por un enlace glicosídico. Cualquier proteína portadora conocida por los expertos en la técnica en vista de la presente divulgación puede ser utilizada. Las proteínas portadoras adecuadas incluyen, pero sin limitaciones, Exotoxina A desintoxicada de *P. aeruginosa* (EPA), flagelina de *E. coli* (FlhC), CRM197, proteína de unión a maltosa (MBP), toxoide diftérico, toxoide tetánico, hemolisina A desintoxicada de *S. aureus*, factor de aglutinación A, factor de aglutinación B, enterotoxina termolábil de *E. coli*, variantes desintoxicadas de enterotoxina termolábil de *E. coli*, toxina del cólera subunidad B (CTB), toxina del cólera, variantes desintoxicadas de la toxina del cólera, proteína Sat de *E. coli*, el dominio pasajero de la proteína Sat de *E. coli*, neumolisina de *Streptococcus pneumoniae*, hemocianina de lapa californiana (KLH), PcrV de *P. aeruginosa*, proteína de membrana externa de *Neisseria meningitidis* (OMPC) y proteína D de *Haemophilus influenzae* no tipificable. Se ha descrito la bioconjugación con diversas proteínas portadoras diferentes que contienen la secuencia de consenso de glicosilación requerida, que muestra que una amplia gama de proteínas puede ser glicosilada con esta tecnología (véase, por ejemplo, WO 06/119987, WO 2015/124769, WO 2015/158403, WO 2015/82571, WO 2017/216286 y WO 2017/67964, que juntos muestran una amplia variedad de proteínas portadoras que se utilizaron con éxito en la bioconjugación).

En ciertas formas de realización, una proteína portadora se modifica, por ejemplo, se modifica de tal manera que la proteína es menos tóxica y/o más susceptible a la glicosilación. En una forma de realización específica, las proteínas portadoras usadas aquí se modifican de tal manera que el número de sitios de glicosilación en las proteínas portadoras se maximiza de manera que permite administrar concentraciones más bajas de la proteína, por ejemplo, en una composición inmunogénica, particularmente en su forma de bioconjugado.

Por lo tanto, en ciertas formas de realización, las proteínas portadoras descritas aquí se modifican para incluir 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más sitios de glicosilación de lo que normalmente se asocia con la proteína portadora (por ejemplo, en relación con el número de sitios de glicosilación asociados con la proteína portadora en su estado nativo o natural, es decir, “de tipo salvaje”). La introducción de sitios de glicosilación en una proteína portadora puede lograrse mediante la inserción de una secuencia de consenso de glicosilación en cualquier parte de la estructura primaria de la proteína mediante, por ejemplo, la adición de nuevos aminoácidos a la estructura primaria de la proteína, de tal manera que un sitio de glicosilación se añade en su totalidad o en parte, o mediante la mutación de los aminoácidos existentes en la proteína con el fin de generar un sitio de glicosilación. El técnico con experiencia en el arte reconocerá que la secuencia de aminoácidos de una proteína se puede modificar fácilmente usando enfoques conocidos en el arte, por ejemplo, enfoques recombinantes que incluyen la modificación de la secuencia de ácido nucleico que codifica la proteína. En formas de realización específicas, las secuencias de consenso de glicosilación se introducen en regiones específicas de la proteína portadora, por ejemplo, en estructuras de superficie de la proteína, en los extremos N o C de la proteína, y/o en los bucles que están estabilizados por puentes disulfuro en la base de la proteína. En algunas formas de

realización, una secuencia de consenso de glicosilación puede extenderse mediante la adición de residuos de lisina para una glicosilación más eficiente.

Los ejemplos de secuencias de consenso de glicosilación que se pueden insertar o generan en una proteína portadora incluyen Asn–X–Ser(Thr), en donde X puede ser cualquier aminoácido excepto Pro (SEQ ID NO: 1); y Asp(Glu)–X–Asn–Z–Ser(Thr), en donde X y Z se seleccionan independientemente entre cualquier aminoácido excepto Pro (SEQ ID NO: 2).

En algunas formas de realización, el O4 antígeno polisacárido glucosilado de *E. coli* se une covalentemente a un residuo de asparagina (Asn) en la proteína portadora (por ejemplo, N–ligado), en donde el residuo de Asn está presente en un sitio de glicosilación que comprende una secuencia de consenso de glicosilación que tiene SEQ ID NO: 1, más preferiblemente que tiene SEQ ID NO: 2. Típicamente, una proteína portadora comprende 1–10 sitios de glicosilación, preferiblemente 2 a 4 sitios de glicosilación, lo más preferiblemente 4 sitios de glicosilación, tal como 1–10, preferiblemente 2–4 y más preferiblemente 4 sitios de glicosilación que comprenden, cada uno, una secuencia de consenso de glicosilación que tiene la secuencia de aminoácidos de SEQ ID NO: 1, y más preferiblemente la secuencia de aminoácidos de SEQ ID NO: 2.

En formas de realización particulares, una proteína portadora es una Exotoxina A desintoxicada de *P. aeruginosa*. Para EPA, diversas variantes de proteínas desintoxicadas se han descrito en la literatura y podrían utilizarse como proteínas portadoras. Por ejemplo, la desintoxicación se puede lograr mediante la mutación y eliminación de los residuos catalíticamente esenciales L552V y ΔE553 de acuerdo con Lukac *et al.*, 1988, *Infect Immun.*, 56: 3095–3098, y Ho *et al.*, 2006, *Hum Vaccin*. Tal como se usa en el presente documento, “EPA” se refiere a un Exotoxina A desintoxicada de *P. aeruginosa*. En esas formas de realización, en donde la proteína portadora es EPA, un O4 antígeno polisacárido glucosilado de *E. coli* se puede unir covalentemente a un residuo de Asn en un sitio de glicosilación que comprende una secuencia de consenso de glicosilación que tiene la SEQ ID NO: 1, y preferiblemente ligado covalentemente a un residuo de Asn en un sitio de glicosilación que comprende una secuencia de consenso de glicosilación que tiene la SEQ ID NO: 2. Preferiblemente, la proteína portadora EPA comprende 1–10 sitios de glicosilación, preferiblemente 2 a 4 sitios de glicosilación, lo más preferiblemente 4 sitios de glicosilación, tal como 1–10, preferiblemente 2–4, y más preferiblemente 4 sitios de glicosilación que comprenden, cada uno, una secuencia de consenso de glicosilación que tiene la secuencia de aminoácidos de SEQ ID NO: 1, y más preferiblemente la secuencia de aminoácidos de SEQ ID NO: 2.

En algunas formas de realización, la proteína portadora EPA comprende cuatro sitios de glicosilación que comprenden, cada uno, una secuencia de consenso de glicosilación, por ejemplo, un sitio de glicosilación que comprende una secuencia de consenso de glicosilación que tiene la SEQ ID NO: 2. Tal como se usa en este documento, “proteína portadora EPA–4” y “EPA–4” se refieren a una proteína portadora de Exotoxina A desintoxicada de *P. aeruginosa* que comprende cuatro sitios de glicosilación que comprenden, cada uno, una secuencia de consenso de glicosilación que tiene SEQ ID NO: 2. Un ejemplo preferido de una proteína portadora EPA–4 es una proteína portadora EPA que comprende la secuencia de aminoácidos de SEQ ID NO: 3.

Composiciones

En otro aspecto, se proporciona en este documento una composición que comprende un bioconjugado de un O4 antígeno polisacárido glucosilado de *E. coli* unido covalentemente a una proteína portadora. Las composiciones proporcionadas en el presente documento pueden incluir cualquier bioconjugado de un O4 antígeno polisacárido glucosilado de *E. coli* unido covalentemente a una proteína portadora (por ejemplo, EPA) descrita en este documento.

En algunas formas de realización, una composición es una composición inmunogénica. Tal como se usa en este documento, una “composición inmunogénica” se refiere a una composición que puede provocar una respuesta inmune en un huésped o sujeto al que se administra la composición. Las composiciones inmunogénicas pueden comprender además un vehículo farmacéuticamente aceptable. En algunas formas de realización, una composición es una composición farmacéutica que comprende además un vehículo farmacéuticamente aceptable. Tal como se usa en este documento, un “vehículo farmacéuticamente aceptable” se refiere a un diluyente, adyuvante, excipiente o vehículo con el que se administra una composición, y que es no tóxico y no debe interferir con la eficacia del ingrediente activo. Por ejemplo, las soluciones salinas y soluciones acuosas de dextrosa y glicerol también pueden emplearse como vehículos líquidos, particularmente para soluciones inyectables. Los excipientes adecuados incluyen almidón, glucosa, lactosa, sacarosa, gelatina, malta, arroz, harina, tiza, gel de sílice, estearato de sodio, monoestearato de glicerol, talco, cloruro de sodio, leche desnatada en polvo, glicerol, propilenglicol, agua, etanol y similares. Otros ejemplos de vehículos farmacéuticamente aceptables adecuados se describen en “Remington’s Pharmaceutical Sciences” por E.W. Martin.

En una forma de realización, una composición de la invención comprende los bioconjugados de la invención en una solución salina tamponada con Tris (TBS) pH 7.4 (por ejemplo, que contiene Tris, NaCl y KCl, por ejemplo a 25 mM, 137 mM y 2.7 mM, respectivamente). En otras formas de realización, las composiciones de la invención comprenden bioconjugados de la invención en aproximadamente 10 mM de tampón $\text{KH}_2\text{PO}_4/\text{Na}_2\text{HPO}_4$ a un pH de aproximadamente 7.0, aproximadamente 5% (p/v) de sorbitol, aproximadamente 10 mM de metionina y

aproximadamente 0.02% (p/v) de polisorbato 80. En otras formas de realización, las composiciones de la invención comprenden bioconjugados de la invención en aproximadamente 10 mM de tampón $\text{KH}_2\text{PO}_4/\text{Na}_2\text{HPO}_4$ a un pH de aproximadamente 7.0, aproximadamente 8% (p/v) de sacarosa, aproximadamente 1 mM de EDTA y aproximadamente 0.02% (p/v) de polisorbato 80 (véase, por ejemplo, el documento WO 2018/077853 para obtener tampones adecuados para bioconjugados de O-antígenos de *E. coli* unidos covalentemente a una proteína portadora EPA).

En algunas formas de realización, las composiciones descritas en el presente documento son formulaciones monovalentes, y contienen un O-antígeno polisacárido de *E. coli*, por ejemplo, en forma aislada o como parte de un glicoconjugado o bioconjugado, tal como el O4 antígeno polisacárido glucosilado de *E. coli*. También se proporcionan en el presente documento composiciones (por ejemplo, composiciones farmacéuticas y/o inmunogénicas) que son composiciones multivalentes, por ejemplo, composiciones bivalentes, trivalentes, tetravalentes, etc. Por ejemplo, una composición multivalente comprende más de un antígeno, tal como un O-antígeno de *E. coli*, glicoconjugado o bioconjugado del mismo. En formas de realización particulares, las composiciones multivalentes proporcionadas en este documento comprenden un bioconjugado de un O4 antígeno polisacárido glucosilado de *E. coli*, y al menos un antígeno adicional.

En una forma de realización, una composición (por ejemplo, una composición farmacéutica y/o inmunogénica) es una composición monovalente que comprende un bioconjugado de un O4 antígeno polisacárido glucosilado de *E. coli* unido covalentemente a una proteína portadora tal como se describe aquí.

En otra forma de realización, una composición (por ejemplo, una composición farmacéutica y/o inmunogénica) es una composición multivalente que comprende un O4 antígeno polisacárido glucosilado de *E. coli* unido covalentemente a una proteína portadora tal como se describe aquí, y al menos un antígeno adicional.

En algunas formas de realización, el antígeno adicional es un antígeno sacárido o polisacárido, más preferiblemente un O-antígeno polisacárido de *E. coli*, tal como O-antígenos de *E. coli* de uno o más de los serotipos O1, O2, O6, O8, O15, O16, O18, O25, O75 y subserotipos de los mismos. En algunas formas de realización, cada uno de los O-antígenos polisacáridos de *E. coli* adicionales es un glicoconjugado, lo que significa que el O-antígeno polisacárido de *E. coli* se une covalentemente a otra especie química, por ejemplo, proteína, péptido, lípido, etc., lo más preferiblemente una proteína portadora, por ejemplo por métodos químicos o enzimáticos. En formas de realización preferidas, cada uno de los O-antígenos polisacáridos de *E. coli* adicionales es un bioconjugado en el que el O-antígeno polisacárido se une covalentemente, por ejemplo, a una proteína portadora, a través de un enlace glicosídico enzimáticamente por la maquinaria de la célula huésped. Las composiciones proporcionadas en este documento en ciertas formas de realización pueden comprender 1–20 glicoconjugados adicionales, más preferiblemente bioconjugados de O-antígenos polisacáridos de *E. coli*, tal como 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20 glicoconjugados adicionales o preferiblemente bioconjugados de O-antígenos polisacáridos de *E. coli*. Se pueden incluir otros antígenos en las composiciones proporcionadas en el presente documento, tales como péptidos, proteínas o antígenos lipídicos, etc.

En algunas formas de realización, una composición (por ejemplo, una composición farmacéutica y/o inmunogénica) comprende un bioconjugado de un O4 antígeno polisacárido glucosilado de *E. coli*, y al menos un antígeno polisacárido adicional seleccionado del grupo que consiste en O1A antígeno polisacárido de *E. coli*, O2 antígeno polisacárido de *E. coli*, O6A antígeno polisacárido de *E. coli*, O8 antígeno polisacárido de *E. coli*, O15 antígeno polisacárido de *E. coli*, O16 antígeno polisacárido de *E. coli*, O18A antígeno polisacárido de *E. coli*, O25B antígeno polisacárido de *E. coli* y O75 antígeno polisacárido de *E. coli*. Preferiblemente, cada uno de los O-antígenos polisacáridos adicionales está unido covalentemente a una proteína portadora, y es más preferiblemente un bioconjugado.

En una forma de realización, un O1A antígeno polisacárido (por ejemplo, en forma aislada o como parte de un glicoconjugado o bioconjugado) se utiliza en una composición proporcionada en este documento (por ejemplo, en combinación con un O4 antígeno polisacárido glucosilado o un bioconjugado del mismo). En una forma de realización específica, el O1A antígeno polisacárido comprende la estructura de fórmula (O1A) como se muestra en la Tabla 1, en donde n es un número entero de 1–100, preferiblemente 3–50, por ejemplo 5–40, por ejemplo 7 a 25, por ejemplo 10 a 20. Preferiblemente, el O1A antígeno polisacárido es parte de un bioconjugado y está unido covalentemente a una proteína portadora, por ejemplo, EPA.

En una forma de realización, un O2 antígeno polisacárido (por ejemplo, en forma aislada o como parte de un glicoconjugado o bioconjugado) se utiliza en una composición proporcionada en este documento (por ejemplo, en combinación con un O4 antígeno polisacárido glucosilado o un bioconjugado del mismo). En una forma de realización específica, el O2 antígeno polisacárido comprende la estructura de fórmula (O2) como se muestra en la Tabla 1, en donde n es un número entero de 1–100, preferiblemente 3–50, por ejemplo 5–40, por ejemplo 7 a 25, por ejemplo 10 a 20. Preferiblemente, el O2 antígeno polisacárido es parte de un bioconjugado y está unido covalentemente a una proteína portadora, por ejemplo, EPA.

En una forma de realización, un OA6 antígeno polisacárido (por ejemplo, en forma aislada o como parte de un glicoconjugado o bioconjugado) se utiliza en una composición proporcionada en este documento (por ejemplo, en combinación con un O4 antígeno polisacárido glucosilado o un bioconjugado del mismo). En una forma de realización

específica, el O6A antígeno polisacárido comprende la estructura de fórmula (O6A) como se muestra en la Tabla 1, en donde n es un número entero de 1–100, preferiblemente 3–50, por ejemplo 5–40, por ejemplo 7 a 25, por ejemplo 10 a 20. Preferiblemente, el O6A antígeno polisacárido es parte de un bioconjugado y está unido covalentemente a una proteína portadora, por ejemplo, EPA.

5 En una forma de realización, un O8 antígeno polisacárido (por ejemplo, en forma aislada o como parte de un glicoconjugado o bioconjugado) se utiliza en una composición proporcionada en este documento (por ejemplo, en combinación con un O4 antígeno polisacárido glucosilado o un bioconjugado del mismo). En una forma de realización específica, el O8 antígeno polisacárido comprende la estructura de fórmula (O8) como se muestra en la Tabla 1, en donde n es un número entero de 1–100, preferiblemente 3–50, por ejemplo 5–40, por ejemplo 7 a 25, por ejemplo 10 a 20. Preferiblemente, el O8 antígeno polisacárido es parte de un bioconjugado y está unido covalentemente a una proteína portadora, por ejemplo, EPA.

15 En una forma de realización, un O15 antígeno polisacárido (por ejemplo, en forma aislada o como parte de un glicoconjugado o bioconjugado) se utiliza en una composición proporcionada en este documento (por ejemplo, en combinación con un O4 antígeno polisacárido glucosilado o un bioconjugado del mismo). En una forma de realización específica, el O15 antígeno polisacárido comprende la estructura de fórmula (O15) como se muestra en la Tabla 1, en donde n es un número entero de 1–100, preferiblemente 3–50, por ejemplo 5–40, por ejemplo 7 a 25, por ejemplo 10 a 20. Preferiblemente, el O15 antígeno polisacárido es parte de un bioconjugado y está unido covalentemente a una proteína portadora, por ejemplo, EPA.

20 En una forma de realización, un O16 antígeno polisacárido (por ejemplo, en forma aislada o como parte de un glicoconjugado o bioconjugado) se utiliza en una composición proporcionada en este documento (por ejemplo, en combinación con un O4 antígeno polisacárido glucosilado o un bioconjugado del mismo). En una forma de realización específica, el O16 antígeno polisacárido comprende la estructura de fórmula (O16) como se muestra en la Tabla 1, en donde n es un número entero de 1–100, preferiblemente 3–50, por ejemplo 5–40, por ejemplo 7 a 25, por ejemplo 10 a 20. Preferiblemente, el O16 antígeno polisacárido es parte de un bioconjugado y está unido covalentemente a una proteína portadora, por ejemplo, EPA.

25 En una forma de realización, un O18A antígeno polisacárido (por ejemplo, en forma aislada o como parte de un glicoconjugado o bioconjugado) se utiliza en una composición proporcionada en este documento (por ejemplo, en combinación con un O4 antígeno polisacárido glucosilado o un bioconjugado del mismo). En una forma de realización específica, el O18A antígeno polisacárido comprende la estructura de fórmula (O18A) como se muestra en la Tabla 1, en donde n es un número entero de 1–100, preferiblemente 3–50, por ejemplo 5–40, por ejemplo 7 a 25, por ejemplo 10 a 20. Preferiblemente, el O18A antígeno polisacárido es parte de un bioconjugado y está unido covalentemente a una proteína portadora, por ejemplo, EPA.

30 En una forma de realización, un O25B antígeno polisacárido (por ejemplo, en forma aislada o como parte de un glicoconjugado o bioconjugado) se utiliza en una composición proporcionada en este documento (por ejemplo, en combinación con un O4 antígeno polisacárido glucosilado o un bioconjugado del mismo). En una forma de realización específica, el O25B antígeno polisacárido comprende la estructura de fórmula (O25B) como se muestra en la Tabla 1, en donde n es un número entero de 1–100, preferiblemente 3–50, por ejemplo 5–40, por ejemplo 7 a 25, por ejemplo 10 a 20. Preferiblemente, el O25B antígeno polisacárido es parte de un bioconjugado y está unido covalentemente a una proteína portadora, por ejemplo, EPA.

35 En una forma de realización, un O75 antígeno polisacárido (por ejemplo, en forma aislada o como parte de un glicoconjugado o bioconjugado) se utiliza en una composición proporcionada en este documento (por ejemplo, en combinación con un O4 antígeno polisacárido glucosilado o un bioconjugado del mismo). En una forma de realización específica, el O75 antígeno polisacárido comprende la estructura de fórmula (O75) como se muestra en la Tabla 1, en donde n es un número entero de 1–100, preferiblemente 3–50, por ejemplo 5–40, por ejemplo 7 a 25, por ejemplo 10 a 20. Preferiblemente, el O75 antígeno polisacárido es parte de un bioconjugado y está unido covalentemente a una proteína portadora, por ejemplo, EPA.

40 En otra forma de realización, una composición (por ejemplo, una composición farmacéutica y/o inmunogénica) comprende al menos el O1A, O2, O4 glucosilado, O6A y O25B antígenos polisacáridos de *E. coli*, preferiblemente bioconjugados de los O1A, O2, O4 glucosilado, O6A y O25B antígenos polisacáridos unidos covalentemente a una proteína portadora, por ejemplo, EPA (es decir, una composición pentavalente).

45 En una forma de realización preferida, una composición (por ejemplo, una composición farmacéutica y/o inmunogénica) comprende al menos el O1A, O2, O4 glucosilado, O6A, O8, O15, O16, O25B y O75 antígenos polisacáridos glucosilados de *E. coli*, preferiblemente bioconjugados de los O1A, O2, O4 glucosilado, O6A, O8, O15, O16, O25B y O75 antígenos polisacáridos unidos covalentemente a una proteína portadora, por ejemplo, EPA (es decir, una composición 9–valente).

50 En otra forma de realización preferida, una composición (por ejemplo, una composición farmacéutica y/o inmunogénica) comprende al menos el O1A, O2, O4 glucosilado, O6A, O8, O15, O16, O18A, O25B y O75 antígenos

polisacáridos de *E. coli*, preferiblemente bioconjugados de los O1A, O2, O4 glucosilado, O6A, O8, O15, O16, O18A, O25B y O75 antígenos polisacáridos unidos covalentemente a una proteína portadora, por ejemplo, EPA (es decir, una composición 10–valente).

5 También se contemplan en el presente documento composiciones que comprenden además opcionalmente O–antígenos adicionales (por ejemplo, en forma aislada, o como parte de un glicoconjugado o bioconjugado) de otros serotipos de *E. coli*.

10 En algunas formas de realización, cada uno de los O1A, O2, O6A, O8, O15, O16, O18A, O25B y/o O75 antígenos polisacáridos de *E. coli* adicionales se une covalentemente a una proteína portadora. El O–antígeno polisacárido se puede unir a una proteína portadora mediante métodos sintéticos u otros métodos químicos, o el O–antígeno polisacárido puede ser parte de un bioconjugado, y es preferiblemente parte de un bioconjugado. Cualquier proteína portadora conocida para los expertos en la técnica en vista de la presente divulgación puede ser utilizada. Las proteínas portadoras adecuadas incluyen, pero sin limitaciones, Exotoxina A desintoxicada de *P. aeruginosa* (EPA), flagelina de *E. coli* (FliC), CRM197, proteína de unión a maltosa (MBP), toxoide diftérico, toxoide tetánico, hemolisina A desintoxicada de *S. aureus*, factor de aglutinación A, factor de aglutinación B, enterotoxina termolábil de *E. coli*,
15 variantes desintoxicadas de enterotoxina termolábil de *E. coli*, toxina del cólera subunidad B (CTB), toxina del cólera, variantes desintoxicadas de la toxina del cólera, proteína Sat de *E. coli*, el dominio pasajero de la proteína Sat de *E. coli*, neumolisina de *Streptococcus pneumoniae*, hemocianina de lapa californiana (KLH), PcrV de *P. aeruginosa*, proteína de membrana externa de *Neisseria meningitidis* (OMPC) y proteína D de *Haemophilus influenzae* no tipificable. Preferiblemente, la proteína portadora es EPA.

20 En algunas formas de realización, cada uno de los O1A, O2, O6A, O8, O15, O16, O18A, O25B y/o O75 antígenos polisacáridos de *E. coli* adicionales, particularmente cuando son parte de un bioconjugado, se une covalentemente a un residuo de asparagina (Asn) en la proteína portadora, en donde el residuo de Asn está presente en un sitio de glicosilación que comprende una secuencia de consenso de glicosilación Asn–X–Ser(Thr), en donde X puede ser cualquier aminoácido excepto Pro (SEQ ID NO: 1), preferiblemente en donde el residuo de Asn está presente en un
25 sitio de glicosilación que comprende una secuencia de consenso de glicosilación Asp(Glu)–X–Asn–Z–Ser(Thr), en donde X y Z se seleccionan independientemente entre cualquier aminoácido excepto Pro (SEQ ID NO: 2). La proteína portadora puede comprender 1–10 sitios de glicosilación, preferiblemente 2 a 4 sitios de glicosilación, lo más preferiblemente 4 sitios de glicosilación, comprendiendo cada uno una secuencia de consenso de glicosilación. En una forma de realización particular, la proteína portadora es la proteína portadora EPA–4, por ejemplo, la proteína portadora EPA–4 que comprende la secuencia de aminoácidos de SEQ ID NO: 3.

30 En una forma de realización particular, se proporciona en este documento una composición (por ejemplo, una composición farmacéutica y/o inmunogénica) que comprende: (i) un bioconjugado de un O4 antígeno polisacárido glucosilado de *E. coli* unido covalentemente a una Exotoxina A desintoxicada de *P. aeruginosa* que comprende SEQ ID NO: 3 (proteína portadora EPA–4), en donde el O4 antígeno polisacárido glucosilado de *E. coli* comprende la estructura de Fórmula (O4–Glc+); (ii) un bioconjugado de un O1A antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora EPA–4, en donde el O1A antígeno polisacárido de *E. coli* comprende la estructura de Fórmula (O1A); (iii) un bioconjugado de un O2 antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora EPA–4, en donde el O2 antígeno polisacárido de *E. coli* comprende la estructura de Fórmula (O2); (iv) un bioconjugado de un O6A antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora EPA–4, en donde el O6A antígeno polisacárido de *E. coli* comprende la estructura de Fórmula (O6A); (v) un bioconjugado de un O8 antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora EPA–4, en donde el O8 antígeno polisacárido de *E. coli* comprende la estructura de Fórmula (O8); (vi) un bioconjugado de un O15 antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora EPA–4, en donde el O15 antígeno polisacárido de *E. coli* comprende la estructura de Fórmula (O15); (vii) un bioconjugado de un O16 antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora EPA–4, en donde el O16 antígeno polisacárido de *E. coli* comprende la estructura de Fórmula (O16); (viii) un bioconjugado de un O25B antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora EPA–4, en donde el O25B antígeno polisacárido de *E. coli* comprende la estructura de Fórmula (O25B); y (ix) un bioconjugado de un O75 antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora EPA–4, en donde el O75 antígeno polisacárido de *E. coli* comprende la estructura de Fórmula (O75), en donde cada una de las fórmulas se proporciona en la Tabla 1, y para cada una de las fórmulas independientemente n es un número entero de 1 a 100, por ejemplo 1 a 50, preferiblemente de 3 a 50, por ejemplo de 5 a 40.

45 En una forma de realización particular, dicha composición (por ejemplo, una composición farmacéutica y/o inmunogénica) comprende además: (x) un bioconjugado de un O18A antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora EPA–4, en donde el O18A antígeno polisacárido de *E. coli* comprende la estructura de Fórmula (O18A) como se muestra en la Tabla 1, en donde n para esta estructura es un número entero de 1 a 100, por ejemplo de 1 a 50, preferiblemente de 3 a 50, por ejemplo de 5 a 40.

50 En algunas formas de realización, una composición proporcionada en el presente documento comprende un bioconjugado de un O4 antígeno polisacárido glucosilado de *E. coli*, y al menos un bioconjugado de un O25B antígeno

polisacárido de *E. coli*, en donde el bioconjugado del O25B antígeno polisacárido de *E. coli* está presente en la composición a una concentración que es de aproximadamente 1.5 a 6 veces, por ejemplo de aproximadamente 2 a 4 veces mayor, tal como 1.5, 2, 3, 4, 5 o 6 veces mayor que la concentración de cualquiera de los otros bioconjugados presentes en la composición.

5 En formas de realización particulares, una composición comprende bioconjugados de O1A, O2, O4 glucosilado, O6A, O8, O15, O16, O25B y O75 antígenos polisacáridos de *E. coli*, en donde los bioconjugados de O1A: O2: O4 glucosilado: O6A: O8: O15: O16: O25B: O75 están presentes en una relación (en peso de O–antígeno polisacárido) de 1:1:1:1:1:1:1:2:1 o 2:1:1:2:1:1:1:4:1.

10 En formas de realización particulares, una composición comprende bioconjugados de O1A, O2, O4 glucosilado, O6A, O8, O15, O16, O18A, O25B y O75 antígenos polisacáridos de *E. coli*, en donde los bioconjugados de O1A: O2: O4 glucosilado: O6A: O8: O15: O16: O18A: O25B: O75 están presentes en una relación (en peso de O–antígeno polisacárido) de 1:1:1:1:1:1:1:1:2:1 o 2:1:1:2:1:1:1:4:1.

15 En algunas formas de realización, una composición proporcionada en el presente documento comprende un bioconjugado de un O4 antígeno polisacárido glucosilado de *E. coli*, y al menos un bioconjugado de un O25B antígeno polisacárido de *E. coli*, en donde el bioconjugado del O25B antígeno polisacárido de *E. coli* está presente en la composición a una concentración de 2 a 50 µg/ml, preferiblemente de 8 a 40 µg/ml, más preferiblemente de 16–32 µg/ml, tal como 16, 18, 20, 22, 24, 26, 28, 30 o 32 µg/ml. En tales formas de realización, la concentración del bioconjugado del O25B antígeno polisacárido de *E. coli* es preferiblemente de aproximadamente 1.5 a 6 veces, por ejemplo de aproximadamente 2 a 4 veces mayor, tal como 1.5, 2, 3, 4, 5 o 6 veces mayor que la concentración de cualquiera de los otros bioconjugados presentes en la composición.

20

En ciertas formas de realización, las composiciones descritas en la presente memoria (por ejemplo, composiciones farmacéuticas y/o inmunogénicas) comprenden o se administran en combinación con un adyuvante. El adyuvante para la administración en combinación con una composición descrita en el presente documento puede administrarse antes (por ejemplo, dentro de las 72 horas, 48 horas, 24 horas, 12 horas, 6 horas, 2 horas, 1 hora, 10 minutos), concomitantemente o después (por ejemplo, dentro de las 72 horas, 48 horas, 24 horas, 12 horas, 6 horas, 2 horas, 1 hora, 10 minutos) de la administración de dicha composición. Tal como se utiliza aquí, el término “adyuvante” se refiere a un compuesto que, cuando se administra conjuntamente o como parte de una composición descrita en este documento, aumenta, mejora y/o potencia la respuesta inmune a un O–antígeno polisacárido de *E. coli* en un bioconjugado, pero cuando el compuesto adyuvante se administra solo no genera una respuesta inmune al O–antígeno polisacárido de *E. coli* en el bioconjugado. En algunas formas de realización, el adyuvante mejora una respuesta inmune a un O–antígeno polisacárido de *E. coli* en un bioconjugado del mismo y no produce una alergia ni otra reacción adversa. Los adyuvantes pueden mejorar una respuesta inmunitaria mediante varios mecanismos, que incluyen, por ejemplo, el reclutamiento de linfocitos, la estimulación de células B y/o T, y la estimulación de macrófagos.

25

30

Los ejemplos de adyuvantes adecuados incluyen, aunque sin limitaciones, sales de aluminio (alum) (tales como hidróxido de aluminio, fosfato de aluminio, sulfato de aluminio y óxido de aluminio, incluyendo nanopartículas que comprenden formulaciones de alum o nanoalum), fosfato de calcio, monofosforil lípido A (MPL) o 3–de–O–monofosforil lípido A acilado (3D–MPL) (véase, por ejemplo, Patente del Reino Unido GB2220211, EP0971739, EP1194166, US6491919), AS01, AS02, AS03 y AS04 (todas de GlaxoSmithKline; véase, por ejemplo, EP1126876, US7357936 para AS04, EP0671948, EP0761231, US5750110 para AS02), MF59 (Novartis), compuestos de imidazopiridina (véase WO2007/109812), compuestos de imidazoquinoxalina (véase WO2007/109813), delta–inulina, dinucleótidos cíclicos sintéticos activadores de STING (por ejemplo, US20150056224), combinaciones de lecitina y homopolímeros de carbómero (por ejemplo, US6676958) y saponinas, tales como QuilA y QS21 (véase, por ejemplo, Zhu D y W Tuo, 2016, Nat Prod Chem Res 3: e113 (doi: 10.4172/2329–6836.1000e113), Matrix M, Iscoms, Iscomatrix, etc., opcionalmente en combinación con QS7 (véase Kensil *et al.*, en Vaccine Design: The Subunit and Adjuvant Approach (eds. Powell & Newman, Plenum Press, NY, 1995); Patente de Estados Unidos N° 5.057.540). En algunas formas de realización, el adyuvante es adyuvante de Freund (completo o incompleto). Otros adyuvantes son emulsiones de aceite en agua (tales como escualeno o aceite de maní), opcionalmente en combinación con estimulantes inmunes, tales como monofosforil lípido A (véase Stoute *et al.*, N. Engl. J. Med. 336, 86–91 (1997)). Otro adyuvante es CpG (Bioworld Today, 15 de noviembre de 1998). Otros ejemplos de adyuvantes son liposomas que contienen estimulantes inmunes tales como MPL y QS21 como en AS01E y AS01B (por ejemplo, US 2011/0206758). Otros ejemplos de adyuvantes son CpG (Bioworld Today, 15 de noviembre de 1998) e imidazoquinolinas (tales como imiquimod y R848). Véase, por ejemplo, Reed G, *et al.*, 2013, *Nature Med.*, 19: 1597–1608. En ciertas formas de realización, el adyuvante contiene un agonista del receptor 4 tipo toll (TLR4). Los agonistas de TLR4 son bien conocidos en la técnica; véase, por ejemplo, Ireton GC y SG Reed, 2013, *Expert Rev Vaccines* 12: 793–807. En ciertas formas de realización, el adyuvante comprende un agonista de TLR4 que comprende lípido A, o un análogo o derivado del mismo, tal como MPL, 3D–MPL, RC529 (por ejemplo, EP1385541), PET–lípido A, GLA (adyuvante de glicopiranosilo lípido, un glicolípido disacárido sintético; por ejemplo, US20100310602, US8722064), SLA (por ejemplo, Carter D *et al.*, 2016, *Clin Transl Immunology* 5: e108 (doi:10.1038/cti.2016.63), que describe un enfoque de estructura–función para optimizar ligandos de TLR4 para vacunas humanas), PHAD (hexaacilo disacárido fosforilado), 3D–PHAD (cuya estructura es la misma que la de GLA), 3D–(6–acil)–PHAD (3D(6A)–PHAD) (PHAD, 3D–PHAD y 3D(6A)PHAD son variantes de lípido A

35

40

45

50

55

60

sintético; véase, por ejemplo, avantilipids.com/divisions/adjuvants, que también proporcionan estructuras de estas moléculas), E6020 (CAS Número 287180–63–6), ONO4007, OM–174 y similares.

En ciertas formas de realización, las composiciones descritas en este documento no comprenden, y no se administra en combinación con, un adyuvante.

- 5 En ciertas formas de realización, las composiciones descritas en este documento se formulan para ser adecuadas para la vía de administración prevista para un sujeto. Por ejemplo, las composiciones (por ejemplo, farmacéuticas y/o inmunogénicas) aquí descritas pueden formularse para administración subcutánea, parenteral, oral, sublingual, bucal, intradérmica, transdérmica, colorrectal, intraperitoneal, rectal, intravenosa, intranasal, intratraqueal, intramuscular, tópica, transdérmica o intradérmica. En una forma de realización específica, una composición proporcionada en el presente documento (por ejemplo, farmacéutica y/o inmunogénica) se formula para inyección intramuscular.

Métodos de uso

Los bioconjugados y las composiciones proporcionadas en este documento pueden usarse para inducir anticuerpos contra un O4 antígeno glucosilado de *E. coli* en un sujeto, y para vacunar a un sujeto contra *E. coli*, en particular *E. coli* patogénicas extraintestinales (ExPEC). Tal como se usa en el presente documento, "sujeto" se refiere a cualquier animal, preferiblemente un mamífero, al que se le administrará o se le ha administrado un bioconjugado o composición proporcionada en este documento. El término "mamífero", tal como se usa en el presente documento, abarca cualquier mamífero. Los ejemplos de mamíferos incluyen, pero sin limitaciones, vacas, caballos, ovejas, cerdos, gatos, perros, ratones, ratas, conejos, cobayos, primates no humanos (NHP) tales como monos o simios, seres humanos, etc. En ciertas formas de realización, un sujeto es un humano. Un sujeto humano puede ser de cualquier edad. En ciertas formas de realización, un sujeto es un humano de aproximadamente dos meses a aproximadamente 18 años de edad, por ejemplo, de 1 año a 18 años de edad. En ciertas formas de realización, un sujeto es un humano de al menos 18 años de edad. En ciertas formas de realización, un sujeto es un humano de 15 a 50 años de edad, por ejemplo, de 18 a 45 años, por ejemplo de 20 a 40 años de edad. En ciertas formas de realización, un sujeto es un varón humano. En ciertas formas de realización, un sujeto es una mujer humana. En ciertas formas de realización, un sujeto está inmunodeprimido. En ciertas formas de realización, un sujeto es un humano de al menos 50 años, al menos 55 años, al menos 60 años, al menos 65 años de edad. En ciertas formas de realización, un sujeto es un humano que no tiene más de 100 años, no tiene más de 95 años, no tiene más de 90 años, no tiene más de 85 años, no tiene más de 80 años o no tiene más de 75 años. En ciertas formas de realización, un sujeto es un humano de por lo menos 60 años y no más de 85 años de edad. En ciertas formas de realización, un sujeto es un humano con salud estable. En ciertas formas de realización, un sujeto es un adulto humano de al menos 60 y no más de 85 años de edad con salud estable. En ciertas formas de realización, un sujeto es un humano que tiene antecedentes de infección del tracto urinario (UTI, es decir, una infección bacteriana en la uretra, la vejiga, los uréteres y/o los riñones), es decir, que ha tenido al menos un episodio de UTI en su vida. En ciertas formas de realización, un sujeto es un humano que tiene antecedentes de UTI en los últimos veinte, quince, doce, diez, nueve, ocho, siete, seis, cinco, cuatro, tres, dos o un año. En ciertas formas de realización, un sujeto es un humano que tiene antecedentes de UTI en los últimos dos años. En ciertas formas de realización, un sujeto es un sujeto humano que tiene antecedentes de UTI recurrente, es decir, que ha tenido al menos dos UTI en seis meses o al menos tres UTI en un año. En ciertas formas de realización, un sujeto es un sujeto humano que tiene antecedentes de UTI recurrente en los últimos dos años. En ciertas formas de realización, un sujeto es un humano de 60 años o más con salud estable. En ciertas formas de realización, un sujeto es un humano de 60 años o más que tiene antecedentes de UTI en los últimos dos años. En ciertas formas de realización, un sujeto es un humano de al menos 60 años y menos de 75 años de edad que tiene antecedentes de UTI en los últimos dos años. En ciertas formas de realización, un sujeto es un sujeto humano de 75 años o más que tiene antecedentes de UTI en los últimos dos años. En ciertas formas de realización, un sujeto es un paciente programado para ser sometidos a procedimientos o cirugías urogenitales y/o abdominales electivas, por ejemplo una biopsia de próstata con aguja guiada por ecografía transrectal (TRUS–PNB).

En un aspecto, se proporciona en este documento un método para inducir anticuerpos contra un O4 antígeno glucosilado de *E. coli* en un sujeto, que comprende administrar al sujeto cualquiera de los bioconjugados de un O4 antígeno glucosilado de *E. coli* unido covalentemente a una proteína portadora descrita aquí, o una composición que comprende un bioconjugado de un O4 antígeno glucosilado de *E. coli* unido covalentemente a una proteína, solo o en combinación además con otros O–antígenos polisacáridos de *E. coli* o glicoconjugados o bioconjugados de los mismos.

En ciertas formas de realización, los anticuerpos inducidos, provocados o identificados contra un O4 antígeno glucosilado de *E. coli* tienen actividad opsonofagocítica. En formas de realización particulares, los anticuerpos inducidos, provocados o identificados son anticuerpos de reacción cruzada capaces de mediar la destrucción opsonofagocítica de cepas de O4 tanto glucosiladas como no glucosiladas de *E. coli*.

En ciertas formas de realización, los anticuerpos inducidos, provocados o identificados contra un O4 antígeno glucosilado de *E. coli* específicamente reconocen O4 antígenos polisacáridos no modificados y modificados por glucosa. En ciertas formas de realización, los anticuerpos inducidos, provocados o identificados contra un O4 antígeno

glucosilado de *E. coli* reconocen específicamente *E. coli* de serotipo O4. En ciertas formas de realización, los anticuerpos inducidos por un bioconjugado de un O4 antígeno glucosilado de *E. coli* se unen preferentemente al O4 antígeno glucosilado en comparación con el O4 antígeno no glucosilado.

5 Los anticuerpos inducidos por los bioconjugados y las composiciones descritas en la presente memoria pueden incluir moléculas de inmunoglobulina y porciones inmunológicamente activas de moléculas de inmunoglobulina, es decir, moléculas que contienen un sitio de unión a antígeno que se une específicamente a un O-antígeno polisacárido de *E. coli*, por ejemplo, un O4 antígeno polisacárido glucosilado.

10 Los anticuerpos inducidos, provocados o identificados usando los bioconjugados o las composiciones proporcionadas en este documento pueden usarse para monitorear la eficacia de una terapia y/o la progresión de la enfermedad. Cualquier sistema de inmunoensayo conocido en la técnica se puede utilizar para este propósito, incluyendo, pero sin limitaciones, sistemas de ensayo competitivos y no competitivos que usan técnicas tales como radioinmunoensayos, ELISA (ensayos inmunoabsorbentes ligados a enzimas), inmunoensayos basados en electroquimioluminiscencia (ECL), inmunoensayos en "sándwich", reacciones de precipitina, reacciones de precipitina de difusión en gel, ensayos de inmunodifusión, ensayos inmunoradiométricos, inmunoensayos fluorescentes, inmunoensayos de proteína A y ensayos de inmunoelectroforesis. Varios de estos ensayos, por ejemplo los inmunoensayos basados en ECL, se pueden hacer en formato multiplex, y por lo general se prefieren formatos de ensayo multiplex.

15 Los anticuerpos inducidos, provocados o identificados usando un bioconjugado de un O4 antígeno polisacárido glucosilado de *E. coli* se puede utilizar para detectar cepas O4 de *E. coli*, particularmente cepas O4 glucosiladas, por ejemplo, a partir de una pluralidad de cepas de *E. coli* y/o para diagnosticar una infección por una cepa O4 u O4 glucosilada de *E. coli*.

20 En otro aspecto, se proporciona aquí un método de vacunación de un sujeto contra *E. coli* (por ejemplo, *E. coli* patógenas extraintestinales, ExPEC), que comprende administrar al sujeto cualquiera de los bioconjugados de un O4 antígeno glucosilado de *E. coli* unido covalentemente a una proteína portadora descrita en la presente memoria, o una composición que comprende un bioconjugado de un O4 antígeno glucosilado de *E. coli* unido covalentemente a una proteína portadora, solo o en combinación además con otro O-antígeno de *E. coli* o glicoconjugados o bioconjugados del mismo. Un experto en la técnica entenderá que el sujeto será vacunado contra cepas de *E. coli* cuyos O antígenos o glicoconjugados o bioconjugados de los mismos están presentes en la composición administrada. Por ejemplo, la administración de una composición que comprende O1A, O2, O4 glucosilado, O6A y O25B antígenos polisacáridos se puede usar para vacunar a un sujeto contra *E. coli* serotipos O1A, O2, O4, O6A y O25B.

30 En ciertas formas de realización, la vacunación es para la prevención de una enfermedad invasiva por ExPEC (IED), por ejemplo, urosepsis, bacteriemia, sepsis, etc. En ciertas formas de realización, la vacunación es para prevenir o reducir la aparición o la gravedad de las infecciones del tracto urinario. En ciertas formas de realización, una IED puede ser adquirida en el hospital, por ejemplo, en pacientes sometidos a procedimientos o cirugías urogenitales y/o abdominales. En ciertas formas de realización, una IED puede ser asociada a la atención de salud, por ejemplo, en pacientes que reciben atención de salud para otra afección, por ejemplo a través de vías centrales, catéteres, etc., por ejemplo en un hospital, centro quirúrgico ambulatorio, centro para enfermedades renales terminales, centro de cuidados a largo plazo, etc. En ciertas formas de realización, la IED puede ser adquirida en la comunidad, por ejemplo, en un paciente que no estuvo recientemente expuesto a riesgos de salud.

35 En otro aspecto, se proporciona aquí un método para inducir una respuesta inmune contra *E. coli* (por ejemplo, ExPEC) en un sujeto, que comprende administrar al sujeto cualquiera de los bioconjugados de un O4 antígeno glucosilado de *E. coli* unido covalentemente a una proteína portadora descrita en este documento, o una composición que comprende un bioconjugado de un O4 antígeno glucosilado de *E. coli* unido covalentemente a una proteína portadora, solo o en combinación además con otro O-antígeno de *E. coli* o glicoconjugados o bioconjugados del mismo. En una forma de realización, el sujeto tiene una infección por *E. coli* (por ejemplo, ExPEC) en el momento de la administración. En una forma de realización preferida, el sujeto no tiene una infección por *E. coli* (por ejemplo, ExPEC) en el momento de la administración.

40 En ciertas formas de realización, las composiciones y los bioconjugados descritos en este documento pueden administrarse a un sujeto para inducir una respuesta inmunitaria que incluye la producción de anticuerpos, preferiblemente anticuerpos que tienen actividad opsonofagocítica. Tales anticuerpos se pueden aislar usando técnicas conocidas por el experto en la técnica (por ejemplo, cromatografía de inmovilización de afinidad, centrifugación, precipitación, etc.).

45 La capacidad de los bioconjugados y las composiciones descritas en la presente memoria para generar una respuesta inmune en un sujeto puede evaluarse usando cualquier enfoque conocido por los expertos en la técnica o descrito en este documento. En algunas formas de realización, la capacidad de un bioconjugado para generar una respuesta inmune en un sujeto puede evaluarse mediante la inmunización de un sujeto (por ejemplo, un ratón, rata, conejo o mono) o un conjunto de sujetos con un bioconjugado descrito en este documento y la inmunización de un sujeto adicional (por ejemplo, un ratón, rata, conejo o mono) o un conjunto de sujetos con un control (PBS). Los sujetos o conjuntos de sujetos posteriormente pueden ser desafiados con ExPEC, y se puede determinar la capacidad de

ExPEC para causar enfermedad (por ejemplo, UTI, bacteriemia u otra enfermedad) en los sujetos o conjuntos de sujetos. Los expertos en la técnica reconocerán que si el sujeto o conjunto de sujetos inmunizados con el control sufren la enfermedad posteriormente al desafío con el ExPEC pero el sujeto o el conjunto de sujetos inmunizados con un bioconjugado o composición del mismo descrito en este documento sufren menos o no sufren la enfermedad, entonces el bioconjugado es capaz de generar una respuesta inmune en un sujeto. La capacidad de un bioconjugado o composición del mismo descrita aquí para inducir antisuero que reacciona en forma cruzada con un O antígeno de ExPEC puede evaluarse mediante, por ejemplo, un inmunoensayo, tal como un ELISA (véase, por ejemplo, Van den Dobbelsteen *et al.*, 2016, *Vaccine*, 34: 4152–4160), o un inmunoensayo basado en ECL.

Por ejemplo, la capacidad de los bioconjugados descritos en este documento para generar una respuesta inmune en un sujeto puede evaluarse usando un ensayo bactericida de suero (SBA) o ensayo de destrucción opsonofagocítica (ensayo OPK, u OPKA), que representa un método establecido y aceptado que ha sido utilizado para obtener la aprobación de vacunas a base de glicoconjugados. Tales ensayos son bien conocidos en la técnica y, brevemente, comprenden las etapas de generación y aislamiento de anticuerpos contra un blanco de interés (por ejemplo, un O antígeno polisacárido, por ejemplo, O4 antígeno polisacárido glucosilado de *E. coli*) mediante la administración a un sujeto (por ejemplo, un ratón, rata, conejo o mono) de un compuesto que genera dichos anticuerpos. Posteriormente, la capacidad bactericida de los anticuerpos se puede evaluar mediante, por ejemplo, el cultivo de las bacterias en cuestión (por ejemplo, *E. coli* del serotipo correspondiente) en presencia de los anticuerpos y complemento y – dependiendo del ensayo– células neutrofílicas y ensayando la capacidad de los anticuerpos para mediar la destrucción y/o neutralización de las bacterias, por ejemplo, utilizando enfoques microbiológico estándar. A modo de ejemplo de ensayo OPK para vacunas de bioconjugados de *E. coli*, véase, por ejemplo, Abbanat *et al.*, 2017, *Clin. Vaccine Immunol.* 24: e00123–17. Un ensayo OPK puede realizarse en formato monoplex o multiplex, de los cuales generalmente se prefiere el formato multiplex (por ejemplo, el ensayo de múltiples serotipos al mismo tiempo). Un ensayo OPK multiplex se denomina a veces en el presente documento como ‘MOPA’.

En algunas formas de realización, los métodos descritos en el presente documento comprenden administrar una cantidad efectiva de bioconjugados de un O4 antígeno glucosilado de *E. coli* unido covalentemente a una proteína portadora descrita en el presente documento, o una composición que comprende un bioconjugado de un O4 antígeno glucosilado de *E. coli* unido covalentemente a una proteína portadora, solo o en combinación además con otro O–antígeno de *E. coli* o glicoconjugados o bioconjugados del mismo. En una forma de realización, una “cantidad eficaz” es una cantidad que vacuna a un sujeto contra *E. coli* (por ejemplo, ExPEC). En otra forma de realización, una “cantidad eficaz” es una cantidad que induce una respuesta inmune contra *E. coli* (por ejemplo, ExPEC) en un sujeto, tal como una respuesta inmune que incluye la producción de anticuerpos, preferiblemente anticuerpos que tienen actividad opsonofagocítica.

En formas de realización particulares, en donde una composición proporcionada en este documento comprende un bioconjugado de un O4 antígeno polisacárido glucosilado de *E. coli* y al menos un bioconjugado de un O25B antígeno polisacárido de *E. coli*, una cantidad eficaz del O25B antígeno polisacárido de *E. coli* es de aproximadamente 1.5 a 6 veces, por ejemplo de aproximadamente 2 a 4 veces mayor, tal como 1.5, 2, 3, 4, 5 o 6 veces mayor que la concentración de cualquiera de los otros bioconjugados presentes en la composición. En tales formas de realización, una cantidad eficaz del O25B antígeno polisacárido de *E. coli* es, por ejemplo, de aproximadamente 5 a 18 µg por administración, tal como 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18 µg por administración.

En ciertas formas de realización, un bioconjugado o una composición de acuerdo con la invención se administra a un sujeto una vez. En ciertas formas de realización, un bioconjugado o una composición de acuerdo con la invención se administra a un sujeto más de una vez, por ejemplo, en un régimen de cebado–refuerzo. En ciertas formas de realización, el tiempo entre dos administraciones es de al menos dos semanas, al menos un mes, al menos dos meses, al menos tres meses, al menos seis meses, al menos un año, al menos dos años, al menos cinco años, al menos diez años o al menos quince años. En los seres humanos, una respuesta inmune deseada típicamente puede ser generada por una única administración de un bioconjugado o una composición de acuerdo con la invención. En ciertas formas de realización, se proporciona una administración repetida después de, por ejemplo, diez años.

Células huésped

En el presente documento se proporcionan células huésped, por ejemplo, células huésped procariotas, capaces de producir O antígenos de *E. coli* y bioconjugados que comprenden tales O antígenos de *E. coli*. Las células huésped proporcionadas en la presente preferiblemente se modifican para comprender (por ejemplo, a través de la ingeniería genética) uno o más de los ácidos nucleicos que codifican la maquinaria de la célula huésped (por ejemplo, glicosiltransferasas) que se utiliza para producir O–antígenos polisacáridos de *E. coli* y/o bioconjugados de los mismos.

Cualquier célula huésped conocida por los expertos en la técnica se puede usar para producir los O antígenos polisacáridos de *E. coli* descritos en este documento (por ejemplo, O4 antígeno polisacárido glucosilado de *E. coli*) y bioconjugados que comprenden los O antígenos polisacáridos de *E. coli* descritos en este documento (por ejemplo, un bioconjugado de O4 antígeno polisacárido glucosilado de *E. coli*) incluyendo arqueobacterias, células huésped procariotas y células huésped eucariotas. En una forma de realización preferida, una célula huésped es una célula

huésped procariota. Las células huésped procariotas ejemplares para usar en la producción de los O antígenos polisacáridos de *E. coli* descritos aquí y bioconjugados que comprenden los O antígenos polisacáridos de *E. coli* descritos en este documento incluyen, pero sin limitaciones, especies de *Escherichia*, especies de *Shigella*, especies de *Klebsiella*, especies de *Xhantomonas*, especies de *Salmonella*, especies de *Yersinia*, especies de *Lactococcus*, especies de *Lactobacillus*, especies de *Pseudomonas*, especies de *Corynebacterium*, especies de *Streptomyces*, especies de *Streptococcus*, especies de *Staphylococcus*, especies de *Bacillus* y especies de *Clostridium*.

En una forma de realización específica, la célula huésped usada para producir los O antígenos polisacáridos de *E. coli* descritos en este documento y bioconjugados que comprenden los O antígenos polisacáridos de *E. coli* descritos aquí es una célula huésped procariota, y es preferiblemente *E. coli*.

En ciertas formas de realización, las células huésped utilizadas para producir los O antígenos polisacáridos de *E. coli* y bioconjugados descritos en este documento están diseñadas para comprender ácidos nucleicos heterólogos, por ejemplo, ácidos nucleicos heterólogos que comprenden grupos de genes *rfb* de un serotipo de O antígeno deseado, ácidos nucleicos heterólogos que codifican una o más proteínas portadoras y/o glicosiltransferasas. En una forma de realización específica, los genes *rfb* heterólogos, y/o ácidos nucleicos heterólogos que codifican proteínas implicadas en las vías de glicosilación (por ejemplo, vías de glicosilación procariotas y/o eucariotas) se pueden introducir en las células huésped descritas en este documento. Tales ácidos nucleicos pueden codificar proteínas, incluyendo, pero sin limitaciones, oligosacaril transferasas y/o glicosiltransferasas.

En este documento se describen las secuencias de diversos genes y grupos de genes que codifican glicosiltransferasas útiles en la preparación de células huésped recombinantes que pueden, por ejemplo, ser utilizadas para preparar O antígenos polisacáridos de *E. coli* y bioconjugados de los mismos. Los expertos en la técnica apreciarán que, debido a la degeneración del código genético, una proteína que tiene una secuencia de aminoácidos específica puede ser codificada por múltiples ácidos nucleicos diferentes. Por lo tanto, los expertos en la técnica entenderán que un ácido nucleico proporcionado en este documento puede ser alterado de manera tal que su secuencia difiera de una secuencia proporcionada en el presente documento, sin afectar la secuencia de aminoácidos de la proteína codificada por el ácido nucleico.

Se proporcionan en este documento células huésped (por ejemplo, células huésped recombinantes) para la producción de un bioconjugado de un O4 antígeno polisacárido glucosilado de *E. coli*, O1A antígeno polisacárido, O2 antígeno polisacárido, O6A antígeno polisacárido, O8 antígeno polisacárido, O15 antígeno polisacárido, O16 antígeno polisacárido, O18A antígeno polisacárido, O25B antígeno polisacárido u O75 antígeno polisacárido. Las células huésped proporcionadas en este documento comprenden ácidos nucleicos que codifican enzimas (por ejemplo, glicosiltransferasas) capaces de producir el O antígeno polisacárido de *E. coli*. Las células huésped proporcionadas en la presente pueden expresar en forma natural los ácidos nucleicos capaces de producir un O antígeno de interés, o las células huésped se pueden preparar para expresar dichos ácidos nucleicos. En ciertas formas de realización, los ácidos nucleicos son heterólogos para las células huésped y se introducen en las células huésped utilizando enfoques genéticos conocidos en la técnica. Por ejemplo, los ácidos nucleicos pueden introducirse en la célula huésped mediante manipulación genética (por ejemplo, el grupo de genes se expresa en un plásmido o plásmidos o se integra en el genoma de la célula huésped (véase, por ejemplo, las publicaciones de solicitudes de patente internacional WO 2014/037585, WO 2014/057109, WO 2015/052344).

En una forma de realización, se proporciona en este documento una célula huésped (por ejemplo, una célula huésped recombinante) capaz de producir un bioconjugado de un O4 antígeno polisacárido glucosilado de *E. coli* unido covalentemente a una proteína portadora. Tal célula huésped comprende, preferiblemente por ingeniería de una célula precursora, una secuencia de ácido nucleico que codifica un gen *gtrS*, que, al leer y entender de los inventores, fue identificado aquí por primera vez como que codifica una enzima ramificadora de polisacárido capaz de transferir glucosa para el O4 antígeno de *E. coli* (es decir, una glicosiltransferasa específica para el O4 antígeno polisacárido de *E. coli*), y en particular para L-Rha a través de un enlace α -1,3-glucosídico. Un ejemplo de una secuencia de aminoácidos de dicha enzima ramificadora se proporciona en la SEQ ID NO: 4. Otros ejemplos comprenden secuencias de aminoácidos que son al menos 80% idénticas a la misma. Los ejemplos de secuencias de ácido nucleico que codifican genes *gtrS* específicos para el O4 antígeno polisacárido de *E. coli* incluyen, pero sin limitación, SEQ ID NO: 5, o secuencias degeneradas de ácido nucleico que codifican SEQ ID NO: 4, o secuencias de ácido nucleico que codifican enzimas GtrS funcionales específicas de O4 que tienen al menos 80% de identidad con SEQ ID NO: 4.

En una forma de realización específica, una célula huésped (por ejemplo, una célula huésped recombinante) capaz de producir un bioconjugado de un O4 antígeno polisacárido glucosilado de *E. coli* unido covalentemente a una proteína portadora comprende una secuencia de nucleótidos que codifica una glicosil transferasa que tiene al menos 80% de identidad de secuencia con SEQ ID NO: 4, tal como aproximadamente 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 95%, 97%, 98%, 99% o 100% de identidad de secuencia con SEQ ID NO: 4. En vista de la redundancia en el código genético, un experto en la técnica puede hacer variantes del ácido nucleico que codifica las secuencias de aminoácidos de glicosil transferasas, por ejemplo, utilizando secuencias optimizadas por codón, si se desea.

En ciertas formas de realización, una célula huésped (por ejemplo, una célula huésped recombinante) capaz de producir un bioconjugado de un O4 antígeno polisacárido glucosilado de *E. coli* unido covalentemente a una proteína portadora, que comprende una secuencia de nucleótidos que codifica una glucosil transferasa (GtrS) que tiene al menos 80% de identidad de secuencia con SEQ ID NO: 4, además comprende una secuencia de nucleótidos que codifica una glucosa translocasa unida a bactoprenol (GtrA) que tiene al menos 80% de identidad de secuencia con SEQ ID NO: 7, y una secuencia de nucleótidos que codifica una bactoprenol glucosil transferasa (GtrB) que tiene al menos 80% de identidad de secuencia con SEQ ID NO: 8. En ciertas formas de realización, dichas secuencias de ácido nucleico codifican proteínas GtrA y GtrB que son al menos aproximadamente 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% o 100% idénticas a SEQ ID NOs: 7 y 8, respectivamente, y tienen actividad de glucosa translocasa unida a bactoprenol (SEQ ID NO: 7) y bactoprenol glucosil transferasa (SEQ ID NO: 8), respectivamente. En vista de la redundancia en el código genético, un experto en la técnica puede hacer variantes del nucleico que codifica las secuencias de aminoácidos de glucosa translocasas unidas a bactoprenol y de bactoprenol glucosil transferasas, por ejemplo, utilizando secuencias optimizadas por codón, si se desea.

Una célula huésped (por ejemplo, una célula huésped recombinante) capaz de producir un bioconjugado de un O4 antígeno polisacárido glucosilado de *E. coli* unido covalentemente a una proteína portadora proporcionada en el presente documento comprende además una secuencia de nucleótidos de un grupo de genes *rfb* para el O4 antígeno polisacárido de *E. coli*. Un ejemplo de un grupo de genes *rfb* útil para la producción del O4 antígeno polisacárido de *E. coli* se proporciona en el presente documento como SEQ ID NO: 9. Otro ejemplo se puede encontrar en el GenBank, locus AY568960. También se pueden usar secuencias de ácido nucleico degeneradas que codifican las mismas enzimas codificadas por esta secuencia, o secuencias que codifican enzimas que son al menos 80% idéntica, preferiblemente al menos 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% idénticas.

En una forma de realización específica, se proporciona en este documento una célula huésped (por ejemplo, una célula huésped recombinante, preferiblemente una célula huésped procariota recombinante, preferiblemente una célula huésped recombinante de *E. coli*) que produce O4 antígeno polisacárido glucosilado, en donde la célula huésped comprende *gtrS*, un grupo de genes *rfb* para el O4 antígeno polisacárido de *E. coli*, y el ácido nucleico que codifica una proteína portadora. Tales células huésped se pueden diseñar utilizando enfoques recombinantes para comprender uno o más plásmidos que comprenden el gen *gtrS*, el grupo de genes *rfb*, y/o ácido nucleico que codifica una proteína portadora, o para comprender algunos o todos los genes relevantes, tales como *gtrS*, el grupo *rfb* y/o el ácido nucleico que codifica la proteína portadora integrada en el genoma de la célula huésped. En ciertas formas de realización, los genes o grupos de genes se han integrado en el genoma de la célula huésped usando recombinación homóloga. Una ventaja de la integración de genes en el genoma de la célula huésped es la estabilidad en ausencia de selección de antibióticos.

En otra forma de realización específica, se proporciona en este documento una célula huésped (por ejemplo, una célula huésped recombinante, preferiblemente una célula huésped procariota recombinante) que produce O4 antígeno polisacárido glucosilado, en donde la célula huésped comprende GtrS (glucosiltransferasa), así como las enzimas codificadas por el grupo de *rfb* O4. En ciertas formas de realización, todas o algunas de las enzimas mencionadas anteriormente son heterólogos para la célula huésped.

En otras formas de realización específicas, se proporciona en este documento una célula huésped (por ejemplo, una célula huésped recombinante, preferiblemente una célula huésped procariota recombinante) que produce O4 antígeno polisacárido glucosilado de *E. coli*, preferiblemente un bioconjugado de O4 antígeno polisacárido glucosilado de *E. coli*, en donde la célula huésped comprende además una secuencia de nucleótidos que codifica una oligosacaril transferasa y/o una secuencia de nucleótidos que codifica una proteína portadora. En una forma de realización específica, la oligosacaril transferasa es heteróloga para la célula huésped. En otra forma de realización específica, la proteína portadora es heteróloga para la célula huésped. Preferiblemente, la célula huésped comprende una secuencia de nucleótidos heteróloga que codifica una glucosil transferasa que tiene al menos 80% de identidad de secuencia con SEQ ID NO: 4. En formas de realización preferidas, los genes *rfb* del grupo O4 son heterólogos para la célula huésped. Preferiblemente, la secuencia que codifica la enzima que es capaz de introducir la cadena lateral de glucosa ramificada en el O4 antígeno, es decir, el gen *gtrS* (que codifica una glucosil transferasa que tiene al menos 80% de identidad de secuencia con SEQ ID NO: 4) es heterólogo para la célula huésped. Un ácido nucleico es heterólogo para la célula huésped si la misma secuencia no está presente naturalmente en dicha célula huésped. El ácido nucleico heterólogo, por ejemplo, puede ser introducido en una célula parental por ingeniería genética, por ejemplo, por transformación (por ejemplo, transformación química o electroporación) y/o recombinación. En ciertas formas de realización, el ácido nucleico heterólogo, tal como un locus *rfb* deseado, secuencia codificadora de *gtrS*, secuencia codificadora de proteína portadora y/o secuencia codificadora de glicosiltransferasa, se integran en el genoma de la célula huésped, preferiblemente una célula huésped bacteriana, preferiblemente una célula huésped de *E. coli*. En formas de realización preferidas, el locus *rfb* endógeno y, si corresponde, la secuencia codificadora de *gtrS* han sido inactivados, preferentemente borrados del genoma de la célula huésped recombinante en comparación con un antecesor de la misma, y preferiblemente se sustituyen por el locus *rfb* heterólogo deseado y, si corresponde, una secuencia codificadora de *gtrS* deseada, respectivamente. En ciertas formas de realización, la célula huésped es una K-12 de *E. coli* (como ejemplo no limitativo, la cepa de *E. coli* W3110 es una cepa K-12) o una cepa B de *E. coli* (como ejemplo no limitativo, la cepa de *E. coli* BL21 es una cepa B), o cualquier otra cepa bien definida de *E. coli*, por

ejemplo, cepas de laboratorio o cepas de producción, en contraste con los aislados de tipo salvaje primarios. En formas de realización preferidas, la célula huésped se deriva de *E. coli* que no expresa O4 antígeno u O4 antígeno glucosilado, mediante la introducción en tal *E. coli* del locus *rfb* O4 y el gen *gtrS* que codifica una glucosil transferasa tiene al menos 80% identidad de secuencia con SEQ ID NO: 4. Las ventajas del uso de cepas bien caracterizadas, tales como *E. coli* K-12 o *E. coli* B, como precursores para las células huésped es la posibilidad de utilizar un proceso de producción similar para diferentes bioconjugados de O antígeno, ya que las características de la cepa de producción están bien definidas. A pesar de que los bioconjugados de diferentes O antígenos se comportan de manera diferente y los procesos de expresión se pueden optimizar según la cepa de producción, al menos el proceso básico para la producción de bioconjugados de O-antígeno será más predecible utilizando tales cepas precursoras bien definidas que cuando se utilizan cepas desconocidas tales como aislados de tipo salvaje como precursores para la producción de cepas huésped. De esta manera, la experiencia con la producción de los bioconjugados de O-antígeno de *E. coli* descritos anteriormente tales como los bioconjugados de O1A, O2, O6A y O25B como se describe, por ejemplo, en WO 2015/124769 y WO 2017/035181 se puede utilizar como base para diseñar la producción de otros bioconjugados de O-antígeno de *E. coli*. A diferencia de *gtrS*, los genes *gtrA* y *gtrB* no son serotipo-específicos, y en ciertas formas de realización son homólogos a la célula huésped (por ejemplo, la *E. coli* K12 cepa W3110 incluye genes *gtrA* y *gtrB* que son capaces de funcionar junto con el gen *gtrS* específico del O4-serotipo introducido recombinantemente que codifica una glucosil transferasa de SEQ ID NO: 4 o una glucosil transferasa que es al menos 80% idéntica a la misma, reemplazando el gen *gtrS* endógeno). En otras formas de realización, uno o ambos genes *gtrA* y *gtrB* (que codifican proteínas GtrA y GtrB que son al menos aproximadamente 80% idénticas a SEQ ID NOs: 7 y 8, respectivamente, y que tienen actividad de glucosa translocasa unida a bactoprenol y bactoprenol glucosil transferasa, respectivamente, también se introducen de forma recombinante en la célula huésped, por ejemplo, en caso de que la célula huésped no tenga genes *gtrA* y/ o *gtrB* endógenos.

También se proporcionan en el presente documento células huésped (por ejemplo, células huésped recombinantes) capaces de producir un bioconjugado de un O1A, O2, O6A, O8, O15, O16, O18A, O25B u O75 antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora. Tales células huésped (por ejemplo, células huésped recombinantes) comprenden la secuencia de nucleótidos de un grupo de genes *rfb* específico del O-antígeno polisacárido. Los grupos de genes *rfb* pueden ser aislados a partir de cepas de *E. coli* de tipo salvaje, y se combinan con ácidos nucleicos que codifican una oligosacaril transferasa (por ejemplo, PglB) y una proteína portadora (por ejemplo, EPA) dentro de una célula huésped para obtener una célula huésped recombinante que produce el O-antígeno de *E. coli* de interés o bioconjugado del mismo. Por ejemplo, tales células huésped pueden ser diseñadas utilizando enfoques recombinantes para comprender uno o más plásmidos que comprenden el grupo de genes *rfb*, oligosacaril transferasa (por ejemplo, PglB) y una proteína portadora (por ejemplo, EPA) utilizando una tecnología de bioconjugación, tal como la descrita en el documento WO 2014/037585, WO 2009/104074 y WO 2009/089396. Preferiblemente, las células huésped comprenden los grupos de genes *rfb* integrados en su genoma. Los ácidos nucleicos que codifican oligosacaril transferasa, la proteína portadora y, cuando corresponda, el gen *gtrS* aplicable están también, en ciertas formas de realización, integrados en el genoma de la célula huésped. Los genes *gtrA* y *gtrB* heterólogos u homólogos están, en ciertas formas de realización, también integrados en el genoma de la célula huésped.

La preparación de bioconjugados para O1A, O2, O6A y O25B antígenos ha sido descrita en detalle en WO 2015/124769 y WO 2017/035181. Los grupos de genes ejemplares para cada O antígeno de *E. coli* (loci *rfb*) se han descrito en Iguchi A, *et al.*, DNA Research, 2014, 1-7 (doi: 10.1093/dnares/dsu043), y en DebRoy C, *et al.*, PLoS One. 2016, 11(1):e0147434 (doi: 10.1371/journal.pone.0147434; corrección en: Plos One. 2016, 11(4):e0154551, doi: 10.1371/journal.pone.0154551). Las secuencias de ácido nucleico para los grupos *rfb* y las secuencias de aminoácidos para las proteínas codificadas en los mismos también se pueden encontrar en bases de datos públicas, tales como GenBank. Las secuencias ejemplares para grupos *rfb* que pueden ser utilizadas en cepas de producción para bioconjugados con antígenos polisacáridos de los serotipos descritos en este documento también se proporcionan en SEQ ID NOs: 9 y 11-19. Así, para cada uno de los bioconjugados deseados mencionados anteriormente, el grupo *rfb* respectivo puede ser introducido en una célula huésped, para obtener células huésped con el grupo *rfb* específico para el O-antígeno deseado, que además contiene ácido nucleico que codifica oligosacaril transferasa y la proteína portadora. Por las razones indicadas anteriormente, preferiblemente las células huésped son células huésped recombinantes, y, preferiblemente, se derivan de cepas con características relativamente bien conocidas, tales como cepas de *E. coli* de laboratorio o de producción, por ejemplo, *E. coli* K12 o *E. coli* BL21, etc. Preferiblemente, los grupos *rfb* son heterólogos para la célula huésped, por ejemplo se introducen en una célula precursora de la célula huésped, y preferiblemente se integran en el genoma de la misma. Preferiblemente, un grupo de genes *rfb* original, si estaba presente en una célula precursora, ha sido sustituido por el grupo de genes *rfb* para el O-antígeno de interés en la célula huésped, para permitir la producción de bioconjugado del O-antígeno de interés. Preferiblemente, la oligosacariltransferasa es heteróloga para la célula huésped, y en ciertas formas de realización el ácido nucleico que codifica tal oligosacariltransferasa está integrado en el genoma de la célula huésped.

Cualquiera de las células huésped proporcionadas en la presente (por ejemplo, las células huésped recombinantes, preferiblemente células huésped procariotas recombinantes) comprenden ácidos nucleicos que codifican enzimas adicionales activas en la N-glicosilación de proteínas, por ejemplo, la célula huésped proporcionada en este

documento puede comprender además un ácido nucleico que codifica una oligosacaril transferasa o ácidos nucleicos que codifican otras glicosiltransferasas.

Las células huésped proporcionados en este documento comprenden un ácido nucleico que codifica una oligosacaril transferasa. Las oligosacaril transferasas transfieren oligosacáridos ligados a lípidos a residuos de asparagina de cadenas polipeptídicas nacientes que comprenden un motivo de consenso de *N*-glicosilación. El ácido nucleico que codifica una oligosacaril transferasa puede ser nativo a la célula huésped, o se puede introducir en la célula huésped utilizando enfoques genéticos. En formas de realización preferidas, la oligosacaril transferasa es heteróloga para la célula huésped. *E. coli* no comprende naturalmente una oligosacaril transferasa, y por lo tanto si se utiliza *E. coli* como célula huésped para la producción de bioconjugados, una oligosacaril transferasa heteróloga está comprendida en tal célula huésped, por ejemplo, después de la introducción por ingeniería genética. La oligosacaril transferasa puede ser de cualquier fuente conocida en la técnica a la vista de la presente divulgación.

En ciertas formas de realización, una alternativa a una oligosacaril transferasa con actividad de *N*-glicosiltransferasa, tal como una *O*-glicosiltransferasa, por ejemplo, como ejemplo no limitativo PglL, se puede utilizar conjuntamente con su propia secuencia de consenso de glicosilación diferente en la proteína portadora, como se describe, por ejemplo, en el documento WO 2016/82597. Otras glicosiltransferasas, tales como *O*-glicosiltransferasas, pueden por lo tanto también utilizarse como una oligosacariltransferasa de acuerdo con la invención.

En ciertas formas de realización preferidas, la oligosacaril transferasa es una oligosacaril transferasa de *Campylobacter*. Por ejemplo, en una forma de realización, la oligosacaril transferasa es una oligosacaril transferasa de *Campylobacter jejuni* (es decir, *pglB*; véase, por ejemplo, Wacker *et al.*, 2002, *Science* 298:1790–1793; véase también, por ejemplo, NCBI Gene ID: 3231775, UniProt nº de acceso O86154). En otra forma de realización, la oligosacaril transferasa es una oligosacaril transferasa de *Campylobacter lari* (véase, por ejemplo, NCBI Gene ID: 7410986).

En formas de realización específicas, la oligosacaril transferasa es PglB oligosacaril transferasa de *Campylobacter jejuni*, incluyendo la proteína natural (de tipo salvaje) o cualquier variante de la misma, tal como las descritas en las Publicaciones de Solicitud de Patente Internacional WO 2016/107818 y WO 2016/107819. PglB puede transferir oligosacáridos unidos a lípidos a residuos de asparagina en las secuencias de consenso SEQ ID NO: 1 y SEQ ID NO: 2. En formas de realización particulares, la PglB oligosacaril transferasa comprende la SEQ ID NO: 6, o una variante de la misma. En ciertas formas de realización, una o más secuencias endógenas de consenso de glicosilación en un PglB de tipo salvaje se han mutado para evitar la autoglicosilación de PglB, por ejemplo, SEQ ID NO: 6 que comprende la mutación N534Q. Los ejemplos de variantes de PglB oligosacaril transferasas adecuadas para usar en las células huésped recombinantes proporcionadas en este documento incluyen la PglB oligosacaril transferasa de SEQ ID NO: 6 que comprende al menos una mutación seleccionada del grupo que consiste de N311V, K482R, D483H, A669V, Y77H, S80R, Q287P y K289R. En una forma de realización particular, una variante de PglB oligosacaril transferasa tiene SEQ ID NO: 6 que comprende la mutación N311V. En otra forma de realización particular, una variante de PglB oligosacaril transferasa tiene SEQ ID NO: 6 que comprende las mutaciones Y77H y N311V. En otra forma de realización particular, una variante de PglB oligosacaril transferasa tiene SEQ ID NO: 6 que comprende las mutaciones N311V, K482R, D483H y A669V. En otra forma de realización particular, una variante de PglB oligosacaril transferasa tiene SEQ ID NO: 6 que comprende las mutaciones Y77H, S80R, Q287P, K289R y N311V. Se encontró y se describe en este documento que ciertas variantes de PglB oligosacaril transferasas tienen rendimientos sorprendentemente mejores en la producción de bioconjugados de *O*-antígenos de *E. coli* de serotipos específicos. La variante mejorada u óptima de PglB para un *O*-antígeno de *E. coli* dado no era previsible. Por lo tanto, en ciertos aspectos, la invención también proporciona métodos para producir bioconjugados de *O*-antígenos de *E. coli* específicos, utilizando variantes de PglB específicas como la oligosacaril transferasa. Otras variantes de PglB que son al menos 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% idénticas a SEQ ID NO: 6 y todavía tienen actividad de oligosacaril transferasa, que preferiblemente tienen uno o más de los aminoácidos específicos en las posiciones indicadas divulgadas en combinación en el presente documento (por ejemplo, 77Y, 80S, 287Q, 289K, 311N, 482K, 483D, 669A; o 311V; o 311V, 482R, 483H, 669V; o 77H, 80R, 287P, 289R, 311V; o 77H, 311V; etc.) también se pueden utilizar para la producción de bioconjugados.

En una forma de realización específica, una célula huésped (por ejemplo, una célula huésped recombinante) capaz de producir un bioconjugado de un O4 antígeno polisacárido glucosilado de *E. coli* unido covalentemente a una proteína portadora comprende además una secuencia de nucleótidos que codifica PglB oligosacaril transferasa de *Campylobacter jejuni* que tiene la secuencia de aminoácidos de SEQ ID NO: 6, o preferiblemente SEQ ID NO: 6 que comprende la mutación N311V, o más preferiblemente SEQ ID NO: 6 que comprende las mutaciones Y77H y N311V.

En otras formas de realización específicas, una célula huésped (por ejemplo, una célula huésped recombinante) capaz de producir un bioconjugado de un O1A, O6A u O15 antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora comprende además una secuencia de nucleótidos que codifica PglB oligosacaril transferasa de *Campylobacter jejuni* que tiene la secuencia de aminoácidos de SEQ ID NO: 6, o preferiblemente SEQ ID NO: 6 que comprende las mutaciones N311V, K482R, D483H y A669V.

En una forma de realización específica, una célula huésped (por ejemplo, una célula huésped recombinante) capaz de producir un bioconjugado de un O16 antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora comprende además una secuencia de nucleótidos que codifica PglB oligosacaril transferasa de *Campylobacter jejuni* que tiene la secuencia de aminoácidos de SEQ ID NO: 6, o preferiblemente SEQ ID NO: 6 que comprende las mutaciones Y77H, S80R, Q287P, K289R y N311V.

En una forma de realización específica, una célula huésped (por ejemplo, una célula huésped recombinante) capaz de producir un bioconjugado de un O75 antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora comprende además una secuencia de nucleótidos que codifica PglB oligosacaril transferasa de *Campylobacter jejuni* que tiene la secuencia de aminoácidos de SEQ ID NO: 6, o preferiblemente SEQ ID NO: 6 que comprende la mutación N311V.

En una forma de realización específica, una célula huésped (por ejemplo, una célula huésped recombinante) capaz de producir un bioconjugado de un O8, O18A, O25B u O2 antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora comprende además una secuencia de nucleótidos que codifica PglB oligosacaril transferasa de *Campylobacter jejuni* que tiene la secuencia de aminoácido de SEQ ID NO: 6, preferiblemente en donde SEQ ID NO: 6 no comprende mutaciones de aminoácidos en las posiciones 77, 80, 287, 289, 311, 482, 483 y 669.

En algunas formas de realización, cualquiera de las células huésped proporcionadas en este documento comprenden un ácido nucleico que codifica una proteína portadora, por ejemplo, una proteína a la que el O-antígeno polisacárido producido por la maquinaria de glicosilación de la célula huésped se puede unir para formar un bioconjugado. La célula huésped puede comprender un ácido nucleico que codifica cualquier proteína portadora conocida por los expertos en la técnica en vista de la presente divulgación, incluyendo, pero sin limitaciones, Exotoxina A desintoxicada de *P. aeruginosa* (EPA), flagelina de *E. coli* (FliC), CRM197, proteína de unión a maltosa (MBP), toxoide diftérico, toxoide tetánico, hemolisina A desintoxicada de *S. aureus*, factor de aglutinación A, factor de aglutinación B, enterotoxina termolábil de *E. coli*, variantes desintoxicadas de enterotoxina termolábil de *E. coli*, toxina del cólera subunidad B (CTB), toxina del cólera, variantes desintoxicadas de la toxina del cólera, proteína Sat de *E. coli*, el dominio pasajero de la proteína Sat de *E. coli*, neumolisina de *Streptococcus pneumoniae*, hemocianina de lapa californiana (KLH), PcrV de *P. aeruginosa*, proteína de membrana externa de *Neisseria meningitidis* (OMPC) y proteína D de *Haemophilus influenzae* no tipificable.

En formas de realización preferidas, una célula huésped comprende además un ácido nucleico que codifica Exotoxina A desintoxicada de *P. aeruginosa* (EPA). Preferiblemente, la proteína portadora EPA comprende 1–10 sitios de glicosilación, preferiblemente 2 a 4 sitios de glicosilación, lo más preferiblemente 4 sitios de glicosilación, tal como 1–10, preferiblemente 2–4, y más preferiblemente 4 sitios de glicosilación que comprenden, cada uno, una secuencia de consenso de glicosilación que tiene la secuencia de aminoácidos de SEQ ID NO: 1, y más preferiblemente tiene la secuencia de aminoácidos de SEQ ID NO: 2. En una forma de realización específica, una célula huésped comprende además un ácido nucleico que codifica la proteína portadora EPA–4 que comprende la SEQ ID NO: 3.

En ciertas formas de realización, las proteínas portadoras utilizadas en la generación de los bioconjugados por las células huésped descritas en este documento comprenden una “etiqueta”, es decir, una secuencia de aminoácidos que permite el aislamiento y/o la identificación de la proteína portadora. Por ejemplo, la adición de una etiqueta a una proteína portadora puede ser útil en la purificación de esa proteína y, por lo tanto, la purificación de vacunas conjugadas que comprenden la proteína portadora etiquetada. Las etiquetas a modo de ejemplo que pueden usarse en este documento incluyen, sin limitación, etiquetas de histidina (HIS) (por ejemplo, etiqueta de hexa-histidina o etiqueta de 6XHis), FLAG-TAG y etiquetas de HA. En ciertas formas de realización, las etiquetas utilizadas en el presente documento son eliminables, por ejemplo, pueden ser eliminadas por agentes químicos o por medios enzimáticos, una vez que ya no son necesarias, por ejemplo, después de que la proteína ha sido purificada. En otras formas de realización, la proteína portadora no comprende una etiqueta.

En ciertas formas de realización, las proteínas portadoras descritas en el presente documento comprenden una secuencia de señal que dirige a la proteína portadora al espacio periplásmico de la célula huésped que expresa la proteína portadora. En una forma de realización específica, la secuencia de señal es de DsbA de *E. coli*, porina A de membrana externa de *E. coli* (OmpA), proteína de unión a maltosa de *E. coli* (MalE), pectato liasa de *Erwinia carotovora* (PelB), FlgI, NikA, o endoxilanasas de *Bacillus* sp. (XynA), enterotoxina LTIIb de *E. coli* termolábil, endoxilanasas XynA de *Bacillus*, o flagelina de *E. coli* (FlgI). En una forma de realización, la secuencia de señal comprende SEQ ID NO: 10. Una secuencia de señal se puede escindir después de la translocación de la proteína al periplasma y puede por lo tanto ya no estar presente en la proteína portadora final de un bioconjugado.

En ciertas formas de realización, se pueden introducir modificaciones adicionales (por ejemplo, usando técnicas recombinantes) en las células huésped descritas en el presente documento. Por ejemplo, los ácidos nucleicos de la célula huésped (por ejemplo, genes) que codifican proteínas que forman parte de una vía de glicosilación que posiblemente compite o interfiere (por ejemplo, compite o interfiere con uno o más genes heterólogos que participan en la glicosilación que se introducen de manera recombinante en la célula huésped) pueden ser eliminados o modificados en el entorno de la célula huésped (genoma) de una manera que los haga inactivos / disfuncionales (es

decir, los ácidos nucleicos de la célula huésped que se eliminan / modifican no codifican una proteína funcional). En ciertas formas de realización, cuando los ácidos nucleicos se eliminan del genoma de las células huésped proporcionadas en este documento son reemplazados por una secuencia deseable, por ejemplo, una secuencia que es útil para la producción de un O antígeno polisacárido o un bioconjugado del mismo.

5 Los genes o grupos de genes ejemplares que se pueden eliminar en las células huésped (y, en algunos casos, reemplazar por otras secuencias de ácido nucleico deseadas) incluyen genes o grupos de genes de células huésped implicadas en la biosíntesis de glicolípidos, tales como *waaL* (véase, por ejemplo, Feldman *et al.*, 2005, *PNAS USA* 102:3016–3021), el grupo de biosíntesis de núcleo de lípido A (*waa*), el grupo de galactosa (*gal*), el grupo de arabinosa (*ara*), el grupo de ácido colónico (*wc*), el grupo de polisacáridos capsulares, genes de la biosíntesis de undecaprenol-
10 p (por ejemplo, *uppS*, *uppP*), genes de reciclaje de und-P, enzimas metabólicas implicadas en la biosíntesis de azúcares activados por nucleótidos, el grupo de antígeno común de enterobacterias (*eca*), y grupos de modificación de O antígenos prófagos como el grupo *gtrABS* o regiones de los mismos. En una forma de realización específica, las células huésped descritas en el presente documento se modifican de modo que no produzcan ningún O antígeno polisacárido distinto de un O antígeno polisacárido deseado, por ejemplo, O4 antígeno polisacárido glucosilado.

15 En una forma de realización específica, el gen de *waaL* se elimina o se inactiva funcionalmente a partir del genoma de una célula huésped (por ejemplo, una célula huésped recombinante) proporcionada en el presente documento. Los términos “*waaL*” y “gen *waaL*” se refieren al gen de ligasa del O antígeno que codifica una enzima unida a membrana con un sitio activo situado en el periplasma. La enzima codificada transfiere O antígeno unido a undecaprenilfosfato (UPP) al núcleo de lípido A, formando lipopolisacárido. La delección o interrupción del gen *waaL* endógeno (por ejemplo,
20 cepas $\Delta waaL$) interrumpe la transferencia del O-antígeno al lípido A, y en su lugar puede potenciar la transferencia del O-antígeno a otra biomolécula, tal como una proteína portadora.

En otra forma de realización específica, uno o más del gen *waaL*, gen *gtrA*, gen *gtrB*, gen *gtrS* y el grupo de genes *rfb* se suprime o se inactiva funcionalmente del genoma original de una célula huésped procarionta proporcionada en este documento.

25 En una forma de realización, una célula huésped usada aquí es *E. coli* que produce un bioconjugado de O4 antígeno polisacárido glucosilado, en donde el gen *waaL* se elimina o se inactiva funcionalmente del genoma de la célula huésped, y se inserta un gen *gtrS* específico de O4 antígeno polisacárido de *E. coli*. En ciertas formas de realización para las cepas de producción de bioconjugados de O4 O-antígeno glucosilado, un gen *gtrS* que codifica una glucosil transferasa que tiene al menos 80% de identidad de secuencia con la SEQ ID NO: 4 se inserta en el lugar de un gen
30 *gtrS* de la cepa parental, con el fin de reemplazar el gen *gtrS* en esa cepa parental por el responsable de la glucosilación del antígeno O4. Un ejemplo de tal cepa parental es *E. coli* K-12 cepa W3110. Los genes *gtrA* y *gtrB* pueden ser homólogos para la cepa parental, o, alternativamente, uno o ambos de estos genes pueden ser heterólogos para la cepa parental. Típicamente, y a diferencia del gen *gtrS*, estos genes *gtrA* y *gtrB* no son específicos para la estructura del O-antígeno.

35 También se proporcionan en este documento métodos de preparación de células huésped recombinantes. Las células huésped recombinantes producidas por los métodos descritos en este documento pueden usarse para producir bioconjugados de O antígenos de *E. coli*. Los métodos comprenden introducir una o más moléculas de ácido nucleico recombinantes en una célula para producir la célula huésped recombinante. Típicamente, las moléculas de ácido nucleico recombinantes son heterólogas. Cualquier método conocido en la técnica en vista de la presente divulgación
40 se puede utilizar para introducir moléculas de ácido nucleico recombinantes en una célula huésped. Los ácidos nucleicos recombinantes pueden introducirse en las células huésped descritas en este documento usando cualquiera de los métodos conocidos por los expertos en la técnica, por ejemplo, electroporación, transformación química, choque térmico, transformación natural, transducción de fagos y conjugación. En formas de realización específicas, los ácidos nucleicos recombinantes se introducen en las células huésped descritas en este documento usando un plásmido. Por
45 ejemplo, los ácidos nucleicos heterólogos pueden ser expresados en las células huésped por un plásmido (por ejemplo, un vector de expresión). En otra forma de realización específica, los ácidos nucleicos heterólogos se introducen en las células huésped descritas en este documento usando el método de inserción en el genoma como se describe, por ejemplo, en la publicación de solicitud de patente internacional WO 2014/037585, WO 2014/057109 o WO 2015/052344.

50 En una forma de realización, un método para preparar una célula huésped recombinante para la producción de un bioconjugado de un O4 antígeno polisacárido glucosilado de *E. coli* unido covalentemente a una proteína portadora comprende introducir una o más moléculas de ácido nucleico recombinante en una célula, preferiblemente una célula de *E. coli*, para producir la célula huésped recombinante. En tales formas de realización, las moléculas de ácido nucleico recombinante introducidas en la célula incluyen (i) una secuencia de nucleótidos de un grupo de genes *rfb*
55 para el O4 antígeno polisacárido de *E. coli*; (ii) una secuencia de nucleótidos que codifica una glucosil transferasa que tiene al menos 80% de identidad de secuencia con SEQ ID NO: 4, en donde la glucosil transferasa es capaz de modificar el O4 antígeno polisacárido de *E. coli* para producir el O4 antígeno polisacárido glucosilado de *E. coli*; (iii) una secuencia de nucleótidos que codifica una proteína portadora; y (iv) una secuencia de nucleótidos que codifica una oligosacaril transferasa capaz de unir covalentemente el O4 antígeno polisacárido glucosilado de *E. coli* a la

5 proteína portadora para producir el bioconjugado. En formas de realización preferidas, la secuencia de nucleótidos que codifica una glucosil transferasa que tiene al menos 80% de identidad de secuencia con SEQ ID NO: 4 reemplaza el gen *gtrS* endógeno. La eliminación del gen *gtrS* endógeno tiene la ventaja de que no interferirá con la generación de la estructura del O4 antígeno polisacárido glucosilado. En ciertas formas de realización, la secuencia de nucleótidos del grupo de genes *rfb* para el O4 antígeno polisacárido de *E. coli* reemplaza el grupo de genes *rfb* endógeno de la cepa parental que se utiliza para hacer la célula huésped recombinante. Si la célula todavía no codifica genes *gtrA* y/o *gtrB*, las secuencias de nucleótidos que codifican una translocasa (*gtrA*) y una glicosiltransferasa (*gtrB*), teniendo al menos 80% de identidad con SEQ ID NOs: 7 y 8, respectivamente, se pueden introducir en la célula. Si la célula ya codifica genes *gtrA* y *gtrB* (tal como, por ejemplo, el caso en *E. coli* K-12 cepa W3110), no hay necesidad de introducir o cambiar estos genes.

10 En una forma de realización específica, la glucosil transferasa (específica para *gtrS* para añadir una rama de glucosa al O4 antígeno) tiene la SEQ ID NO: 4.

15 En una forma de realización específica, la oligosacaril transferasa es PglB de *C. jejuni*. En tal forma de realización, la oligosacaril transferasa comprende la secuencia de aminoácidos de SEQ ID NO: 6. En otra de tales formas de realización, la oligosacaril transferasa comprende la secuencia de aminoácidos de SEQ ID NO: 6 que comprende la mutación N311V. En otra de tales formas de realización, la oligosacaril transferasa comprende la secuencia de aminoácidos de SEQ ID NO: 6 que comprende las mutaciones Y77H y N311V.

20 En otra forma de realización específica, la proteína portadora comprende al menos un sitio de glicosilación que comprende una secuencia de consenso de glicosilación que tiene la SEQ ID NO: 1, preferiblemente la SEQ ID NO: 2. En otra forma de realización específica, la proteína portadora es EPA, preferiblemente EPA-4, tal como EPA-4 que comprende la SEQ ID NO: 3.

25 Las cepas de *E. coli* que se utilizan de forma rutinaria en biología molecular como una herramienta y como un organismo modelo pueden, por ejemplo, ser utilizadas como progenitores para células huésped en ciertas formas de realización de acuerdo con la invención. Los ejemplos no limitantes incluyen *E. coli* cepas K12 (por ejemplo, tales como W1485, W2637, W3110, MG1655, DH1, DH5 α , DH10, etc.), cepas B (por ejemplo, BL-21, REL606, etc.), cepas C o cepas W. En una forma de realización particular, la cepa huésped se deriva de la cepa parental W3110. Esta cepa puede, por ejemplo, ser obtenida de la *E. coli* Genetic Stock Center at Yale. Para obtener más información sobre *E. coli*, véase, por ejemplo, Ecoliwiki.net.

Métodos para producir conjugados y bioconjugados

30 También se proporcionan métodos de producción de glicoconjugados de los O antígenos polisacáridos de *E. coli* descritos en este documento. Los glicoconjugados, incluyendo bioconjugados, se pueden preparar *in vitro* o *in vivo*, por ejemplo, usando las células huésped recombinantes descritos en este documento para la producción.

35 En algunas formas de realización, los glicoconjugados se pueden preparar por síntesis química, es decir, se pueden preparar fuera de las células huésped (*in vitro*). Por ejemplo, un O antígeno polisacárido de *E. coli* se puede conjugar con proteínas portadoras utilizando métodos conocidos por los expertos en la técnica, incluyendo por medio de la utilización de grupos reactivos de activación en el polisacárido / oligosacárido así como la proteína portadora. Véase, por ejemplo, Pawlowski *et al.*, 2000, *Vaccine* 18:1873-1885; y Robbins, *et al.*, 2009, *Proc Natl Acad Sci USA* 106:7974-7978). Tales enfoques comprenden la extracción de polisacáridos / oligosacáridos antigénicos a partir de células huésped, la purificación de los polisacáridos / oligosacáridos, la activación química de los polisacáridos / oligosacáridos, y la conjugación de los polisacáridos / oligosacáridos a una proteína portadora.

40 En algunas formas de realización, las células huésped descritas en el presente documento pueden usarse para producir bioconjugados que comprenden un O antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora. Los métodos para producir tales bioconjugados utilizando células huésped son conocidos en la técnica. Véase, por ejemplo, WO 2003/074687 y WO 2006/119987. Dichos métodos comprenden cultivar cualquiera de las células huésped recombinantes descritas en este documento en condiciones para la producción del bioconjugado. Los bioconjugados pueden aislarse, separarse y/o purificarse a partir de células huésped recombinantes usando cualquier método conocido en la técnica a la vista de la presente divulgación. Por ejemplo, los bioconjugados se pueden purificar mediante cualquier método conocido en la técnica para la purificación de una proteína, por ejemplo, mediante cromatografía (por ejemplo, cromatografía de intercambio iónico, intercambio aniónico, de afinidad y en columna de tamaño), centrifugación, solubilidad diferencial o mediante cualquier otra técnica estándar para la purificación de proteínas. Véase, por ejemplo, los métodos descritos en el documento WO 2009/104074. Además, los bioconjugados pueden fusionarse con secuencias de polipéptidos heterólogos para facilitar la purificación. Las condiciones reales utilizadas para purificar un bioconjugado particular dependerán, en parte, de factores tales como la carga neta, la hidrofobicidad y/o la hidrofiliidad del bioconjugado, y serán evidentes para los expertos en la materia. La preparación de bioconjugados para O1A, O2, O6A y O25B, así como las composiciones de vacunas que comprenden estos, se han descrito, por ejemplo, en el documento WO 2015/124769 y en el documento WO 2017/035181.

También se proporcionan bioconjugados producidos por los métodos descritos en este documento, es decir, usando las células huésped recombinantes descritas en este documento.

En algunas formas de realización, un método para preparar un bioconjugado de un O-antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora comprende: (i) proporcionar una célula huésped recombinante que comprende (a) secuencia de nucleótidos de un grupo de genes *rfb* para el O-antígeno polisacárido; (b) una secuencia de nucleótidos que codifica una proteína portadora, preferiblemente EPA, que comprende al menos un sitio de glicosilación que comprende una secuencia de consenso de glicosilación que tiene la SEQ ID NO: 1, preferiblemente la SEQ ID NO: 2, y más preferiblemente que comprende cuatro sitios de glicosilación, cada uno de los cuales comprende una secuencia de consenso de glicosilación que tiene la SEQ ID NO: 2; y (c) una secuencia de nucleótidos que codifica una oligosacaril transferasa, por ejemplo PglB oligosacaril transferasa o una variante de la misma.

En ciertas formas de realización, los O-antígenos polisacáridos de *E. coli* producidos usando las células huésped recombinantes descritas en este documento están unidos covalentemente a la proteína portadora en una relación particular de polisacárido a proteína en peso (p/p). Esta relación de cantidad de O-antígeno polisacárido en peso unido covalentemente a la proteína portadora en peso se conoce como "relación de glicano / proteína" o "relación de polisacárido / proteína" o "relación de PS / proteína". En algunas formas de realización, el O-antígeno polisacárido se une covalentemente a la proteína portadora en una relación de polisacárido a proteína (p/p) de aproximadamente 1:20 a 20:1, preferiblemente de 1:10 a 10:1, más preferiblemente de 1:3 a 3:1. En ciertas formas de realización no limitantes para los bioconjugados descritos en el presente documento, la relación de glicano / proteína es de aproximadamente 0.1 a 0.5, tal como 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45 o 0.5. En tales formas de realización, la relación en peso del O-antígeno polisacárido: proteína es de aproximadamente 1:10 a 1:2, tal como 1:10: 1:9: 1:8, 1:7, 1:6, 1:5, 1:4, 1:3 o 1:2, dependiendo del serotipo del O-antígeno particular. En ciertas formas de realización, la relación de glicano / proteína es de aproximadamente 0.15 a aproximadamente 0.45. En general, se prefiere una mayor relación glicano / proteína de O-antígeno polisacárido a proteína portadora, debido a que una alta cantidad de proteína portadora puede llevar a la interferencia inmunológica en algunos casos. También, una mayor relación glicano / proteína ayudaría a conseguir suficiente O-antígeno polisacárido dosificado en forma de bioconjugado, mientras se mantiene la cantidad de proteína portadora relativamente baja, lo que es especialmente beneficioso para las composiciones multivalentes donde múltiples serotipos deben ser cubiertos por la composición, por ejemplo, composiciones que comprenden bioconjugados de al menos 4 O-antígenos diferentes, al menos 5 O-antígenos diferentes, al menos 6 O-antígenos diferentes, al menos 7 O-antígenos diferentes, al menos 8 O-antígenos diferentes, al menos 9 O-antígenos diferentes, al menos 10 O-antígenos diferentes, etc.

Una relación glicano / proteína de un conjugado de acuerdo con la invención se puede determinar mediante la determinación de la cantidad de proteína y la cantidad de glicano. cantidad de proteína puede determinarse por medición de la absorbancia UV a 280 nm (A280). La cantidad de glicano puede determinarse sobre la base de una cromatografía iónica con detección amperométrica por pulsos (IC-PAD) de un azúcar en la unidad de repetición (por ejemplo, de Man para O8 en la Tabla 1, y de GlcNAc para los otros glicanos en la Tabla 1), después de lo cual la información estructural de la unidad de repetición se puede utilizar para calcular la cantidad de glicano total (por ejemplo, la unidad de repetición de O1A tiene una masa molar de 845 Da y un mol de tal unidad de una repetición contiene un mol de GlcNAc, lo que permite el cálculo de la cantidad total de glicanos cuando la cantidad de GlcNAc ha sido determinada por IC-PAD).

En algunas formas de realización, un bioconjugado de un O25B antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora producida usando una célula huésped recombinante de acuerdo con las células y los métodos descritos en el presente documento tiene un determinado grado de acetilación en la posición 2 del azúcar L-Rh. El grado de O-acetilación del O25B antígeno polisacárido en un bioconjugado es preferiblemente al menos 30%, preferiblemente al menos 50%, tal como al menos 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% o 100%.

Del mismo modo, el grado de O-acetilación de un O16 antígeno polisacárido de *E. coli* en un bioconjugado es preferiblemente al menos 30%, preferiblemente al menos 50%, tal como al menos 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% o 100%.

En formas de realización específicas, un método de preparación de un bioconjugado de un O-antígeno polisacárido comprende proporcionar una célula huésped recombinante que comprende la secuencia de ácido nucleico que codifica una enzima oligosacaril transferasa particular, especialmente una PglB oligosacaril transferasa o una variante de la misma, dependiendo del bioconjugado de O-antígeno polisacárido a ser producido. La variante de enzima oligosacaril transferasa particular puede afectar el rendimiento del bioconjugado producido por la célula huésped. Típicamente, se prefiere un mayor rendimiento, ya que el rendimiento tendrá un impacto en los costes para producir un bioconjugado específico, lo que es especialmente importante para composiciones multivalentes que comprenden varios bioconjugados diferentes. En algunas formas de realización, el método comprende además aislar el bioconjugado de la célula huésped recombinante.

En una forma de realización particular, cuando el O-antígeno es O1A, O6A u O15 antígeno polisacárido, la PglB oligosacaril transferasa comprende las mutaciones de aminoácidos de N311V, K482R, D483H y A669V, en donde las

mutaciones de aminoácidos son en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6.

5 En otra forma de realización particular, cuando el O-antígeno es O4 antígeno polisacárido glucosilado, la PglB oligosacaril transferasa comprende la mutación de aminoácido N311V, o las mutaciones de aminoácidos de Y77H y N311V, en donde las mutaciones de aminoácidos son en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6.

10 En otra forma de realización particular, cuando el O-antígeno es O16 antígeno polisacárido, la PglB oligosacaril transferasa comprende las mutaciones de aminoácidos de Y77H, S80R, Q287P, K289R y N311V, en donde las mutaciones de aminoácidos son en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6.

En otra forma de realización particular, cuando el O-antígeno es O75 antígeno polisacárido, la PglB oligosacaril transferasa comprende la mutación de aminoácido de N311V, en donde las mutaciones de aminoácidos son en relación con la PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6.

15 En otra forma de realización particular, cuando el O-antígeno es O8, O18A, O25B u O2 antígeno polisacárido, la PglB oligosacaril transferasa comprende la secuencia de aminoácidos de SEQ ID NO: 6, en donde SEQ ID NO: 6 no comprende mutaciones de aminoácidos en las posiciones 77, 80, 287, 289, 311, 482, 483 y 669. En ciertas formas de realización de la misma, la PglB oligosacaril transferasa comprende la secuencia de aminoácidos de SEQ ID NO: 6.

20 En algunas formas de realización, la proteína portadora se selecciona del grupo que consiste en Exotoxina A desintoxicada de *P. aeruginosa* (EPA), flagelina de *E. coli* (FliC), CRM197, proteína de unión a maltosa (MBP), toxoide diftérico, toxoide tetánico, hemolisina A desintoxicada de *S. aureus*, factor de aglutinación A, factor de aglutinación B, enterotoxina termolábil de *E. coli*, variantes desintoxicadas de enterotoxina termolábil de *E. coli*, toxina del cólera subunidad B (CTB), toxina del cólera, variantes desintoxicadas de la toxina del cólera, proteína Sat de *E. coli*, el dominio pasajero de la proteína Sat de *E. coli*, neumolisina de *Streptococcus pneumoniae*, hemocianina de lapa californiana (KLH), PcrV de *P. aeruginosa*, proteína de membrana externa de *Neisseria meningitidis* (OMPC), y proteína D de *Haemophilus influenzae* no tipificable.

25 En ciertas formas de realización, la proteína portadora es exotoxina A desintoxicada de *Pseudomonas aeruginosa* (EPA). Preferiblemente, la proteína portadora EPA comprende 1-10, preferiblemente 2-4, más preferiblemente 4 sitios de glicosilación. Preferiblemente, cada sitio de glicosilación comprende una secuencia de consenso de glicosilación que tiene la secuencia de aminoácidos de SEQ ID NO: 2. En una forma de realización específica, una célula huésped comprende un ácido nucleico que codifica la proteína portadora EPA-4 que comprende la SEQ ID NO: 3.

30 En ciertas formas de realización, la célula huésped recombinante es una célula de *E. coli*, por ejemplo, una *E. coli* cepa K-12, como la cepa W3110.

35 También se proporcionan en este documento bioconjugados de O-antígenos polisacáridos producidos usando células huésped recombinantes que codifican las enzimas oligosacaril transferasas según los apareamientos de O-antígeno / PglB oligosacaril transferasa indicados anteriormente. También se proporcionan composiciones que comprenden tales bioconjugados. En ciertas formas de realización, una composición comprende al menos 2, preferiblemente al menos 3, más preferiblemente al menos 5, aún más preferiblemente al menos 7 de tales bioconjugados.

40 En algunas formas de realización, los bioconjugados de O-antígenos polisacáridos producidos por células huésped recombinantes que codifican las enzimas oligosacaril transferasas según los apareamientos de O-antígeno / PglB oligosacaril transferasa indicados anteriormente preferiblemente tienen uno o más de los atributos preferidos descritos en el presente documento, por ejemplo, una relación glicano / proteína y/o una cantidad o relación de proteína portadora multiglicosilada.

FORMAS DE REALIZACIÓN

45 La Forma de Realización 1 es un método de preparación de un bioconjugado de un O_x antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora, comprendiendo el método:

(i) proporcionar una célula huésped recombinante que comprende:

a. una secuencia de nucleótidos de un grupo de genes *rfb* para el O_x-antígeno polisacárido;

50 b. una secuencia de nucleótidos que codifica la proteína portadora que comprende al menos un sitio de glicosilación que comprende una secuencia de consenso de glicosilación que tiene la SEQ ID NO: 1, preferiblemente que tiene la SEQ ID NO: 2; y

c. una secuencia de nucleótidos que codifica una oligosacaril transferasa PglBy; y

ES 2 951 442 T3

(ii) cultivar la célula huésped recombinante en condiciones para la producción del bioconjugado;

en donde:

cuando el O_x-antígeno es O1A antígeno polisacárido, el PglB_y comprende las mutaciones de aminoácidos de N311V, K482R, D483H y A669V;

- 5 cuando el O_x-antígeno es O4 antígeno polisacárido glucosilado, el PglB_y comprende la mutación de aminoácidos N311V o las mutaciones de aminoácidos Y77H y N311V, y la célula huésped recombinante comprende además una secuencia que codifica una glucosiltransferasa GtrS que tiene al menos 80% de identidad con SEQ ID NO: 4 y es capaz de modificar un O4 antígeno polisacárido de *E. coli* mediante la adición de glucosa para producir el O4 antígeno polisacárido glucosilado de *E. coli*, y secuencias de nucleótidos que codifican una translocasa GtrA y una glicosiltransferasa GtrB que tienen al menos 80% de identidad de secuencia con las SEQ ID NOs: 7 y 8, respectivamente, en donde la translocasa es capaz de translocar glucosa unida a bactoprenol y la glicosiltransferasa es capaz de glucosilar bactoprenol;

cuando el O_x-antígeno es O6A antígeno polisacárido, el PglB_y comprende las mutaciones de aminoácidos de N311V, K482R, D483H y A669V;

- 15 cuando el O_x-antígeno es O8 antígeno polisacárido, el PglB_y no comprende mutaciones de aminoácidos en las posiciones 77, 80, 287, 289, 311, 482, 483 y 669;

cuando el O_x-antígeno es O15 antígeno polisacárido, el PglB_y comprende las mutaciones de aminoácidos de N311V, K482R, D483H y A669V;

- 20 cuando el O_x-antígeno es O16 antígeno polisacárido, el PglB_y comprende las mutaciones de aminoácidos de Y77H, S80R, Q287P, K289R y N311V;

cuando el O_x-antígeno es O18A antígeno polisacárido, el PglB_y no comprende mutaciones de aminoácidos en las posiciones 77, 80, 287, 289, 311, 482, 483 y 669;

y cuando el O_x-antígeno es O75 antígeno polisacárido, el PglB_y comprende la mutación de aminoácido de N311V;

- 25 en donde en cada caso las mutaciones de aminoácidos son en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6, y

en donde los O1A, O4 glucosilado, O6A, O8, O15, O16, O18A y O75 antígenos polisacáridos tienen las estructuras de las Fórmulas (O1A), (O4-Glc+), (O6A), (O8), (O15), (O16), (O18A) y (O75), respectivamente, tal como se muestra en la Tabla 1, y cada n es independientemente un número entero de 1 a 100, preferiblemente de 3 a 50, por ejemplo de 5 a 40, por ejemplo de 7 a 25, por ejemplo de 10 a 20.

- 30 La Forma de Realización 2 es el método de la forma de realización 1, en donde el O_x-antígeno es O1A antígeno polisacárido, y el PglB_y comprende las mutaciones de aminoácidos de N311V, K482R, D483H y A669V en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6.

- 35 La Forma de Realización 3 es el método de la forma de realización 1, en donde el O_x-antígeno es O4 antígeno polisacárido glucosilado, y el PglB_y comprende la mutación de aminoácido N311V o las mutaciones de aminoácidos Y77H y N311V en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de la SEQ ID NO: 6.

La Forma de Realización 4 es el método de la forma de realización 3, en donde la célula huésped recombinante comprende además una secuencia que codifica un GtrS que tiene la secuencia de aminoácidos de SEQ ID NO: 4, y secuencias de nucleótidos que codifican un GtrA y un GtrB que tienen las secuencias de aminoácidos de SEQ ID NOs: 7 y 8, respectivamente.

- 40 La Forma de Realización 5 es el método de la forma de realización 1, en donde el O_x-antígeno es O6A antígeno polisacárido, y el PglB_y comprende las mutaciones de aminoácidos de N311V, K482R, D483H y A669V en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6.

- 45 La Forma de Realización 6 es el método de la forma de realización 1, en donde el O_x-antígeno es O8 antígeno polisacárido, y el PglB_y no comprende mutaciones de aminoácidos en las posiciones 77, 80, 287, 289, 311, 482, 483 y 669 en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6.

La Forma de Realización 7 es el método de la forma de realización 1, en donde el O_x-antígeno es O15 antígeno polisacárido, y el PglB_y comprende las mutaciones de aminoácidos de N311V, K482R, D483H y A669V en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6.

- La Forma de Realización 8 es el método de la forma de realización 1, en donde el O_x-antígeno es O16 antígeno polisacárido, y el PglB_y comprende las mutaciones de aminoácidos de Y77H, S80R, Q287P, K289R y N311V en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6.
- 5 La Forma de Realización 9 es el método de la forma de realización 1, en donde el O_x-antígeno es O18A antígeno polisacárido, y el PglB_y no comprende mutaciones de aminoácidos en las posiciones 77, 80, 287, 289, 311, 482, 483 y 669 en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6.
- La Forma de Realización 10 es el método de la forma de realización 1, en donde el O_x-antígeno es O75 antígeno polisacárido, y el PglB_y comprende la mutación de aminoácidos de N311V en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6.
- 10 La Forma de Realización 11 es un método de preparación de un bioconjugado de un O_x antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora, comprendiendo el método:
- (i) proporcionar una célula huésped recombinante que comprende:
- (a) una secuencia de nucleótidos de un grupo de genes *rfb* para el O_x-antígeno polisacárido;
- (b) una secuencia de nucleótidos que codifica la proteína portadora que comprende al menos un sitio de glicosilación que comprende una secuencia de consenso de glicosilación que tiene la SEQ ID NO: 1, preferiblemente que tiene la SEQ ID NO: 2; y
- 15 (c) una secuencia de nucleótidos que codifica una oligosacaril transferasa PglB_y; y
- (ii) cultivar la célula huésped recombinante en condiciones para la producción del bioconjugado,
- en donde el PglB_y comprende la mutación de aminoácidos N311V en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6,
- 20 en donde el O_x-antígeno es O1A antígeno polisacárido, O4 antígeno polisacárido glucosilado, O6A antígeno polisacárido, O15 antígeno polisacárido, O16 antígeno polisacárido o O75 antígeno polisacárido, y cuando el O_x-antígeno es O4 antígeno polisacárido glucosilado, la célula huésped recombinante comprende además una secuencia que codifica una glucosiltransferasa GtrS que tiene al menos 80% de identidad con SEQ ID NO: 4 y que es capaz de modificar un O4 antígeno polisacárido de *E. coli* mediante la adición de glucosa para producir el O4 antígeno polisacárido glucosilado de *E. coli*, y secuencias de nucleótidos que codifican una translocasa GtrA y una glicosiltransferasa GtrB que tienen al menos 80% de identidad de secuencia con las SEQ ID NOs: 7 y 8, respectivamente, en donde la translocasa es capaz de translocar glucosa ligada a bactoprenol y la glicosiltransferasa es capaz de glucosilar bactoprenol, y
- 25 en donde los O1A, O4 glucosilado, O6A, O15, O16 y, O75 antígenos polisacáridos tienen las estructuras de Fórmulas (O1A), (O4- Glc+), (O6A), (O15), (O16) y (O75), respectivamente, como se muestra en la Tabla 1, y cada n es independientemente un número entero de 1 a 100, preferiblemente de 3 a 50, por ejemplo de 5 a 40, por ejemplo de 7 a 25, por ejemplo de 10 a 20.
- 30 La Forma de Realización 12 es el método de cualquiera de las formas de realización 1 a 11, que comprende además aislar el bioconjugado a partir de la célula huésped recombinante.
- 35 La Forma de Realización 13 es el método de cualquiera de las formas de realización 1 a 12, en donde la proteína portadora se selecciona del grupo que consiste en Exotoxina A desintoxicada de *P. aeruginosa* (EPA), flagelina de *E. coli* (FliC), CRM197, proteína de unión a maltosa (MBP), toxoide diftérico, toxoide tetánico, hemolisina A desintoxicada de *S. aureus*, factor de aglutinación A, factor de aglutinación B, enterotoxina termolábil de *E. coli*, variantes desintoxicadas de enterotoxina termolábil de *E. coli*, toxina del cólera subunidad B (CTB), toxina del cólera, variantes desintoxicadas de la toxina del cólera, proteína Sat de *E. coli*, el dominio pasajero de la proteína Sat de *E. coli*, neumolisina de *Streptococcus pneumoniae*, hemocianina de lapa californiana (KLH), PcrV de *P. aeruginosa*, proteína de membrana externa de *Neisseria meningitidis* (OMPC) y proteína D de *Haemophilus influenzae* no tipificable.
- 40 La Forma de Realización 14 es el método de la forma de realización 13, en donde la proteína portadora es exotoxina A desintoxicada de *Pseudomonas aeruginosa* (EPA).
- 45 La Forma de Realización 15 es el método de la forma de realización 14, en donde la proteína portadora EPA comprende 1-10, preferiblemente 2-4, más preferiblemente 4, de los sitios de glicosilación.
- La Forma de Realización 16 es el método de la forma de realización 15, en donde cada sitio de glicosilación comprende una secuencia de consenso de glicosilación que tiene la SEQ ID NO: 2.

ES 2 951 442 T3

La Forma de Realización 17 es el método de la forma de realización 16, en donde la proteína portadora EPA comprende la SEQ ID NO: 3.

La Forma de Realización 18 es el método de cualquiera de las formas de realización 1–17, en donde la célula huésped recombinante es una célula de *E. coli*, por ejemplo, una *E. coli* cepa K–12, tal como la cepa W3110.

- 5 La Forma de Realización 19 es un bioconjugado producido por el método de cualquiera de las formas de realización 1–18.

La Forma de Realización 20 es una composición que comprende un bioconjugado de la forma de realización 19.

La Forma de Realización 21 es una composición que comprende al menos 2, preferiblemente al menos 3, más preferiblemente al menos 5, aún más preferiblemente al menos 7 bioconjugados de la forma de realización 19.

- 10 La Forma de Realización 22 es una composición de la forma de realización 20 o 21, que comprende un bioconjugado de O4 antígeno polisacárido glucosilado de *E. coli* unido covalentemente a una proteína portadora, en donde el O4 antígeno polisacárido glucosilado tiene la estructura de Fórmula (O4–Glc+) tal como se muestra en la Tabla 1, y n es un número entero de 1 a 100, preferiblemente de 3 a 50, por ejemplo de 5 a 40, por ejemplo de 7 a 25, por ejemplo de 10 a 20.

- 15 La Forma de Realización 23 es una composición de cualquiera de las formas de realización 20 a 22, que comprende además al menos un bioconjugado de O25B antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora, en donde el O25 antígeno polisacárido tiene la estructura de Fórmula (O25B) tal como se muestra en Tabla 1, y n es un número entero de 1 a 100, preferiblemente de 3 a 50, por ejemplo de 5 a 40, por ejemplo de 7 a 25, por ejemplo de 10 a 20.

- 20 La Forma de Realización 24 es una composición de cualquiera de las formas de realización 20 a 23, que comprende además al menos un bioconjugado de O2 antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora, en donde el O2 antígeno polisacárido tiene la estructura de Fórmula (O2) tal como se muestra en Tabla 1, y n es un número entero de 1 a 100, preferiblemente de 3 a 50, por ejemplo de 5 a 40, por ejemplo de 7 a 25, por ejemplo de 10 a 20.

- 25 La Forma de Realización 25 es una composición de cualquiera de las formas de realización 20 a 24, que comprende: (i) un bioconjugado de O1A antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora, (ii) un bioconjugado de O2 antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora, (iii) un bioconjugado de O4 antígeno polisacárido glucosilado de *E. coli* unido covalentemente a una proteína portadora, (iv) un bioconjugado de O6A antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora, (v) un bioconjugado de O8 antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora, (vi) un bioconjugado de O15 antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora, (vii) un bioconjugado de O16 antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora, (viii) un bioconjugado de O25B antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora, y (ix) un bioconjugado de O75 antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora, en donde los
30 O1A, O2, O4 glucosilado, O6A, O8, O15, O16, O25B y O75 antígenos polisacáridos tienen las estructuras de Fórmulas (O1A), (O2), (O4–Glc+), (O6A), (O8), (O15), (O16), (O25B) y (O75), respectivamente, tal como se muestra en la Tabla 1, y cada n es independientemente un número entero de 1 a 100, preferiblemente de 3 a 50, por ejemplo de 5 a 40, por ejemplo de 7 a 25, por ejemplo de 10 a 20.

- 40 La Forma de Realización 26 es una composición de la forma de realización 25, que comprende además: (x) un bioconjugado de O18A antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora, en donde el O18A antígeno polisacárido tiene la estructura de Fórmula (O18A) tal como se muestra en la Tabla 1, y n es un número entero de 1 a 100, preferiblemente de 3 a 50, por ejemplo de 5 a 40, por ejemplo de 7 a 25, por ejemplo de 10 a 20.

La Forma de Realización 27 es una composición de cualquiera de las formas de realización 20 a 26, en donde la composición es una composición inmunogénica.

- 45 La Forma de Realización 28 es un método de vacunación de un sujeto contra *E. coli*, en particular *E. coli* patogénica extraintestinal (ExPEC), que comprende administrar al sujeto el bioconjugado de la forma de realización 19, o la composición o la composición inmunogénica de cualquiera de las formas de realización 20 a 27.

- 50 La Forma de Realización 29 es el bioconjugado de la forma de realización 19, o la composición o la composición inmunogénica de cualquiera de las formas de realización 20 a 27 para usar en la vacunación contra *E. coli* patogénica extraintestinal (ExPEC).

La Forma de Realización 30 es una célula huésped recombinante para la preparación de un bioconjugado de un O_x antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora, donde la célula huésped recombinante comprende:

ES 2 951 442 T3

(a) una secuencia de nucleótidos de un grupo de genes *rfb* para el O_x-antígeno polisacárido;

(b) una secuencia de nucleótidos que codifica la proteína portadora que comprende al menos un sitio de glicosilación que comprende una secuencia de consenso de glicosilación que tiene la SEQ ID NO: 1, preferiblemente que tiene la SEQ ID NO: 2; y

5 (c) una secuencia de nucleótidos que codifica una oligosacaril transferasa PglB_y;

en donde:

cuando el O_x-antígeno es O1A antígeno polisacárido, el PglB_y comprende las mutaciones de aminoácidos de N311V, K482R, D483H y A669V;

10 cuando el O_x-antígeno es O4 antígeno polisacárido glucosilado, el PglB_y comprende la mutación de aminoácido N311V o las mutaciones de aminoácidos Y77H y N311V, y la célula huésped recombinante comprende además una secuencia que codifica una glucosiltransferasa GtrS que tienen al menos 80% de identidad con SEQ ID NO: 4 y que es capaz de modificar un O4 antígeno polisacárido de *E. coli* mediante la adición de glucosa para producir el O4 antígeno polisacárido glucosilado de *E. coli*, y secuencias de nucleótidos que codifican una translocasa GtrA y una glicosiltransferasa GtrB que tienen al menos 80% de identidad de secuencia con SEQ ID NOs: 7 y 8, respectivamente,
15 en donde la translocasa es capaz de translocar glucosa unida a bactoprenol y la glicosiltransferasa es capaz de glucosilar bactoprenol;

cuando el O_x-antígeno es O6A antígeno polisacárido, el PglB_y comprende las mutaciones de aminoácidos de N311V, K482R, D483H y A669V;

20 cuando el O_x-antígeno es O8 antígeno polisacárido, el PglB_y no comprende mutaciones de aminoácidos en las posiciones 77, 80, 287, 289, 311, 482, 483 y 669;

cuando el O_x-antígeno es O15 antígeno polisacárido, el PglB_y comprende las mutaciones de aminoácidos de N311V, K482R, D483H y A669V;

cuando el O_x-antígeno es O16 antígeno polisacárido, el PglB_y comprende las mutaciones de aminoácidos de Y77H, S80R, Q287P, K289R y N311V;

25 cuando el O_x-antígeno es O18A antígeno polisacárido, el PglB_y no comprende mutaciones de aminoácidos en las posiciones 77, 80, 287, 289, 311, 482, 483 y 669; y

cuando el O_x-antígeno es O75 antígeno polisacárido, el PglB_y comprende la mutación de aminoácido de N311V;

en donde, en cada caso, las mutaciones de aminoácidos son en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6, y

30 en donde los O1A, O4 glucosilado, O6A, O8, O15, O16, O18A y O75 antígenos polisacáridos tienen las estructuras de Fórmulas (O1A), (O4-Glc+), (O6A), (O8), (O15), (O16), (O18A) y (O75), respectivamente, tal como se muestra en la Tabla 1, y cada n es independientemente un número entero de 1 a 100, preferiblemente de 3 a 50, por ejemplo de 5 a 40, por ejemplo de 7 a 25, por ejemplo de 10 a 20.

35 La Forma de Realización 31 es la célula huésped recombinante de la forma de realización 30, en donde el O_x-antígeno es O1A antígeno polisacárido, y el PglB_y comprende las mutaciones de aminoácidos de N311V, K482R, D483H y A669V en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6.

40 La Forma de Realización 32 es la célula huésped recombinante de la forma de realización 30, en donde el O_x-antígeno es O4 antígeno polisacárido glucosilado, y el PglB_y comprende la mutación de aminoácido N311V o las mutaciones de aminoácido Y77H y N311V en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6.

La Forma de Realización 33 es la célula huésped recombinante de la forma de realización 32, en donde la célula huésped recombinante comprende además una secuencia que codifica un GtrS que tiene la secuencia de aminoácidos de SEQ ID NO: 4, y secuencias de nucleótidos que codifican un GtrA y un GtrB que tiene las secuencias de aminoácidos de SEQ ID NOs: 7 y 8, respectivamente.

45 La Forma de Realización 34 es la célula huésped recombinante de la forma de realización 30, en donde el O_x-antígeno es O6A antígeno polisacárido, y el PglB_y comprende las mutaciones de aminoácidos de N311V, K482R, D483H y A669V en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6.

- La Forma de Realización 35 es la célula huésped recombinante de la forma de realización 30, en donde el O_x-antígeno es O8 antígeno polisacárido, y el PglB_y no comprende mutaciones de aminoácidos en las posiciones 77, 80, 287, 289, 311, 482, 483 y 669 en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6.
- 5 La Forma de Realización 36 es la célula huésped recombinante de la forma de realización 30, en donde el O_x-antígeno es O15 antígeno polisacárido, y el PglB_y comprende las mutaciones de aminoácidos de N311V, K482R, D483H y A669V en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6.
- La Forma de Realización 37 es la célula huésped recombinante de la forma de realización 30, en donde el O_x-antígeno es O16 antígeno polisacárido, y el PglB_y comprende las mutaciones de aminoácidos de Y77H, S80R, Q287P, K289R y N311V en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6.
- 10 La Forma de Realización 38 es la célula huésped recombinante de la forma de realización 30, en donde el O_x-antígeno es O18A antígeno polisacárido, y el PglB_y no comprende mutaciones de aminoácidos en las posiciones 77, 80, 287, 289, 311, 482, 483 y 669 en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6.
- 15 La Forma de Realización 39 es la célula huésped recombinante de la forma de realización 30, en donde el O_x-antígeno es O75 antígeno polisacárido, y el PglB_y comprende la mutación de aminoácido de N311V en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6.
- 20 La Forma de Realización 40 es la célula huésped recombinante de cualquiera de las formas de realización 30 a 39, en donde la proteína portadora se selecciona del grupo que consiste en Exotoxina A desintoxicada de *P. aeruginosa* (EPA), flagelina de *E. coli* (FliC), CRM197, proteína de unión a maltosa (MBP), toxoide diftérico, toxoide tetánico, hemolisina A desintoxicada de *S. aureus*, factor de aglutinación A, factor de aglutinación B, enterotoxina termolábil de *E. coli*, variantes desintoxicadas de enterotoxina termolábil de *E. coli*, toxina del cólera subunidad B (CTB), toxina del cólera, variantes desintoxicadas de la toxina del cólera, proteína Sat de *E. coli*, el dominio pasajero de la proteína Sat de *E. coli*, neumolisina de *Streptococcus pneumoniae*, hemocianina de lapa californiana (KLH), PcrV de *P. aeruginosa*, proteína de membrana externa de *Neisseria meningitidis* (OMPC) y proteína D de *Haemophilus influenzae* no tipificable.
- 25 La Forma de Realización 41 es la célula huésped recombinante de cualquiera de las formas de realización 30–40, en donde la proteína portadora es exotoxina A desintoxicada de *Pseudomonas aeruginosa* (EPA).
- La Forma de Realización 42 es la célula huésped recombinante de la forma de realización 41, en donde la proteína portadora EPA comprende 1–10, preferiblemente 2–4, más preferiblemente 4, de los sitios de glicosilación.
- 30 La Forma de Realización 43 es la célula huésped recombinante de la forma de realización 42, en donde cada sitio de glicosilación comprende una secuencia de consenso de glicosilación que tiene la SEQ ID NO: 2.
- La Forma de Realización 44 es la célula huésped recombinante de la forma de realización 43, en donde la proteína portadora EPA comprende la SEQ ID NO: 3.
- 35 La Forma de Realización 45 es la célula huésped recombinante de cualquiera de las formas de realización 30 a 44, en donde la célula huésped recombinante es una célula de *E. coli*, por ejemplo, una *E. coli* cepa K–12, tal como la cepa W3110.
- La Forma de Realización 46 es un bioconjugado de acuerdo con la forma de realización 19, en donde el bioconjugado es un bioconjugado de O4 antígeno polisacárido glucosilado de *E. coli* unido covalentemente a una proteína portadora.
- 40 La Forma de Realización 47 es un bioconjugado de acuerdo con la forma de realización 46, en donde la proteína portadora es una proteína portadora EPA que comprende la SEQ ID NO: 3.
- La Forma de Realización 48 es un bioconjugado de acuerdo con la forma de realización 46 o 47, en donde el O4 antígeno polisacárido glucosilado tiene las estructuras de Fórmula (O4–Glc+) tal como se muestra en la Tabla 1, y n es un número entero de 5 a 40.
- 45 La Forma de Realización 49 es una composición que comprende un bioconjugado de acuerdo con cualquiera de las formas de realización 46–48.
- La Forma de Realización 50 es una composición de acuerdo con la forma de realización 49, que comprende además uno o más conjugados que comprenden, cada uno, un antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora.
- 50 La Forma de Realización 51 es una composición de acuerdo con la forma de realización 50, en donde el uno o más conjugados comprenden antígeno polisacárido de *E. coli* de uno o más de los siguientes serotipos de *E. coli*: O1A, O2, O6A, O8, O15, O16, O18A, O25B y O75, en donde los O1A, O2, O6A, O8, O15, O16, O25B y O75 antígenos

polisacáridos tienen las estructuras de Fórmulas (O1A), (O2), (O6A), (O8), (O15), (O16), (O18A), (O25B) y (O75), respectivamente, tal como se muestra en la Tabla 1, y cada n es independientemente un número entero de 1 a 100, preferiblemente de 3 a 50, por ejemplo de 5 a 40, por ejemplo de 7 a 25, por ejemplo de 10 a 20.

5 La Forma de Realización 52 es una composición de acuerdo con la forma de realización 51, que comprende conjugados de serotipos de *E. coli*: O1A, O2, O6A, O8, O15, O16, O18A, O25B y O75.

La Forma de Realización 53 es una composición de acuerdo con la forma de realización 52, en donde cada uno de los conjugados es un bioconjugado.

EJEMPLOS

10 Los siguientes ejemplos de la invención son para ilustrar adicionalmente la naturaleza de la invención. Se debe entender que los siguientes ejemplos no limitan la invención, y el alcance de la invención estará determinado por las reivindicaciones adjuntas.

Ejemplo 1: Datos epidemiológicos de las infecciones por *E. coli*

15 Para determinar la distribución de O-serotipo de la *E. coli* causante de bacteriemia, se realizaron estudios mundiales de vigilancia. Entre 2011 y 2017, se recolectaron más de 3200 aislados de *E. coli* del torrente sanguíneo de pacientes ≥ 60 años de edad hospitalizados en países de América del Norte, Europa, la región de Asia-Pacífico, y América del Sur. Cada cepa se analizó para determinar el serotipo del O antígeno mediante técnicas de aglutinación clásicas y O-genotipificación basada en secuencia. Véase la Tabla 2.

20 Se analizaron muestras de sangre humana aislada para determinar la identidad de los agentes patógenos en la misma y sus patrones de resistencia a antibióticos. Se obtuvieron aislados de *E. coli* de las muestras luego del análisis. La identidad de *E. coli* se verificó por MALDI-TOF MS. Se realizó un análisis adicional de los aislados de *E. coli* usando un ensayo de aglutinación basado en antisueros para determinar su serotipo de O-antígeno (DebRoy *et al.* (2011) Animal health research reviews / Conference of Research Workers in Animal Diseases 12, 169–185). Los aislados no tipificables por el método de aglutinación se analizaron adicionalmente por secuenciación de genoma completo, seguido de O-genotipificación basada en secuencias de genes *wzy* y *wzx* específicas de O-serotipo.

25 **Tabla 2:** Distribución de los O-serotipos de *E. coli* asociados con bacteriemia más comunes a partir de una colección de 3217 aislados de sangre recogidos a nivel mundial entre 2011 y 2017, en base a O-serotipificación por aglutinación más O-genotipificación de los aislados no tipificables por aglutinación. Los sujetos fueron hospitalizados en los siguientes países: EE.UU., Canadá, Argentina, Brasil, Reino Unido, Alemania, España, Italia, Países Bajos, Francia, Japón, Tailandia, Corea del Sur y Australia.

O-serotipo	Prevalencia n (%)
O25	737 (22.9%)
O2	268 (8.3%)
O6	261 (8.1%)
O1	255 (7.9%)
O75	145 (4.5%)
O15	110 (3.4%)
O8	104 (3.2%)
O16	103 (3.2%)
O4	96 (3.0%)
O18	91 (2.8%)

30 La estratificación de la ubicación geográfica en el conjunto global de *E. coli* asociada con bacteriemia mostró una prevalencia de los principales 10 O-serotipos independientemente de la ubicación, lo que sugiere que estos son los O-serotipos predominantes globalmente asociados con *E. coli* causante de bacteriemia.

35 En el conjunto global de aislados de *E. coli* resistente a múltiples fármacos y asociada a bacteriemia (n=345), es decir, aquellas cepas que son resistentes a al menos tres clases de fármacos antimicrobianos clínicamente relevantes, la prevalencia de los 10 O-serotipos principales es del 75.4%.

A partir de toda la información del análisis de epidemiología en conjunto, los 10 O-serotipos predominantes podrían cubrir un estimado del 60–80% de las infecciones de bacteriemia asociadas a *E. coli*, suponiendo la cobertura de subporciones de las cepas no tipificables.

40 Una vacuna multivalente que cubriera una proporción significativa de los serotipos de *E. coli* causante de bacteriemia sería muy útil. Los O-serotipos de la Tabla 2 serían así buenos candidatos para una vacuna multivalente basada en O-antígeno. Dicha vacuna podría prepararse beneficiosamente utilizando la tecnología de bioconjugación.

Uno de los serotipos de los 10 principales (Tabla 2) es O4. Por consiguiente, sería beneficioso preparar una vacuna de bioconjugado que incluya O–antígeno polisacárido de *E. coli* serotipo O4 acoplado a una proteína portadora.

Ejemplo 2: Caracterización de aislados clínicos de O4 contemporáneos para genes que codifican enzimas modificadoras de O–antígenos

5 Dos variantes de O4 antígeno polisacárido de *E. coli* han sido descritas (véase, por ejemplo, Jann B, *et al.*, 1993, Carbohydr. Res. 248: 241–250), uno que tiene una estructura no ramificada (estructura mostrada como (O4–Glc–) en la Tabla 1) y otra variante sustituida con una rama lateral de glucosa adicional (estructura mostrada como (O4–Glc+) en la Tabla 1). La proporción en que estas dos variantes se encuentran en aislados clínicos contemporáneos no se conocía. Aunque ambas variantes reaccionan con O4 antisueros, tampoco se sabía si existen diferencias inmunológicas entre estas variantes. Por otra parte, hasta ese momento no se había identificado una enzima responsable de unir la rama lateral de glucosa para generar el (O4–Glc+) antígeno polisacárido, y es probable que una secuencia de codificación putativa de la misma resida fuera del grupo de genes *rfb* O4.

15 Un conjunto de 32 aislados clínicos O4 de *E. coli* confirmado por aglutinación originalmente aislados durante el período de 2011 a 2012 de sujetos de los Estados Unidos y la Unión Europea se sometieron a análisis de secuencia de genoma completo. Las secuencias del grupo de genes *rfb* extraídas de los 32 aislados O4 secuenciados fueron alineadas con las de la cepa de referencia y se compararon a nivel de los nucleótidos. A excepción de algunos polimorfismos de nucleótido único que ocurren naturalmente, los aislados caracterizados mostraron todos un grupo *rfb* que era idéntico a la cepa de referencia O4, lo que indica que las cepas O4 de *E. coli*, independientemente de su estado de ramificación Glc, tienen un grupo de genes *rfb* idénticos. Por lo tanto, para generar el O4–Glc+ antígeno polisacárido de *E. coli*, probablemente se necesite un gen con secuencia desconocida que codifica una enzima de ramificación específica de O4 de *E. coli* y que debe residir en algún lugar fuera del grupo de genes *rfb* de O4 *E. coli*. La secuencia de este gen desconocido debe identificarse y emplearse si uno quiere producir bioconjugados con los O4–Glc+ antígenos polisacáridos de *E. coli* en una cepa que de otro modo solo produciría bioconjugados con O4–Glc– antígenos polisacáridos de *E. coli*.

25 Los datos de la secuencia de genoma completo se analizaron a continuación en cuanto a la presencia de genes fuera del grupo de genes *rfb* que pueden codificar enzimas modificadoras de O–antígenos. Los homólogos de *gtrAB* en *Shigella flexneri* se identificaron por primera vez en *E. coli* O4. Un marco de lectura abierto aguas abajo de *gtrAB* en *E. coli* fue luego identificado putativamente como el gen *gtrS* O4–específico de *E. coli*, que podría codificar la enzima ramificadora GtrS específica de O4 de *E. coli* responsable de la adición de una rama de glucosa al O4 antígeno de *E. coli*.

La secuencia de aminoácidos de la enzima GtrS específica de O4 se proporciona como SEQ ID NO: 4. Un ejemplo de secuencia de ácido nucleico que codifica esta proteína se proporciona como SEQ ID NO: 5.

35 De los aislados de O4 de *E. coli* caracterizados, se descubrió que aproximadamente el 80% lleva el gen *gtrS* identificado aquí (26 de 32). La prevalencia de la secuencia *gtrS* O4–específica de *E. coli* también fue determinada por PCR usando cebadores específicos de secuencia en un conjunto independiente de 20 aislados clínicos de O4 de *E. coli* confirmados por aglutinación aislados durante el período de 2014 a 2016 de sujetos de los Estados Unidos y la Unión Europea. Este análisis demostró que 17 de 20 aislados llevaban la secuencia *gtrS* de O4, lo que corresponde a una prevalencia del 85%.

40 **Ejemplo 3: Clonación de O4 *gtrS* en *E. coli* W3110, producción y confirmación estructural de O4 bioconjugados Glc–modificados**

Para evaluar si los bioconjugados que comprenden O4–antígeno polisacárido modificado con una glucosa de ramificación se podrían preparar, se construyeron cepas de producción de bioconjugado de *E. coli* O4–antígeno EPA con la enzima de ramificación putativa. Para ello, el gen endógeno O16–*gtrS* se sustituyó por el gen putativo O4–*gtrS* (SEQ ID NO: 5, véase el Ejemplo 2) y el grupo O16 *rfb* fue sustituido con el grupo O4 *rfb* en *E. coli* cepa W3110 $\Delta wzzE$ –*wecG* $\Delta waaL$ $\Delta wbbI$ –*J–K* por recombinación homóloga. Alternativamente, en algunas cepas, el grupo O4 *rfb* se codificó en un plásmido.

Posteriormente, los plásmidos que codifican una proteína portadora de exotoxina A desintoxicada de *Pseudomonas aeruginosa* (EPA) (una variante que tiene ya sea 2 o 4 sitios de glicosilación de consenso, denominada ‘EPA–2’ y ‘EPA–4’, respectivamente), y oligosacaril transferasa PglB se introdujeron en las cepas. Los bioconjugados O4–EPA modificados con Glc fueron producidos cultivando las cepas de producción de *E. coli* en cultivos de biorreactor, y la inducción de la expresión de PglB y EPA por IPTG y arabinosa, respectivamente. Los bioconjugados O4–EPA se extrajeron a partir del extracto periplásmico de biomasa.

55 Para confirmar la composición de polisacárido detallada y la unión de los bioconjugados O4–EPA, se realizaron múltiples experimentos de RMN en los bioconjugados que tienen proteína portadora EPA–4 (datos no mostrados). Las asignaciones obtenidas estuvieron de acuerdo con la literatura publicada (Jansson, P.E., *et al.*, 1984, Carbohydr. Res. 134(2): 283–291; Jann B, *et al.*, 1993, Carbohydr. Res. 248: 241–250). El espectro 1D registrado a 313K mostró una

gran señal de HOD y pequeñas señales nítidas a partir del O4 pentasacárido RU con cinco señales anoméricas, dos NAc y dos H6 (Rha y FucNAc).

5 Las asignaciones de protones 1D se confirmaron mediante el uso de experimentos 2D de RMN de correlación protón-protón y protón-carbono. En primer lugar, los experimentos 2D TOCSY (120 ms) demostraron los picos cruzados esperados de H1 y H6 (para Rha y FucNAc) para el O4 pentasacárido RU y pequeños picos de la terminal de RU y EPA. En la región de metilo, TOCSY mostró picos cruzados de H6 a H1 para α -RHA y de H6 a H5 para α -FucNAc para el O4 RU. Otros picos observados fueron de EPA aminoácidos y Rha terminal (tRha). En segundo lugar, un espectro de RMN de carbono contenía picos individuales bien dispersos y de diagnóstico para el O4 RU. Los carbonos fueron perfilados indirectamente a través de los protones unidos mediante el uso del experimento de HSQC. El experimento de HSQC-DEPT dio picos invertida para los grupos CH₂. El HSQC dio picos cruzados para los grupos O4 pentasacárido RU [5 anomérico, anillo, dos N-acetilo y dos metilo (Rha y FucNAc)], así como EPA aminoácidos en regiones características. Cada uno de los pares de protón / carbono para el O4 pudo ser asignado en base a las asignaciones de protones y la literatura.

15 Así, los experimentos de caracterización estructural confirmaron que los bioconjugados de O4 con ramificación de Glc (que comprenden estructuras de antígeno polisacárido como se indica por la Fórmula (O4-Glc+) en la Tabla 1) podría ser producidos usando el gen putativo de *E. coli* O4-*gtrS* identificado en el Ejemplo 2.

Ejemplo 4: Inmunogenicidad de un bioconjugado de O4 Glc-ramificado en conejos

20 Se usaron bioconjugados de O4 Glc-modificados (es decir, que tienen glicanos con la estructura de Fórmula (O4-Glc+) tal como se muestra en la Tabla 1) para la inmunización de conejos mediante la aplicación de un protocolo rápido en conejos (Eurogentec). Los sueros de los conejos inmunizados se analizaron por ELISA en cuanto a los títulos de IgG anti-O4 contra O4 lipopolisacárido purificado (LPS) con (Glc+; es decir, que contiene O4 polisacárido glucosilado) o sin ramificación de Glc (Glc-; es decir, que contiene O4 polisacárido no glucosilado). La inmunización con el bioconjugado dio por resultado títulos elevados de IgG en ambos conejos (FIG. 1). En ambos casos, los títulos de anticuerpos inducidos por el bioconjugado de O4 eran más altos contra Glc+ LPS en comparación con Glc- LPS.

25 Los sueros también se combinaron y se utilizaron en estudios ELISA de células enteras con equipos de prueba de aislados de O4 de *E. coli* con estado *gtrS* caracterizado. Se evaluaron cinco aislados de O4 de *E. coli gtrS*-negativos (sin ramificación de Glc) y seis *gtrS*-positivos (con ramificación de Glc) y una cepa de control negativa. Los sueros combinados de conejos inmunizados con un bioconjugado de O4 modificado por Glc contenían altos títulos de IgG que reconoce específicamente los aislados de O4 evaluados (FIG. 2). En concordancia con el LPS ELISA, todos los aislados de O4 evaluados fueron reconocidos por los sueros inmunes. Los aislados *gtrS*-positivos mostraron una unión general más alta que los aislados *gtrS*-negativos (FIG. 2). En particular, los siguientes aislados fueron *gtrS*-positivos: Y1382, E551, OC24334, stGVXN4983, stGVXN4994 y OC24794, y los siguientes aislados fueron *gtrS*-negativos: A2625, stGVXN4988, OC24784, OC24787 y OC24788.

35 Los perfiles de LPS extraídos del conjunto de prueba de aislados *gtrS*-positivos y negativos en geles de poliacrilamida teñidos con plata no revelaron diferencias marcadas entre los aislados que expresan formas modificadas y no modificadas del antígeno O4, lo que confirma que las diferencias observadas no se explican por diferencias cuantitativas en los niveles de expresión de LPS (datos no mostrados).

40 Se realizaron transferencias Western de LPS extraído por medio de sueros inmunes combinados para evaluar el reconocimiento de O4 O-antígeno por las IgG provocadas en respuesta a la inmunización con un bioconjugado de O4 modificado con Glc. Se observó la unión de O4 LPS modificados y no modificados por las IgG de conejos inmunizados con O4 modificado e incluyó el reconocimiento específico de bandas de LPS que abarcan una amplia gama de tamaños, incluyendo bandas de LPS de alto peso molecular (FIG. 3).

45 En los otros experimentos a continuación, cuando se hace referencia a un bioconjugado de 'O4' o a cepas de producción o 'EcoO4', significa el bioconjugado o la cepa de producción de O4 ramificado con Glc (que tiene una estructura de glicano (O4-Glc+) en la Tabla 1), a menos que se indique específicamente lo contrario (los términos 'O4' y 'O4-Glc+' por lo tanto se utilizan indistintamente para bioconjugados o cepas de producción en estos experimentos).

Ejemplo 5: Inmunogenicidad de un bioconjugado de O4 Glc-ramificado en ratas

50 Las ratas Sprague Dawley fueron inmunizadas por vía intramuscular 3 veces con tampón de formulación o (O4-Glc+)-EPA bioconjugado (es decir, bioconjugado de O4 antígeno polisacárido glucosilado unido covalentemente a una proteína portadora EPA; la proteína portadora era EPA-2, como se describe en el Ejemplo 3 anterior) en 3 dosis diferentes (0.04 μ g, 0.40 μ g o 4.0 μ g). Los niveles de anticuerpos en suero se midieron por ELISA el día 0, 14 y 42 después de la inmunización.

La inmunización con 0.04 μ g, 0.40 μ g y 4.00 μ g de (O4-Glc+)-EPA bioconjugado indujo un aumento significativo en los niveles de anticuerpos IgG el día 42 después de la inmunización cuando se compara con el tampón de formulación

(FIG. 4A). Los anticuerpos inducidos por (O4–Glc+)–conjugado fueron funcionales, es decir, capaces de mediar la destrucción de la cepa (O4–Glc+) de *E. coli* (FIG. 4B).

Los niveles de anticuerpos inducidos por 0.04 µg, 0.40 µg y 4.0 µg de (O4–Glc+)–EPA bioconjugado aumentaron significativamente el día 42 en comparación con los detectados en condición basal (día 42 versus día 0, P = 0.006 para todas las dosis) y el día 14 después de la inmunización (día 42 versus día 14, P = 0.006 para todas las dosis) (FIG. 5). En el grupo que recibió 4.0 µg de bioconjugado, los títulos también se incrementaron significativamente el día 14 en comparación con el día 0, lo que indica que una sola dosis de 4.0 µg de (O4–Glc+)–EPA bioconjugado induce un aumento significativo en los títulos de IgG (día 14 versus día 0, P = 0.012). El aumento significativo de los títulos de IgG observado entre el día 14 y el día 42, para las tres concentraciones de bioconjugado ensayadas, mostró que una tercera dosis de (O4–Glc+)– EPA bioconjugado es capaz de reforzar las respuestas de anticuerpos (FIG. 5).

La funcionalidad de los anticuerpos inducidos por O4–Glc+–EPA conjugado en las ratas inmunizadas por vía intramuscular 3 veces con tampón de formulación o el bioconjugado a 4.00 µg/dosis se determinó por el ensayo de destrucción opsonofagocítica (OPKA) usando cepas de *E. coli* O4(Glu+) y O4(Glu–). Los anticuerpos inducidos por el bioconjugado (O4–Glc+)–EPA fueron funcionales, es decir, capaces de mediar la destrucción de una cepa de *E. coli* (O4–Glc+) (FIG. 4B, FIG. 6). En particular, los anticuerpos inducidos por el bioconjugado (O4–Glc+)–EPA fueron capaces de mediar la destrucción de ambos (O4–Glc+) y (O4–Glc–), es decir, que tienen glicanos con la estructura de Fórmula (O4–Glc– en la Tabla 1, es decir, cepas de *E. coli* O4 polisacárido sin ramificación de Glc) (FIG. 6).

En conclusión, los anticuerpos inducidos por el bioconjugado O4–Glc+–EPA son de reacción cruzada y capaces de mediar la destrucción de cepas de *E. coli* O4 con y sin ramificación de glucosa.

Ejemplo 6: Cepas de producción para bioconjugados de O–antígeno de *E. coli* y productos de bioconjugados resultantes

Además de los bioconjugados (O4–Glc+)–EPA preparados como se describe anteriormente, se produjeron nueve (9) bioconjugados más. En particular, los bioconjugados producidos adicionalmente incluyeron bioconjugado O1A–EPA de *E. coli*, bioconjugado O2–EPA, bioconjugado O6A–EPA, bioconjugado O8–EPA, bioconjugado O15–EPA, bioconjugado O16–EPA, bioconjugado O18A–EPA, bioconjugado O25B–EPA y bioconjugado O75–EPA. Las estructuras químicas de los glicanos de estos conjugados se pueden ver en las fórmulas respectivas de la Tabla 1. Una composición que comprende los 10 bioconjugados se denomina aquí como ‘ExPEC10V’. Una composición que comprende los bioconjugados O1A–EPA, O2–EPA, O6A–EPA y O25B–EPA se denomina ‘ExPEC4V’ (y se ha descrito previamente en, por ejemplo, WO 2015/124769 y WO 2017/035181).

Cepa parental de *Escherichia coli* W3110

La cepa W3110 de *E. coli* K12 no patógena se usó como cepa parental para la construcción de las 10 cepas de producción. La cepa W3110 de *E. coli* K12 se obtuvo a partir del Coli Genetic Stock Center (Universidad de Yale, New Haven (CT), EE. UU., número de producto CGSC#4474). Su genotipo relevante fue descrito previamente (*E. coli* W3110, F–, lambda–, IN(rrnD–rrnE)1, rph–1) y su secuencia genómica se publicó anteriormente (Hayashi K, *et al.*, 2006, Mol. Syst. Biol. 2006.0007 (doi:10.1038/msb4100049). La cepa de *E. coli* W3110 fue modificada genéticamente para permitir la producción de cada uno de los bioconjugados de O–antígeno de *E. coli* (Tabla 3).

Cepas de producción de bioconjugados

Las composiciones de “ExPEC4V” y “ExPEC10V” ambas comprenden los bioconjugados O2–EPA y O25B–EPA a partir de las mismas cepas de producción. La composición de “ExPEC4V” comprende el bioconjugado O1A–EPA a partir de las cepas de producción stGVXN4411 o stLMTB10217, mientras que la composición de “ExPEC10V” comprende el bioconjugado O1A–EPA a partir de la cepa de producción stLMTB10217. La composición de “ExPEC4V” comprende el bioconjugado O6A–EPA a partir de la cepa de producción stGVXN4112, mientras que la composición de “ExPEC10V” comprende el bioconjugado O6A–EPA a partir de la cepa de producción stLMTB10923. Además, la composición de “ExPEC10V” comprende los bioconjugados O4–EPA (es decir, (O4–Glc+)– EPA), O8–EPA, O15–EPA, O16–EPA, O18A–EPA y O75–EPA a partir de cepas de producción que no son usadas para “ExPEC4V”. Diferentes cepas de producción podrían variar en los plásmidos para la expresión de la proteína portadora EPA y/o la oligosacaril transferasa PglB, como se indica a continuación. Una visión general de varias cepas de producción se da en la Tabla 3 a continuación.

Tabla 3: Descripción general de la ingeniería genética de las cepas de producción de *E. coli* para bioconjugados de O–antígeno para composiciones de vacunas ExPEC4V y ExPEC10V

Serotipo	Nombre de la cepa	Mutaciones genómicas			Plásmidos	
		Grupo de genes <i>rfb</i>	<i>waaL</i>	<i>gtrABS</i>	<i>pgIB</i>	<i>epa</i>
O1A (ExPEC4V)	stGVXN4411	$\Delta rfb::O1A$ <i>rfb</i> upecGVXN_032	$\Delta waaL$	–	pGVXN970	pGVXN1076

O1A (ExPEC4V; ExPEC10V)	stLMTB102 17	$\Delta rfb::O1A$ <i>rfb</i> upecGVXN_032	$\Delta waaL$	–	pGVXN12 21	pGVXN1076
O2	stGVXN490 6	$\Delta rfb::O2$ <i>rfb</i> upecGVXN_116	$\Delta waaL$	–	pGVXN97 1	pGVXN1076
O4	BVEC–L– 00684	$\Delta rfb::O4$ <i>rfb</i> CCUG11450	$\Delta waaL$	$\Delta gtrS::gtr$ S O4	pGVXN12 17	pGVXN1076
O6A (ExPEC4V)	stGVXN411 2	$\Delta rfb::O6A$ <i>rfb</i> CCUG11309	$\Delta waaL$	–	pGVXN11 4	pGVXN659
O6A (ExPEC10V)	stLMTB109 23	$\Delta rfb::O6A$ <i>rfb</i> CCUG11309	$\Delta waaL$	–	pGVXN12 21	pGVXN1076
O8	stLMTB117 34	$\Delta rfb::O8$ <i>rfb</i> E2420	$\Delta waaL$	$\Delta gtrABS$	pGVXN97 0	pGVXN1076
O15	stLMTB117 38	$\Delta rfb::O15$ <i>rfb</i> OC24891	$\Delta waaL$	$\Delta gtrABS$	pGVXN12 21	pGVXN1076
O16	stLMTB117 39	$\Delta rfb::O16$ <i>rfb</i> OC24208	$\Delta waaL$	$\Delta gtrABS$	pGVXN23 81	pGVXN1076
O18A	BVEC–L– 00559	$\Delta rfb::O18A$ <i>rfb</i> OC24255	$\Delta waaL$	$\Delta gtrABS$	pGVXN97 0	pGVXN1076
O25B	stGVXN445 9	$\Delta rfb::O25B$ <i>rfb</i> upecGVXN_138	$\Delta waaL$	$\Delta gtrABS$	pGVXN97 0	pGVXN1076
O75	stLMTB117 37	$\Delta rfb::O75$ <i>rfb</i> CCUG31	$\Delta waaL$	$\Delta gtrABS$	pGVXN12 17	pGVXN1076

Grupo de genes (*rfb*) de biosíntesis de O–antígeno

5 En todas las cepas de producción de O–antígeno de *E. coli*, el grupo de genes (*rfb*) de biosíntesis de antígeno O16::IS5 genómico W3110 de *E. coli* de origen natural fue sustituido por los grupos de biosíntesis específicos de O–antígeno seleccionados a partir de cepas de *E. coli* del serotipo seleccionado, que codifica las estructuras de O–antígeno de serotipo específico (véase la Tabla 1 para estas estructuras de O–antígeno). Se seleccionaron los diez grupos *rfb* donantes o se confirmaron después del análisis de genoma completo de aislados de sangre de *E. coli*. El reemplazo del grupo de genes *rfb* O16::IS5 de W3110, que es defectuoso en la biosíntesis de O–antígeno, se ha logrado en un único evento de recombinación homóloga. En el caso de los grupos de genes *rfb* O16 y O18A, el ADN donante se recombina a través de los genes flanqueantes *gnd* y *rmlCA*, mientras que el grupo de genes *rfb* para las otras cepas se recombina a través de los genes flanqueantes *gnd* y *galF*. Las secuencias de los grupos *rfb* en las cepas de producción se proporcionan en SEQ ID NOs: 9 y 11–19.

Gen de O–antígeno ligasa (*waaL*)

15 Todas las cepas de producción de O–antígeno de *E. coli* llevan una delección introducida artificialmente de la ligasa de O–antígeno genómico de *E. coli* W3110 codificada por el gen *waaL*. En las cepas $\Delta waaL$, la transferencia del O–antígeno al lípido A se interrumpe, lo que en cambio dirige la transferencia del O–antígeno a la proteína portadora para aumentar el rendimiento del producto.

Genes de glucosilación de O–antígeno (*gtrABS*)

20 En las cepas de producción de *E. coli* O8, O15, O16, O18A, O25B y O75, los genes *gtrABS* genómicos de *E. coli* W3110, que son responsables de la glucosilación de O16 O–antígeno, se han eliminado. Mientras que los genes *gtrA* y *gtrB* en diferentes serotipos son altamente homólogos e intercambiables, el gen *gtrS* codifica una glicosil transferasa de O–antígeno serotipo–específica. En *E. coli* W3110 GtrS puede transferir un residuo de glucosa (Glc) al azúcar GlcNAc en el motivo α -L-Rha-(1→3)-D-GlcNAc del O16 O–antígeno de *E. coli*. En las cepas de producción de *E. coli* O1A, O2 y O6A, no se ha producido supresión ni sustitución del gen *gtrABS*. Estos O–antígenos carecen del motivo α -L-Rha-(1→3)-D-GlcNAc que es el sustrato natural para *E. coli* O16 *gtrS*. En la cepa de producción de *E.*

coli O4, el gen *gtrS*W3110 ha sido reemplazado por el gen *gtrS* O4 de *E. coli* para acomodar la glucosilación apropiada del O4 O-antígeno de *E. coli*.

Oligosacaril transferasa PglB

5 Todas las cepas de producción de O-antígeno de *E. coli* expresaron una variante de la glicosil transferasa PglB de *C. jejuni*, que puede transferir el O-antígeno sobre una secuencia de consenso de aminoácidos en una proteína portadora mediante *N*-glucosilación. PglB tiene un amplio reconocimiento de sustrato, pero debido a los bajos rendimientos del producto se prepararon varias cepas de producción que expresan una variante de PglB que tiene especificidades de sustrato modificadas, lo que dio por resultado un mejor rendimiento del producto (véase, por ejemplo, WO 2016/107818, WO 2016/107819). El gen *pglB* se colocó detrás de un promotor inducible de β -D-1-tiogalactopiranosido de isopropilo (IPTG) en un plásmido. La Tabla 4 a continuación enumera las variantes de PglB codificadas por los plásmidos usados para la producción de las cepas de producción de O-antígeno de *E. coli* para los bioconjugados para las composiciones de ExPEC4V y ExPEC10V descritas anteriormente. Otros plásmidos con variación en la cadena principal del vector, marcador de resistencia a antibióticos y/o variantes alternativas de PglB también se han probado con éxito para la producción de bioconjugados.

15 **Tabla 4:** Plásmidos PglB y EPA utilizados en cepas de producción de O-antígeno de *E. coli*

Nombre del plásmido	Gen	Descripción ¹
pGVXN114	<i>pglB</i>	Uso de codón de <i>C. jejuni</i> ; SpR
pGVXN970	<i>pglB</i>	Uso de codón optimizado de <i>E. coli</i> ; SpR
pGVXN971	<i>pglB</i> ^{N534Q}	Uso de codón optimizado de <i>E. coli</i> ; el sitio de glucosilación natural de PglB fue inactivado; SpR
pGVXN1217	<i>pglB</i> ^{N311V}	Uso de codón optimizado de <i>E. coli</i> ; sustrato optimizado PglB; SpR
pGVXN1221	<i>pglB</i> ^{N311V,K482R,D483H,A669V}	Uso de codón optimizado de <i>E. coli</i> ; sustrato optimizado PglB; SpR
pGVXN2381	<i>pglB</i> ^{Y77H,S80R,Q287P,K289R,N311V}	Uso de codón optimizado de <i>E. coli</i> ; sustrato optimizado PglB; SpR
pGVXN659	EPA-4	EPA con cuatro sitios de bioconjugación; AmpR
pGVXN1076	EPA-4	EPA con cuatro sitios de bioconjugación; KanR
¹ SpR, resistente a espectinomocina; AmpR, resistente a ampicilina; KanR, resistente a kanamicina.		

Proteína portadora (EPA)

20 Todas las cepas de producción de O-antígeno de *E. coli* expresaron un toxoide de ADP-ribosiltransferasa (EPA) de *P. aeruginosa* genéticamente desintoxicado como proteína portadora para el O-antígeno. El toxoide EPA difiere de la toxina EPA de tipo salvaje en dos residuos: Leu552 se cambió a Val y Glu553 (en el dominio catalítico) se eliminó. Se informó que las deleciones de Glu553 redujeron significativamente la toxicidad. Además de la mutación desintoxicación, se introdujeron cuatro motivos de sitio de *N*-glucosilación de consenso (EPA-4). El gen *epa* se colocó detrás de un promotor inducible de L-arabinosa (Ara) en un plásmido (Tabla 4). La Tabla 4 se limita a los plásmidos utilizados en cepas de producción para bioconjugados utilizados en las composiciones "ExPEC4V" y "ExPEC10V" descritas anteriormente. Los plásmidos con variación en el esqueleto del vector, marcador de resistencia a antibióticos y/o variantes de EPA, por ejemplo, que varían en el número de motivos de sitio de *N*-glucosilación de consenso (por ejemplo, que tienen dos de tales motivos, EPA-2), también se han probado con éxito para la producción de bioconjugados.

30 **Ejemplo 7: Optimización de la oligosacariltransferasa para la generación de bioconjugados con antígeno O4 glucosilado (O4-Glc+)**

La optimización del rendimiento para la producción de bioconjugados se puede lograr mediante la modificación de la oligosacaril transferasa PglB de *C. jejuni*, lo que puede conducir a un grado más eficiente o mayor de *N*-glucosilación del O-antígeno de interés a la proteína portadora EPA. En una cepa de *E. coli* para la producción de bioconjugado

con O-antígeno polisacárido O4 glucosilado (O4-Glc+), se aplicó dicha estrategia de optimización y dio por resultado una variante de PglB específica para (O4-Glc+) que mejora el rendimiento del producto bioconjugado.

En este enfoque, una cepa que produce un O-antígeno polisacárido O4-Glc+ que contiene un plásmido de expresión de EPA se transformó con una variedad de diferentes plásmidos de expresión PglB, cada uno de los cuales contenía diferentes sustituciones de aminoácidos en la proteína PglB, alterando la especificidad del sustrato. El nivel de producción de bioconjugado y el perfil de cada cepa se evaluó a nivel del matraz de agitación en experimentos de choque osmótico, y la lectura se realizó mediante inmunoensayos de electroforesis capilar en el extracto periplásmico usando anticuerpos monoclonales específicos de O4-Glc+.

Una de las variantes de PglB probadas que contiene una sustitución de aminoácido N311V demostró mejorar el rendimiento del producto de O4 bioconjugados glucosilados significativamente (FIG. 7A).

En una mejora adicional donde se modificó aún más la variante de PglB N311V, una sustitución de aminoácidos Y77H mejoró adicionalmente el rendimiento del producto específico de O4-Glc+ y mostró un mayor grado de producto di- y tri-glicosilado en comparación con la variante N311V PglB, donde otras modificaciones resultaron ser neutrales o tuvieron un efecto negativo en el rendimiento del producto (FIG. 7B). El plásmido pLMTB4008 (SpR) codifica una variante de PglB de *E. coli* optimizada por uso de codón, optimizada por sustrato (O4-Glc+) con mutaciones Y77H y N311V.

La variante PglB con especificidad de sustrato optimizada para O-antígeno polisacárido O4-Glc+, que contiene sustituciones de aminoácidos N311V y Y77H en relación con la glicosil transferasa PglB de *C. jejuni* de tipo salvaje (wt), demostró duplicar el rendimiento del bioconjugado en comparación con la variante PglB-N311V optimizada en primera vuelta.

Del mismo modo, con el uso de tamices, las variantes de PglB de rendimiento óptimo también se determinaron para la producción de bioconjugado de O-antígeno de los otros nueve serotipos en la composición ExPEC10V.

Para los bioconjugados que tienen el O1A, O6A u O15 antígeno polisacárido, PglB con mutaciones de aminoácidos N311V, K482R, D483H y A669V demostró brindar los rendimientos más altos.

Para los bioconjugados que tienen el O2, O8, O18A u O25B antígeno polisacárido, el PglB de tipo salvaje (es decir, que no tiene mutaciones de aminoácidos en las posiciones 77, 80, 287, 289, 311, 482, 483 y 669) se encontró para dar el más alto rendimientos.

Para los bioconjugados que tienen el O16 antígeno polisacárido, PglB con mutaciones de aminoácidos Y77H, S80R, Q287P, K289R y N311V demostró brindar los rendimientos más altos.

Para los bioconjugados que tienen el O75 antígeno polisacárido, PglB con mutación de aminoácido N311V demostró brindar los rendimientos más altos.

Puede verse a partir de estos resultados que la variante óptima de PglB es diferente para diferentes O-antígenos, y que la variante óptima de PglB para producir un bioconjugado con un O-antígeno polisacárido dado es impredecible.

Ejemplo 8: Bioconjugados de O-antígenos de 10 serotipos de *E. coli* y sus atributos de calidad

Los residuos de O-glicano de los O-antígenos blanco son estructuralmente diversos y tienen unidades de repetición variables. La especificidad y la afinidad de la glicosil transferasa PglB está ligada a la estructura de glicano. Por lo tanto, la preparación de un bioconjugado que tenga los atributos de calidad deseados, por ejemplo, pureza, relación glicano / proteína, etc., es un reto, no una tarea sencilla. La combinación correcta de proteínas portadoras PglB y EPA determina el rendimiento y puede influir en la eficiencia de la glicosilación. Mediante la optimización del PglB y las proteínas portadoras, se produjeron bioconjugados con los atributos de calidad deseados. También puede ser importante mantener un valor umbral más bajo de proteína portadora total, particularmente cuando uno o más bioconjugados de O-antígeno se combinan y se administran en una única composición o vacuna, porque cantidades muy altas de proteína portadora pueden conducir a una interferencia inmunológica. A fin de evitar tal fenómeno, se prefieren los conjugados que tienen una relación glicano / proteína mayor. Por lo tanto, para la vacuna ExPEC10V, se desarrollaron bioconjugados con una relación de glicosilación al menos comparable (a la vacuna ExPEC4V anteriormente descrita que ha sido objeto de ensayos clínicos).

Cada uno de los bioconjugados fue producido cultivando las células huésped respectivas (Ejemplo 6, Tabla 3) en biorreactores (volúmenes de 10 L y/o 200 L) y la expresión de los bioconjugados, siguiendo los métodos descritos previamente. Cada sustancia farmacológica fue fabricada en lotes por fermentación microbiana en lotes para generar una biomasa que contiene los bioconjugados expresados del correspondiente serotipo de polisacárido. Las células se cultivaron y se indujeron con IPTG y arabinosa. Los bioconjugados se aislaron a partir del periplasma de las células en los cultivos del biorreactor mediante choque osmótico seguido por purificación cromatográfica. Este proceso se llevó a cabo para cada uno de los 10 bioconjugados.

Los bioconjugados de O-antígeno de *E. coli* preparados de este modo que son sustancias farmacológicas (DS) para ExPEC10V y ExPEC4V mostraron atributos de calidad crítica comparables: (1) la pureza relacionada con el proceso (medida por RP-HPLC) fue mayor al 95%, (2) la relación polisacárido / proteína varió entre aproximadamente 0.1–0.5, principalmente entre 0.15 y 0.45, (3) la endotoxina bacteriana (Ph. Eur. 2.2.3) fue de menos de 0.5 UE/μg de polisacárido. La longitud media de las cadenas de polisacáridos individuales fue típicamente entre aproximadamente 10–20 unidades de repetición (medidas utilizando SDS-PAGE de alta resolución).

Las estructuras de las unidades de repetición de polisacáridos fueron confirmadas (por RMN y MS/MS de los conjugados, intactos o digeridos con tripsina) para ser los que se muestran en las fórmulas para los serotipos correspondientes en la Tabla 1, para todos los 10 bioconjugados que son DS para la composición ExPEC10V descrita anteriormente.

El serotipo O18 tuvo los rendimientos más bajos de producción de bioconjugado entre los 10 serotipos de los cuales se hicieron bioconjugados para la composición ExPEC10V.

El producto farmacológico (DP) ExPEC10V comprende una mezcla de los 10 DS monovalentes descritos antes.

Ejemplo 9: Toxicología de la vacuna ExPEC10V

Se llevó a cabo un estudio de toxicidad piloto de dosis única y tolerancia local (no GLP) con ExPEC10V en conejos NZW hembras. Un grupo (n = 2) recibió una inyección intramuscular (IM) (el Día 0) del control (solución salina), y un segundo grupo (n = 4) recibió una inyección IM de ExPEC10V a razón de 105.6 μg de polisacárido total (PS)/dosis (9.6: 9.6: 9.6: 9.6: 9.6: 9.6: 9.6: 9.6: 9.6: 9.6: 19.2: 9.6 μg de PS por dosis para, respectivamente, los O-serotipos O1A, O2, O4, O6A, O8, O15, O16, O18A, O25B y O75) usando un volumen de dosificación de 0.6 ml (176 μg PS/ml). Se realizó la necropsia el Día 2.

No se observó mortalidad. Además, no hubo efectos relacionados con la vacuna para las observaciones clínicas (incluyendo efectos en el sitio de inyección utilizando el puntaje de Draize), peso corporal, consumo de alimento y temperatura corporal. Histopatológicamente, no se observaron cambios relacionados con la vacuna en el sitio de administración o el ganglio linfático de drenaje (ilíaco). Se observó un aumento mínimo en la formación de centros germinales en el bazo en uno de cuatro animales tratados (Día 2), y se consideró una respuesta inmunológica normal a la vacuna inyectada. En general, la administración de una dosis IM única de ExPEC10V a conejos hembras fue bien tolerada.

Ejemplo 10: Inmunogenicidad de la formulación mezclada de ExPEC10V en conejos

Se ha demostrado anteriormente que una vacuna ExPEC4V (que comprende bioconjugados de serotipos O1A, O2, O6A y O25B de *E. coli*) es inmunogénica para estos cuatro serotipos en ratas, conejos y seres humanos (véase, por ejemplo WO 2015/124769, WO 2017/035181; Huttner et al, 2017, Lancet Infect Dis, [http://dx.doi.org/10.1016/S1473-3099\(17\)30108-1](http://dx.doi.org/10.1016/S1473-3099(17)30108-1); RW Frenck Jr, et al, abstract 5587, ASM Microbe 2018). Los nuevos bioconjugados de la invención con el serotipo O4 glucosilado de *E. coli* demostraron ser inmunogénicos en los Ejemplos 4 y 5 más arriba. La inmunogenicidad de los bioconjugados de serotipos O8, O15, O16, O18A y O75 de *E. coli* (todos con EPA-2 como proteína portadora en este experimento) cuando se administraron por separado (monovalente) a ratas confirmó que también cada uno de estos bioconjugados era inmunogénico, ya que los datos de ELISA indicaron que cada uno de estos bioconjugados podía provocar altos niveles de anticuerpos específicos de O-antígeno de *E. coli* (no mostrado).

La inmunogenicidad de la vacuna 10-valente que contenía una mezcla de los 10 bioconjugados como se describe anteriormente también fue probada. Los conejos blancos de Nueva Zelanda (NZW) (hembras, de 12–16 semanas de edad) recibieron 3 inmunizaciones intramusculares con ExPEC10V o solución salina administradas con 2 semanas de diferencia (Tabla 5; administración los días 0, 14 y 27). Los 10 polisacáridos que son parte de la vacuna ExPEC10V utilizada en estos experimentos fueron conjugados a la proteína portadora EPA que contiene 4 sitios de glicosilación (EPA-4). La vacuna se formuló en 3 dosis diferentes: Grupo 1 ('dosis alta'): 8 ug/dosis de O1A, O2, O6A, O4, O8, O15, O16, O18 y O75 y 16 ug/dosis de O25B; Grupo 2 ('dosis media'): 4 ug/dosis de O2, O4, O8, O15, O16, O18 y O75, 8 ug/dosis de O1A y O6A y 16 ug/dosis de O25B; Grupo 3 ('dosis baja'): 0.4 ug/dosis de O2, O4, O8, O15, O16, O18 y O75, 0.8 ug/dosis de O1A y O6A y 1.6 ug/dosis de O25B. Los animales del grupo de control (Grupo 4) recibieron solo solución salina (solución de cloruro de sodio al 0.9% (p/v)) (Tabla 5).

Las respuestas de anticuerpos se evaluaron el día 0 (antes de la inmunización) y los días 14, 27 y 42 después de la inmunización. Los niveles de anticuerpos en suero inducidos por cada uno de los bioconjugados incluidos en la vacuna y la proteína portadora EPA se midieron por ELISA (IgG total), con LPS específico del tipo como material de recubrimiento. Los títulos de anticuerpos se informaron como valores de EC50 que corresponden a la concentración efectiva media máxima basado en duplicados de curvas de titulación de 12 pasos representadas en un modelo de regresión no lineal logística de 4 parámetros. La actividad funcional se determinó mediante OPK.

Tabla 5. Descripción de los grupos experimentales

Grupos experimentales	Dosis ($\mu\text{g/PS}$) O1A:O2:O6A:O25B:O4:O8:O15:O16:O18A:O75	Tamaño de muestra
Grupo 1 (dosis alta)	8:8:8:16:8:8:8:8:8:8	7
Grupo 2 (dosis media)	8:4:8:16:4:4:4:4:4:4	7
Grupo 3 (dosis baja)	0.8:0.4:0.8:1.6:0.4:0.4:0.4:0.4:0.4:0.4	7
Grupo 4 (control)	Solución de cloruro de sodio al 0.9% (p/v)	7

Los resultados se muestran en la FIG. 8 y se resumen en la Tabla 6.

Tabla 6. Resumen de las respuestas de anticuerpos específicos de O-antígeno de *E. coli* inducidos por ExPEC10V en conejos NZW

5

ExPEC10V	Respuestas de anticuerpos el día 14 después de la vacunación									
Dosis	O1A	O2	O6A	O25B	O4	O8	O15#	O16	O18A	O75
Alta	*	**	**	*	**	Ns	**	**	*	Ns
Media	*	**	**	**	**	Ns	**	**	Ns	Ns
Baja	*	*	*	*	*	Ns	**	**	Ns	Ns

ExPEC10V	Respuestas de anticuerpos el día 27 después de la vacunación									
Dosis	O1A	O2	O6A	O25B	O4	O8	O15#	O16	O18A	O75
Alta	**	**	**	**	**	*	**	**	**	**
Media	**	**	**	**	**	*	**	**	*	**
Baja	**	**	**	**	**	*	**	**	**	**

ExPEC10V	Respuestas de anticuerpos el día 42 después de la vacunación									
Dosis	O1A	O2	O6A	O25B	O4	O8	O15#	O16	O18A	O75
Alta	**	**	**	**	**	**	**	**	**	**
Media	**	**	**	**	**	**	**	**	**	**
Baja	**	**	**	**	**	**	**	**	**	**

Los cuadrados de color gris oscuro muestran las respuestas de anticuerpos específicos de serotipo en las que los valores de p fueron estadísticamente significativos. Los cuadrados de color gris claro muestran las respuestas de anticuerpos específicos de serotipo en las que los valores de p no fueron estadísticamente significativos (ns). Prueba de Wilcoxon de suma de rangos con corrección de Bonferroni para comparaciones múltiples. Comparaciones de animales vacunados con ExPEC10V (Grupo 1, 2 y 3) versus control de solución salina (Grupo 4). * $p \leq 0.05$, ** $p \leq 0.01$. # Los valores de P fueron estadísticamente significativos después de excluir un animal con valor atípico del grupo de control (análisis de sensibilidad).

15

- 5 La alta dosis de ExPEC10V (Grupo 1) indujo niveles significativamente más altos de anticuerpos IgG en todos los puntos temporales investigados (Días 14, 27 y 42 después de la inmunización) en comparación con el control de solución salina para O1A, O2, O4, O6A, O16, O18A y O25B (FIG. 8, Tabla 6). Se observaron títulos de anticuerpos significativamente más altos inducidos por conjugados O8 y O75 en comparación con el control de solución salina los Días 27 y 42 después de la inmunización (FIG. 8, Tabla 6).
- 10 La dosis media de ExPEC10V (Grupo 2) y la dosis baja (Grupo 3) indujeron niveles significativamente más altos de anticuerpos en todos los puntos temporales investigados (Días 14, 27 y 42 después de la inmunización) en comparación con el control de solución salina para O1A, O2, O4, O6A, O16 y O25B (FIG. 8, Tabla 6). Se observaron títulos de anticuerpos significativamente más altos inducidos por conjugados O8, O18A y O75 en comparación con el control de solución salina los Días 27 y 42 después de la inmunización, lo que sugiere que la dosis de refuerzo en conejos aumenta la respuesta a estos O-serotipos (FIG. 8, Tabla 6).
- 15 Para los conjugados O15, el análisis de sensibilidad omitiendo un animal con valor atípico del grupo de control mostró que las tres dosis de la vacuna ExPEC10V indujeron un aumento significativo en las respuestas de anticuerpos en comparación con el control de solución salina los Días 14, 27 y 42 después de la inmunización (FIG. 8, Tabla 6).
- Los anticuerpos inducidos por la proteína portadora EPA fueron significativamente mayores que los títulos de anticuerpos de EPA en el grupo tratado con solución salina (control) para las tres dosis de ExPEC10V probadas (alta, media y baja) en todos los puntos temporales investigados (Días 14, 27 y 42) (FIG. 8).
- 20 Las comparaciones entre dosis (no mostradas) mostraron que el Día 14 después de la vacunación, la dosis alta de ExPEC10V indujo respuestas de anticuerpos significativamente más altas en comparación con la dosis baja para la mayoría de los conjugados ensayados (O1A, O2, O4, O6A, O15, O16, O18A y O25B). La dosis media de ExPEC10V también indujo respuestas de anticuerpos significativamente más altas en comparación con la dosis baja para O1A, O2, O4, O18A, O25B y O75. Para el conjugado O8, las tres formulaciones de ExPEC10V indujeron niveles similares de anticuerpos el Día 14 después de la vacunación.
- 25 La dosis baja de ExPEC10V indujo un aumento significativo en las respuestas de anticuerpos el Día 42 después de la vacunación (después de una dosis de cebado y dos dosis de refuerzo) en comparación con las dosis alta y media de ExPEC10V para los conjugados O1A, O2, O4, O16, O25B y O75. Estos resultados están en línea con otras experiencias con vacunas conjugadas, en donde, por ejemplo, no se observó ninguna relación clara entre la dosis y la magnitud de la respuesta de anticuerpos a la vacunación primaria en lactantes vacunados con vacuna neumocócica conjugada (Poolman JT, *et al.* Expert Rev Vaccines. 2013, 12(12):1379–94).
- 30 No hubo diferencias significativas entre las tres dosis de ExPEC10V ensayadas el Día 42 después de la vacunación para los conjugados O6A, O8 y O15. Para el conjugado O18A, la dosis alta de ExPEC10V indujo una respuesta de anticuerpos significativamente más alta en comparación con la dosis media el Día 42 después de la vacunación.
- 35 Para la proteína portadora (EPA), la dosis alta y media de ExPEC10V indujo respuestas de anticuerpos significativamente mayores en comparación con la dosis baja el Día 14 después de la vacunación. La dosis alta de la vacuna también indujo respuestas de anticuerpos significativamente más altas en comparación con la dosis baja el Día 42 después de la vacunación.
- En conclusión, las tres formulaciones de ExPEC10V (alta, media y baja), administradas por inyección intramuscular los Días 0, 14, 27 son inmunogénicas en conejos.
- 40 Hasta ahora, los anticuerpos funcionales capaces de matar cepas de *E. coli* inducidas por esta vacuna en conejos se demostraron para los serotipos O1A, O2, O4, O6A, O15, O16 y O25B.
- 45 En un experimento adicional, un lote GMP de la vacuna ExPEC10V (véase el Ejemplo 8 anterior para la producción) se preparó y se inyectó en conejos NZW como parte de un estudio de toxicología (Tabla 7). En este estudio, los conejos NZW (machos y hembras) recibieron 3 inyecciones intramusculares (0.6 ml) de la vacuna ExPEC10V (Días 1, 15 y 29) y un grupo de control recibió solución de cloruro de sodio al 0.9% (p/v) (solución salina). Cada dosis de la vacuna contenía 9.6 µg de polisacárido (PS) para los serotipos O1A, O2, O4, O6A, O8, O15, O16, O18A y O75 y 19.2 µg de PS para los serotipos O25B, correspondiente a 105.6 µg de PS total (176 µg de PS total/ml) y 382.8 µg de EPA total (638 µg de EPA/ml). Los títulos de IgG contra O-antígenos y proteína portadora (EPA) se determinaron a partir de muestras recogidas durante el período pretratamiento (día 1) y los días 31 y 50 después de la inmunización.
- 50 Se observó un aumento significativo en las respuestas de anticuerpos contra todos los O-antígenos y la proteína portadora EPA el día 31 y 50 después de la vacunación en el grupo que recibió ExPEC10V en comparación con el grupo de control que recibió solo solución salina (Fig. 9, Tabla 8). Para el serotipo O1A, también se observó una respuesta de anticuerpos significativamente más alta el día 1 (línea basal) cuando los animales vacunados se compararon con los controles. Estos resultados sugieren que algunos animales fueron preexposados a *E. coli* o tienen anticuerpos que reaccionan de forma cruzada con O1A-LPS.

Tabla 7. Grupos experimentales y dosis de ExPEC10V utilizada en conejos NZW

Grupos	Tratamiento	Dosis	Días de dosificación	Principal (día 31) (machos / hembras)	Recuperación (día 50) (machos / hembras)
1	control	0	1, 15, 29	10	10
2	ExPEC10V	105.6 µg PS*	1, 15, 29	10	10

* Cada dosis (0.6 ml de volumen de dosificación) contiene 9.6:9.6:9.6:9.6:9.6:9.6:9.6:19.2:9.6 µg de polisacárido (PS) para los serotipos O1A, O2, O4, O6A, O8, O15, O16, O18A, O25B, O75, respectivamente (176 µg de PS total/ml). Cada dosis contiene 382.8 µg de proteína EPA (638 µg de EPA/ml).

Tabla 8. Inmunogenicidad de ExPEC10V en conejos NZW como parte de un estudio toxicológico

Tratamiento	Respuestas de anticuerpos el día 31 después de la vacunación									
ExPEC10V	O1A	O2	O6A	O25B	O4	O8	O15	O16	O18A	O75
Día 31	****	****	****	****	****	****	****	****	****	****
Día 50	****	****	****	****	****	****	****	****	****	****

- 5 Respuestas de anticuerpos inducidas por ExPEC10V. Los cuadrados de color gris claro muestran los serotipos en los que se observó un aumento significativo en las respuestas de anticuerpos en el grupo vacunado comparado con el control. Modelo de Tobit con una prueba de razón de verosimilitud. **** $P \leq 0.0001$.

Ejemplo 11: Ensayo de fase 1/2a con la vacuna ExPEC10V en humanos

- 10 En la actualidad, no existe una vacuna disponible para prevenir la IED. Los serotipos que comprenden la vacuna ExPEC10V (O1A, O2, O4, O6A, O8, O15, O16, O18A, O25B y O75) se seleccionaron para abordar la enfermedad invasiva causada por la mayoría de las cepas clínicamente relevantes de ExPEC que también representan la mayoría de los aislados de ExPEC que causan IED resistente a los antimicrobianos, incluyendo ST131. Los serotipos seleccionados son representativos de los diez O-serotipos de ExPEC prevalentes que causan infecciones del torrente sanguíneo en la población de más edad y responsable de aproximadamente el 70% de las infecciones del torrente sanguíneo causadas por ExPEC.

Puesto que no se espera que el mecanismo de acción de las vacunas conjugadas en la prevención de la enfermedad invasiva vaya a ser afectado por mecanismos de resistencia a los antibióticos, se cree que la vacuna ExPEC10V proporciona protección contra IED causada por serotipos fármaco-resistentes y fármaco-sensibles O2, O4, O6A, O8, O15, O16, O18A, O25B y O75.

- 20 Existe experiencia clínica precedente con ExPEC4V, un candidato de vacuna anterior, que comprendía un subconjunto de cuatro de los conjugados de O-antígeno de *E. coli* (O1A, O2, O6A y O25B) también hallados en ExPEC10V. Basándose en los resultados de cuatro estudios clínicos (dos estudios de fase 1 completados, un estudio de fase 2 completado y un estudio de fase 2 en curso), ExPEC4V fue bien tolerado por los participantes del estudio y no se observaron señales de seguridad relacionadas con la vacuna en dosis de hasta 16 µg de polisacárido (PS) por serotipo (O1A, O2, O6A y O25B). La mayoría de los eventos adversos (AE) fueron de Grado 1 y 2; se registraron muy pocos AE de Grado 3. Se observaron AE locales solicitados de inicio tardío (AE que se inician después del Día 5 posterior a la vacunación) principalmente con las dosis más altas de ExPEC4V. En cada estudio, la vacuna ExPEC4V mostró ser inmunogénica, lo que demuestra una respuesta inmune a la vacuna dependiente de la dosis, y aumentos de títulos de inmunoglobulina G (IgG) específicos de O-antígeno, tal como se mide por un ensayo de inmunoabsorción ligado a enzimas (ELISA). La actividad funcional de los anticuerpos se demostró con un ensayo de destrucción opsonofagocítica (OPKA) optimizado por ExPEC4V. el coanálisis de resultados de pruebas ELISA y OPKA mostró correlación entre las respuestas de ensayo (coeficientes de correlación de Pearson ≥ 0.61 y ≥ 0.48 para el Día 30 y el Día 360, respectivamente, en un ensayo clínico de Fase 2 [estudio 4V-BAC2001]), lo que justifica el uso de ELISA como medida primaria de los títulos de anticuerpos ExPEC4V y para predecir la actividad de anticuerpo funcional. El análisis de los datos de inmunogenicidad ha demostrado la duración de la respuesta inmune durante tres años tras la vacunación con ExPEC4V. Ahora también se ha observado que los sueros de seres humanos vacunados con ExPEC4V y que tenían altos títulos de anticuerpos opsonofagocíticos específicos de serotipo, cuando se transfirieron pasivamente a ratones que fueron posteriormente desafiados por vía intraperitoneal con cepas de *E. coli* de serotipo O25B u O2, fueron capaces de mediar la protección *in vivo* (no se muestra). Por lo tanto, los anticuerpos humanos opsonofagocíticos específicos de ExPEC4V median la muerte bacteriana *in vivo*, lo que está en línea con otras vacunas conjugadas en las que el mecanismo de protección propuesto es mediante la inducción de anticuerpos opsonofagocíticos que median la muerte bacteriana.

5 ExPEC10V incluye un total de diez serotipos y aumenta la cobertura de aproximadamente el 50% (ExPEC4V) a aproximadamente el 70% de las infecciones del torrente sanguíneo causadas por ExPEC en adultos de 60 años y mayores. Basado en la experiencia clínica con ExPEC4V, y en los datos preclínicos para ExPEC10V como se discute en los ejemplos anteriores, se espera que la administración de ExPEC10V induzca respuestas inmunes a *E. coli* serotipos O1A, O2, O4, O6A, O8, O15, O16, O18A, O25B y O75 también en los seres humanos.

10 Un estudio aleatorizado, de observador ciego, primero en humanos, de fase 1/2a, para evaluar la seguridad, la reactogenicidad y la inmunogenicidad de tres dosis diferentes de la vacuna ExPEC10V se lleva a cabo en seres humanos de 60 a 85 años de edad con salud estable (estudio 10V–BAC1001). El diseño del estudio incluye 2 cohortes: un total de 1.004 participantes son enrolados en el estudio con 404 participantes (100 participantes / dosis de ExPEC10V) con edades ≥60 a ≤85 años con salud estable en la Cohorte 1 y un adicional de 600 participantes de ≥60 años de edad con salud estable con antecedentes de UTI en los últimos 5 años en la Cohorte 2.

15 ExPEC10V es un candidato de vacuna 10–valente en desarrollo para la prevención de la enfermedad por *Escherichia coli* patogénica extraintestinal (ExPEC) (IED) en adultos de 60 años de edad y mayores. ExPEC10V se compone de los O–antígenos polisacáridos (PS) de ExPEC serotipos O1A, O2, O4, O6A, O8, O15, O16, O18A, O25B y O75 bioconjugados por separado a la proteína portadora, una forma genéticamente desintoxicada de exotoxina A (EPA) derivada de *Pseudomonas aeruginosa*, y su producción se ha descrito anteriormente. El O4 PS es la forma glucosilada, que tiene la estructura de Fórmula (O4–Glc+) de la Tabla 1.

OBJETIVOS Y CRITERIOS DE VALORACIÓN

20 COHORTE 1 – Período de Fase 1/2a con observador ciego y etiqueta abierta, período de seguimiento a largo plazo (N = 404):

Objetivos	Criterios de valoración
Primarios	
Evaluar la seguridad y la reactogenicidad de diferentes dosis de ExPEC10V en participantes ≥60 a ≤85 años de edad	Eventos adversos (AE) locales y sistémicos recogidos durante 14 días después de la vacunación (del día 1 al día 15) AE no solicitados recogidos desde la administración de la vacuna del estudio hasta 29 días después de la vacunación (del día 1 al día 30) Eventos adversos serios (SEA) recogidos desde la administración de la vacuna del estudio hasta el día 181
Evaluar la inmunogenicidad dependiente de la dosis de ExPEC10V el día 15 en participantes ≥60 a ≤85 años de edad	Títulos de anticuerpos para ExPEC10V, según lo determinado por inmunoensayo basado en electroquimioluminiscencia multiplex (ECL) y ensayo opsonofagocítico multiplex (MOPA) el día 15
Secundarios	
Evaluar la correlación entre los títulos séricos por inmunoensayo basado en ECL multiplex (anticuerpo total) y MOPA (anticuerpos funcionales) el día 15	Títulos de anticuerpos para ExPEC10V, según lo determinado por inmunoensayo basado en ECL multiplex y MOPA el día 15
Evaluar la inmunogenicidad dependiente de la dosis de ExPEC10V los Días 30 y 181 participantes ≥60 a ≤85 años de edad	Títulos de anticuerpos para ExPEC10V, según lo determinado por inmunoensayo basado en ECL multiplex y MOPA los días 30 y 181

Objetivos	Criterios de valoración
Evaluar, en el período de seguimiento a largo plazo (LTFU), la seguridad de la dosis de ExPEC10V seleccionada para un mayor desarrollo clínico basado en el análisis primario del día 30 en participantes ≥60 a ≤85 años de edad	SEA relacionados con la vacuna del estudio o los procedimientos del estudio recogidos desde el día 182 hasta el final del estudio
Evaluar, en el período LTFU, la inmunogenicidad de la dosis de ExPEC10V seleccionada para un mayor desarrollo clínico basado en el análisis primario del Día 30	Títulos de anticuerpos para ExPEC10V, según lo determinado por inmunoensayo basado en ECL multiplex y MOPA el Año 1 (día 366), Año 2 (día 731) y Año 3 (Día 1096)

ES 2 951 442 T3

COHORTE 2 – Período a doble ciego con período de seguimiento a largo plazo a doble ciego (N = 600):

Objetivos	Criterios de valoración
Primarios	
Evaluar la seguridad y la reactogenicidad de la dosis seleccionada de ExPEC10V en participantes ≥ 60 años de edad con antecedentes de UTI en los últimos 5 años	AE locales y sistémicos solicitados recogidos durante 14 días después de la vacunación (del día 1 al día 15) AE no solicitados recogidos desde la administración de la vacuna del estudio hasta 29 días después de la vacunación (del día 1 al día 30) SAE recogidos desde la administración de la vacuna del estudio hasta el día 181
Evaluar la inmunogenicidad de la dosis seleccionada de ExPEC10V el día 30 en participantes ≥ 60 años de edad con antecedentes de UTI en los últimos 5 años	Títulos de anticuerpos para ExPEC10V, según lo determinado por inmunoensayo basado en ECL multiplex y MOPA el Día 30
Secundarios	
Evaluar la correlación entre los títulos séricos por inmunoensayo basado en ECL multiplex (anticuerpo total) y MOPA (anticuerpo funcional) el día 30 en participantes ≥ 60 años de edad con antecedentes de UTI en los últimos 5 años	Títulos de anticuerpos para ExPEC10V, según lo determinado por inmunoensayo basado en ECL multiplex y MOPA el Día 30
Evaluar la inmunogenicidad de la dosis seleccionada de ExPEC10V los Días 15 y 181 en participantes ≥ 60 años de edad con antecedentes de UTI en los últimos 5 años	Títulos de anticuerpos para ExPEC10V, según lo determinado por inmunoensayo basado en ECL multiplex y MOPA los días 15 y 181

Objetivos	Criterios de valoración
Evaluar, en el período LTFU, la seguridad de la dosis seleccionada de ExPEC10V en participantes ≥ 60 años de edad con antecedentes de UTI en los últimos 5 años	SAE relacionados con la vacuna del estudio o los procedimientos del estudio recogidos desde el día 182 hasta el final del estudio
Evaluar, en el período LTFU, la inmunogenicidad de la dosis seleccionada de ExPEC10V en participantes ≥ 60 años de edad con antecedentes de UTI en los últimos 5 años	Títulos de anticuerpos para ExPEC10V, según lo determinado por inmunoensayo basado en ECL multiplex y MOPA el Año 1 (día 366), Año 2 (día 731) y Año 3 (Día 1096)
Exploratorios	
<ul style="list-style-type: none"> • Evaluar el efecto de ExPEC10V en el microbioma intestinal (heces) mediante análisis de metagenómica 	<ul style="list-style-type: none"> • Metagenómica de muestras de heces de un subconjunto seleccionado de participantes para evaluar el efecto de ExPEC10V en: <ul style="list-style-type: none"> La prevalencia de patógenos (por ejemplo, Clostridium difficile) en la flora intestinal Prevalencia de serotipos de ExPEC10V en la flora intestinal

DISEÑO GENERAL

- 5 Se trata de un estudio multicéntrico, aleatorizado, de intervención, que incluye dos cohortes.

Para la Cohorte 1, el estudio tiene un diseño de observador ciego, con control activo, y se incluye un total de 404 participantes adultos de edad ≥ 60 a ≤ 85 años con salud estable, con o sin antecedentes de UTI. El diseño del estudio

para la Cohorte 1 se compone de tres periodos: un período de selección máximo de 28 días, un período de seguimiento de 181 días con observador ciego con vacunación el Día 1 y un periodo LTFU de etiqueta abierta que dura desde el día 182 hasta 3 años (Día 1096) después de la vacunación (FIG. 10A). Solo los participantes del grupo de dosis seleccionada de ExPEC10V (aproximadamente 100 participantes) y los participantes del grupo Pevnar 13 avanzan al período LTFU. El final de la Cohorte 1 es la visita del Año 3 del último participante (Día 1096).

Para la Cohorte 2, el estudio tiene un diseño a doble ciego, controlado con placebo, y se incluye un total de 600 participantes adultos de edad ≥ 60 años con salud estable y antecedentes de UTI en los últimos 5 años. El enrolamiento comienza después de la finalización del análisis primario de Fase 1/2a y la selección de dosis de ExPEC10V a partir de la Cohorte 1. El diseño del estudio para la Cohorte 2 se compone de tres periodos: un período de selección máximo de 28 días, un período de seguimiento a doble ciego de 181 días con vacunación el Día 1, y un período de LTFU a doble ciego que dura desde el Día 182 hasta los 3 años (Día 1096) después de la vacunación (FIG. 10B). Todos los participantes de la Cohorte 2 avanzan al período LTFU. El fin del estudio es la visita del Año 3 del último participante (Día 1096) en la Cohorte 2.

Cohorte 1: Fase 1

En la Fase 1 de la Cohorte 1, un total de 84 participantes son enrolados en un enfoque escalonado siguiendo procedimientos de escalación de dosis en etapas, con realización de evaluaciones de seguridad antes de avanzar de una etapa a la siguiente. Un Comité de Revisión de Datos (DRC) interno se pone en marcha para este estudio a fin de revisar los datos de los exámenes físicos (tanto basales como dirigidos), los datos demográficos basales y los datos de seguridad posteriores a la vacunación el día 14 (incluyendo AE locales y sistémicos solicitados, AE no solicitados, SAE, datos de laboratorio clínico y signos vitales) de estos 84 participantes de la Fase 1. En esta fase del estudio, los participantes fueron enrolados y aleatorizados en seis pasos:

Paso 1: Cuatro participantes centinela fueron enrolados y aleatorizados; dos participantes en el grupo de dosis baja de ExPEC10V (Tabla 11), y un participante en cada uno en los grupos de ExPEC4V y Pevnar 13.

Paso 2: Veinticuatro participantes fueron enrolados y aleatorizados; 18 participantes en el grupo de dosis baja de ExPEC10V (Tabla 11), y tres participantes en cada uno en los grupos de ExPEC4V y Pevnar 13.

Paso 3: Cuatro participantes centinela fueron enrolados y aleatorizados; dos participantes en el grupo de dosis media de ExPEC10V (Tabla 11), y un participante en cada uno de los grupos de ExPEC4V y Pevnar 13.

Paso 4: Veinticuatro participantes fueron enrolados y aleatorizados; 18 participantes en el grupo de dosis media de ExPEC10V (Tabla 11), y tres participantes en cada uno de los grupos de ExPEC4V y Pevnar 13.

Paso 5: Cuatro participantes centinela fueron enrolados y aleatorizados; dos participantes en el grupo de dosis alta de ExPEC10V (Tabla 11), y un participante en cada uno de los grupos de ExPEC4V y Pevnar 13.

Paso 6: Veinticuatro participantes fueron enrolados y aleatorizados; 18 participantes en el grupo de dosis alta de ExPEC10V (Tabla 11), y tres participantes en cada uno de los grupos de ExPEC4V y Pevnar 13.

Todos los participantes recibieron una única inyección intramuscular (IM) de ExPEC10V (1 de 3 dosis), ExPEC4V o Pevnar 13 el Día 1 según los grupos de vacunación del estudio asignados. Los cuatro participantes centinela de cada uno de los Pasos 1, 3 y 5 fueron contactados por teléfono 24 horas después de la vacunación para recolectar la información de seguridad. Los datos de seguridad 24 horas después de la vacunación bajo ciego de cada grupo de los cuatro participantes centinela fueron revisados por el investigador principal (PI), el médico responsable del estudio (SRP) y el líder médico del patrocinador (SML). La aleatorización de participantes adicionales para el siguiente paso se detuvo hasta que se completara esta evaluación de seguridad centinela del Día 2.

En ausencia de cualquier hallazgo clínicamente significativo, otros 24 participantes (para los Pasos 2, 4 y 6) fueron enrolados y aleatorizados a uno de tres grupos de vacunación del estudio (Tabla 11) para recibir una única inyección IM de ExPEC10V (1 de 3 dosis), ExPEC4V o Pevnar 13 el Día 1.

Después de la vacunación de 24 participantes adicionales en cada nivel de dosis (dosis baja en el Paso 2, dosis media en el Paso 4, y dosis alta en el Paso 6), los datos de seguridad 14 días después de la vacunación de los 28 participantes (4+24) en cada nivel de dosis fueron revisados por el DRC antes de pasar al siguiente nivel de dosis o la Fase 2a.

Cohorte 1: Fase 2a

Sobre la base de una seguridad y reatogenicidad aceptables (en ausencia de cualquier problema de seguridad o cualquier evento que cumpla con una regla de detención del estudio específica) según lo determinado por el DRC después de la revisión de los datos de seguridad 14 días después de la vacunación para los 84 participantes iniciales, los 320 participantes restantes de la Cohorte 1 son aleatorizados y dosificados en la Fase 2a del estudio. Estos 320 participantes adicionales fueron enrolados y aleatorizados en paralelo en una proporción de 2:2:2:1:1 a uno de los

cinco grupos de vacunación del estudio para recibir una única inyección IM de ExPEC10V (1 de 3 dosis), ExPEC4V o Pevnar 13 el Día 1 (Tabla 11).

5 Además de realizar la revisión de seguridad de los 14 días para los 84 participantes iniciales, el DRC también evalúa los datos de seguridad de la Cohorte 1 en el transcurso del estudio y revisa cualquier evento que cumpla con una regla específica de detención de la vacunación del estudio o cualquier otro problema de seguridad que pueda surgir.

10 Para la Cohorte 1, el análisis primario se produce cuando todos los participantes hayan terminado la visita del Día 30 (Visita 4) o hayan discontinuado el estudio anticipadamente. El análisis final se produce cuando todos los participantes hayan terminado la visita del Día 181 o hayan discontinuado el estudio anticipadamente. Para los participantes que progresan al período de seguimiento a largo plazo (LTFU) de etiqueta abierta (grupo de dosis seleccionada de ExPEC10V y grupo de Pevnar 13), los análisis de seguimiento anuales incluyen datos de seguridad e inmunogenicidad (inmunoensayo basado en ECL multiplex y MOPA) recogidos hasta al momento de la visita del Año 1 (Día 366), Año 2 (Día 731) y Año 3 (Día 1096) después de la vacunación.

Cohorte 2

15 En la Cohorte 2, la seguridad, la reactogenicidad y la inmunogenicidad de la dosis seleccionada de ExPEC10V (basado en los resultados del análisis primario de la Cohorte 1) se evalúan en participantes de ≥ 60 años con salud estable y antecedentes de UTI en los últimos 5 años. Para la Cohorte 2, el estudio tiene un diseño a doble ciego, controlado con placebo, y un total de 600 participantes son enrolados y aleatorizados en paralelo en una proporción de 2:1 (400 participantes en el grupo de ExPEC10V y 200 en el grupo del placebo).

20 Todos los participantes reciben una única inyección IM de la dosis seleccionada de ExPEC10V o placebo el Día 1 según los grupos de vacunación del estudio asignados (Tabla 11).

25 Para la Cohorte 2, el análisis primario incluye datos de seguridad e inmunogenicidad, y se produce cuando todos los participantes hayan terminado la visita del Día 30 (Visita 4) o hayan discontinuado el estudio anticipadamente. El análisis final se produce cuando todos los participantes hayan terminado la visita del Día 181 o hayan discontinuado el estudio anticipadamente. Para todos los participantes, los análisis de seguimiento anuales incluyen datos de seguridad e inmunogenicidad (inmunoensayo basado en ECL multiplex y MOPA) recogidos hasta el momento de la visita del Año 1 (Día 366), Año 2 (Día 731) y Año 3 (Día 1096) después de la vacunación.

Se realiza un análisis de muestra de materia fecal en un subconjunto seleccionado de los participantes para evaluar el efecto de ExPEC10V sobre la prevalencia de patógenos (por ejemplo, *Clostridium difficile*) y serotipos de ExPEC10V en la flora intestinal utilizando metagenómica.

30 NÚMERO DE PARTICIPANTES

Un total de 1004 participantes son enrolados en el estudio; 404 participantes en la Cohorte 1 y 600 participantes en la Cohorte 2.

GRUPOS DE INTERVENCIÓN

Descripción de las intervenciones

35 **ExPEC10V:** Vacuna bioconjugada de *E. coli* en solución tamponada con fosfato que contiene PS O–antígeno de ExPEC serotipos O1A, O2, O4, O6A, O8, O15, O16, O18A, O25B y O75 bioconjugados por separado a la proteína portadora EPA. Una sola inyección de 0.5 ml IM (deltoides) de una de las tres dosis de ExPEC10V el Día 1.

40 **ExPEC4V:** Vacuna bioconjugada de *E. coli* en solución tamponada con solución salina que contiene PS O–antígeno de ExPEC serotipos O1A, O2, O6A, O25B (4:4:4:8 μ g PS/serotipos de ExPEC) bioconjugadas por separado a la proteína portadora EPA. Una sola inyección de 0.5 ml IM (deltoides) de ExPEC4V el Día 1.

Pevnar 13: Suspensión estéril de sacáridos de los antígenos capsulares de *Streptococcus pneumoniae* serotipos 1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A, 19F y 23F, unidos individualmente a la proteína de difteria CRM197 no tóxica. Única inyección de 0.5 ml IM (deltoides) el Día 1, suministrada en una jeringa precargada monodosis.

Placebo: Solución salina normal. Una sola inyección de 0.5 ml IM (deltoides) de placebo el Día 1.

45 Los materiales de intervención del estudio de ExPEC se describen en la Tabla 9.

Tabla 9. Vacunas del estudio de BAC1001MV ExPEC

Brazo del estudio	O1A (µg)	O2 (µg)	O4 (µg)	O6A (µg)	O8 (µg)	O15 (µg)	O16 (µg)	O18A (µg)	O25B (µg)	O75 (µg)	EPA (µg)	PS (Total) (µg)
Dosis baja de ExPEC10V	4	4	4	4	4	4	4	4	8	4	160	44
Dosis media de ExPEC10V	8	4	4	8	4	4	4	4	16	4	221	60
Dosis alta de ExPEC10V	8	8	8	8	8	8	8	8	16	8	320	88
ExPEC10V	4	4	-	4	-	-	-	-	8	-	72	20
ExPEC4V												

EPA = una forma genéticamente desintoxicada de la exotoxina A derivada de *Pseudomonas aeruginosa*; PS = polisacárido.
 ExPEC4V se compone de los O-antígenos polisacáridos (PS) de ExPEC serotipos O1A, O2, O6A y O25B bioconjugados por separado a la proteína portadora EPA.
 ExPEC10V se compone de los O-antígenos polisacáridos (PS) de ExPEC serotipos O1A, O2, O4, O6A, O8, O15, O16, O18A, O25B y O75 bioconjugados por separado a la proteína portadora EPA.
 La dosis se basa solo en el PS. El EPA (µg) son valores medidos.

ExPEC10V se compone de 10 sustancias farmacológicas (DS) monovalentes. Para este estudio clínico, se producen 2 concentraciones diferentes (media y alta) de producto farmacológico (DP) (Tabla 10). Una tercera concentración (baja) se obtiene en la clínica por dilución de la concentración alta 1:1 con tampón de dilución, que es el mismo que el tampón de formulación. Cada DP se formula en tampón de fosfato de sodio / potasio a pH 7.0 (0.02% [p/p] Polisorbato 80, 5% [p/p] sorbitol, metionina 10 mM).

Tabla 10: Composición de la vacuna ExPEC10V para el estudio clínico de fase 1/2a

Ingrediente	Cantidad (µg/ml)^a		
Activo^a	Concentración baja^b	Concentración media	Concentración alta
O-antígeno polisacárido			
EcoO1A	8	16	16
EcoO2	8	8	16
EcoO4	8	8	16
EcoO6A	8	16	16
EcoO8	8	8	16
EcoO15	8	8	16
EcoO16	8	8	16
EcoO18A	8	8	16
EcoO25B	16	32	32
EcoO75	8	8	16
Proteína portadora			
EPA	320	441	640
Excipientes			
KH ₂ PO ₄	6.19 mM		
Na ₂ HPO ₄	3.81 mM		
Sorbitol	5% (p/p)		
Metionina	10 mM		
Polisorbato 80	0.02% (p/p)		

EPA = exotoxina A de *P. aeruginosa* genéticamente desintoxicada utilizada como proteína portadora.

^a El ingrediente activo es un conjugado biológicamente sintetizado compuesto por el antígeno PS y una proteína portadora (EPA); la dosis se calcula sobre la fracción de PS solamente.

^b La "concentración baja" se obtiene en la clínica por dilución de la "concentración alta" 1:1 con tampón de dilución.

EVALUACIONES DE SEGURIDAD

Las evaluaciones de seguridad clave incluyen AE solicitados locales y sistémicos, AE no solicitados, SAE, exámenes físicos, mediciones de signos vitales y análisis de laboratorio clínico.

EVALUACIONES DE INMUNOGENICIDAD

Las evaluaciones de inmunogenicidad clave de los sueros recogidos incluyen la evaluación de los niveles de anticuerpos IgG totales específicos de serotipo de ExPEC10V y ExPEC4V producidos por la vacuna tal como se mide por un inmunoensayo basado en ECL multiplex, y anticuerpos funcionales específicos de serotipo de ExPEC10V y ExPEC4V tal como se mide por un ensayo de destrucción opsonofagocítica (OPKA) en formato multiplex (MOPA). No se realizan evaluaciones de inmunogenicidad de los títulos de anticuerpos neumocócicos provocados por Prevnar 13.

Los niveles de anticuerpos séricos inducidos por ExPEC10V se miden mediante un inmunoensayo basado en electroquimioluminiscencia (ECL) multiplex. Este ensayo combina electrodos de carbono de alta unión en un formato multipunto de microplaca de 96 pocillos que está recubierto con diferentes O-LPS antígenos de *E. coli* o la proteína portadora EPA. Los niveles de anticuerpos antígeno-específicos presentes en las muestras de suero se detectan usando un anticuerpo secundario (IgG anti-humano) marcado con SULFO-TAG. SULFO-TAG emite luz en presencia de estimulación eléctrica a una intensidad que aumenta proporcionalmente a la cantidad de anticuerpos IgG unidos. Este ensayo fue calificado de acuerdo con las recomendaciones de la Conferencia Internacional de Armonización (ICH).

Los niveles de anticuerpos funcionales inducidos por ExPEC10V se miden mediante un ensayo de opsonofagocítico multiplex (MOPA). Brevemente, las muestras de suero inactivadas por calor se diluyen en serie y se incuban con diferentes cepas de *E. coli* que son específicamente resistentes a diferentes tipos de antibióticos. Después de eso, el complemento humano y las células fagocíticas (HL60) se añaden a la reacción y, después de un segundo periodo de incubación, una alícuota de la mezcla de reacción se transfiere a diferentes placas de filtro de membrana hidrófila de PVDF que contienen medio suplementado con antibiótico específico que permiten selectivamente el crecimiento de

una cepa que es resistente a ese antibiótico particular. Después del cultivo durante la noche, se cuentan las unidades formadoras de colonias (UFC) para determinar el número de bacterias supervivientes. Este ensayo fue calificado de acuerdo con las recomendaciones de la ICH.

5 Para los anticuerpos de serotipo ExPEC10V medidos por el inmunoensayo basado en ECL multiplex y MOPA, y EPA medido por el inmunoensayo basado en ECL multiplex solamente, se evalúan las siguientes medidas de inmunogenicidad y se tabulan por grupo de vacunación del estudio, para todos los puntos temporales de inmunogenicidad:

– proporción de participantes con un aumento ≥ 2 veces y ≥ 4 veces en los títulos de anticuerpos en suero desde el Día 1 (antes de la vacunación)

10 – título por media geométrica (GMT)

– GMR: veces de cambio desde la línea basal, calculado a partir del valor posbasal / basal.

Para el período LTFU, se proporcionan resúmenes descriptivos de inmunogenicidad para cada serotipo.

La selección de dosis para las fases posteriores considera la totalidad de la evidencia disponible al momento del análisis principal de la Cohorte 1 (resultados del Día 30).

Tabla 11: Cohorte 1: Cronograma de vacunación

Grupo de vacunación del estudio	Fase 1				Fase 2a			Total
	Paso 1	Paso 2	Paso 3	Paso 4	Paso 5	Paso 6	Paso 7	
G1	Dosis baja de ExPEC10V*	18					80	100
G2	Dosis media de ExPEC10V*		2	18			80	100
G3	Dosis alta de ExPEC10V*				2	18	80	100
G4	ExPEC4V**	1	3	1	1	3	40	52
G5	Prevnar 13***	1	3	1	1	3	40	52
Total		4	24	4	4	24	320	404

* ExPEC10V se compone de los O-antígenos polisacáridos (PS) de ExPEC serotipos O1A, O2, O4, O6A, O8, O15, O16, O18A, O25B y O75 bioconjugados por separado a la proteína portadora, una forma genéticamente desintoxicada de exotoxina A (EPA) derivado de *Pseudomonas aeruginosa*.

** ExPEC4V se compone de los O-antígenos polisacáridos (PS) de ExPEC serotipos O1A, O2, O6A y O25B bioconjugados por separado a la proteína portadora, una forma genéticamente desintoxicada de exotoxina A (EPA) derivada de *Pseudomonas aeruginosa*.

*** Pevnar 13, una vacuna conjugada 13-valente neumocócica (proteína diftérica CRM197), es una suspensión estéril de sacáridos de los antígenos capsulares de *Streptococcus pneumoniae* serotipos 1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A, 19F y 23F, unidos de forma individual a la proteína de difteria CRM197 no tóxica.

Tabla 11: Cohorte 2: Cronograma de vacunación

Grupos de vacunación del estudio	Vacunación el Día 1	Total
G6	ExPEC10Va	400
G7	Placebo	200
Total		600

^a ExPEC10V se compone de los O-antígenos polisacáridos (PS) del ExPEC serotipos O1A, O2, O4, O6A, O8, O15, O16, O18A, O25B y O75 bioconjugados por separado a la proteína portadora, una forma genéticamente desintoxicada de exotoxina A (EPA) derivada de *Pseudomonas aeruginosa*.

5 La relación de aleatorización para los participantes enrolados en la Cohorte 2 del estudio es de 2:1 (ExPEC10V: Placebo). La dosis de ExPEC10V utilizada en la Cohorte 2 se basa en los resultados del análisis primario (Día 30) de la Cohorte 1.

ESTADO

El enrolamiento y la vacunación de la Cohorte 1 del estudio descrito anteriormente se completó. El estudio está en curso en forma ciega. Sobre la base de una revisión continua de los datos de seguridad, no se identificaron problemas de seguridad importantes, y la vacuna ExPEC10V tiene un perfil de seguridad aceptable.

10 El análisis de la inmunogenicidad de las muestras clínicas de la Cohorte 1 está en curso de manera ciega. Los datos de ECL fueron 100% chequeados dentro de los Límites de Aceptación de Calidad (AQL) y subidos para la gestión de datos. El análisis de las muestras MOPA está en curso. La apertura del ciego de los datos y el análisis estadístico se lleva a cabo mediante el uso de una organización de investigación clínica (CRO).

15 Las vacunaciones en la Cohorte 2 se inician una vez que la dosis de ExPEC10V para esa cohorte se haya identificado basado en el análisis primario finalizado de los resultados del Día 30 a partir de la Cohorte 1.

SECUENCIAS

SEQ ID NO: 1 (secuencia de consenso de glicosilación)

Asn-X-Ser(Thr), en donde X puede ser cualquier aminoácido excepto Pro

SEQ ID NO: 2 (secuencia de consenso de glicosilación optimizada)

20 Asp(Glu)-X-Asn-Z-Ser(Thr), en donde X y Z se seleccionan independientemente entre cualquier aminoácido excepto Pro

SEQ ID NO: 3 (proteína portadora EPA que comprende 4 secuencias de consenso de glicosilación (EPA-4))

G SGGGDQ**NATG SGGG**KLAEEA FDLWNECAKA CVLCLKDGVR SSRMSVDP AI ADTNGQGV LH
 YSMVLEGGND ALKLAIDNAL SITS DGLTIR LEGGVEPNKP VRYSYTRQAR GSWSLNWLVP IGHEKPSNIK
 VFIHELNAGN QLSHMSPIYT IEMGD ELLAK LARDATFFVR AHESNEMQPT LAISHAGVSV VMAQAQPRRE
 KRWSEWASGK VLCLLDPLDG VYNYLAQQRC NLDDT WEGKI YRVLAGNPAK HDLDIK**DNNN** STPTVISHRL
 HFPEGGLAA LTAHQACHLP LEAFTRHRQP RGWEQLEQCG YPVQRLVALY LAARLSWNQV
 DQVIRNALAS PGSGGDLGEA IREQPEQARL ALTAAAEESE RFVRQGTGND EAGAASADV
 SLTCPVAK**DQ** **NRTK**GECAGP ADSGDALLER NYPTGAEFLG DGGDVSFSTR GTQNWTVERL
 LQAHRQLEER GYVFGYHGT FLEAAQSIVF GGVRARSQDL DAIWRGFYIA GDPALAYGYA
 QDQEPDARGR IRNGALLRVY VPRWSLPGFY RTGLTLAAPE AAGEVERLIG HPLPLRLDAI TGPEEEGGRV
 TILGWPLAER TVVIPS IPT DPRNVGGDL D PSSIPDKEQA ISALPDYASQ PGKPPREDLK **LGSGGG**DQ**NA**
T

SEQ ID NO: 4 (secuencia de aminoácidos de O4 GtrS)

25 MNNLIMNNWCKLSIFIIFILLWLRPPDILTNAQFWAEDSVFWYKDAYENGFLSSLTTPRNGYFQTVSTFIVGLTALLNP
 DYAPFVSNFFGIMIRSVIIFLFTERFNFLTLTTRIFLSIYFLCMPGLDEVHANITNAHWYLSLYVSMILIARNPSSKSWRF
 HDIFFILLSGLSGPFIIFILAASCFFKFINNCKDHISVRSFINFYLRQPYALMIVCALIQGTSIILTFNGTRSSAPLGFSDVISSI
 ISSNIFLFTFVPWDIAKAGWDNLLLSYFLSVSILSCAAFVFKGTWRMKVFATLPLLIIFSMAPQLTDSAPQLPTLINGQ
 GSRYFVNIHIAIFSLLCVYLLECVRGKVATLFSKIYLTILLFVMGCLNFVITPLPNMNWREGATLINNAKTGDVISIQVLP
 GLTLELRKK

SEQ ID NO: 5 (Ejemplo de secuencia de ácido nucleico de O4 *gtrS*)

ATGAATAATTTAATTATGAATAACTGGTGTAATTATCTATATTTATTATTGCATTTATTTTGGCTATGGCTTAGAAGG
 CCGGATATACTCACAAACGCACAATTTTGGGCAGAAGATCCCGTTTTCTGGTATAAGGACGCCTATGAGAACGGA
 5 TATTAATCCAGATTATGCACCTTTTGTCTAATTTTTTGGCATAATGATTTCGCTCAGTAATTATATGGTTTTTAT
 TTACAGAAAGATTCAACTTCCTCACATTGACTACTAGGATTTTCTTATCTATTTATTTTCTATGCATGCCTGGATTG
 GATGAAGTTCATGCAAATATAACAAATGCACATTGGTATTTGTCATTATATGTATCAATGATCCTGATAGCTCGCA
 ATCCAAGTTCAAAATCATGGAGGTTTCATGATATATTCTTTATCTTGCTATCCGGGCTCAGTGGCCATTTATAAT
 10 TTTTCATTTTAGCAGCTTCATGCTTTAAATTTATAAATAATTGTAAAGATCATATTAGTGTAAGATCTTTCATAAATTT
 CTACTTGGCTCAGCCATACGCATTAATGATTGTTTGGCCTTTAATTCAAGGAACCTTCTATAATTCTAACTTTCAATG
 GCACACGTTCCCTCAGCACCCTAGGATTCAGTTTTGATGTGATTTTCGCTATTATATCATCGAATATTTTTTATTT
 ACATTTGTCCATGGGATATTGCAAAGGCTGGTGGGATAATTTACTGTTATCTTATTTTTTGTCTGTTTCGATTTT
 15 GTCGTGTGCGGCCTTTGTTTTGTTAAAGGTACGTGGCGAATGAAAGTATTTGCAACTTTACCATTGCTAATTATA
 ATATTTTCAATGGCAAACCACAATTGACAGACTCGGCACCTCAATTGCCAACACTTATTAATGGGCAAGGTTCAA
 GATACTTCGTAATATACATATTGCGATATTCTTTGCTATGTGTTACTTACTTGAGTGCCTCAGGGGAAAGT
 GGCAACTTTATTTTCCAAAATATACTTAACAATTTTGTATTTCGTGATGGGATGTTTGAATTTTGTATCACCCAC
 TCCCAAACATGAACTGGAGGGAAGGTGCTACTTTGATTAATAATGCAAAAACCTGGTGTATGTCATTTGATTCAAG
 TGCTACCACCTGGCCTAACACTTGAACCTAAGGAAAAAATAA

SEQ ID NO: 6 (Ejemplo de secuencia de PglB ('de tipo salvaje'))

MLKKEYLKNPYLVLFAMILAYVFSVFCRFYVWVWVASEFNEYFFNQLMIISNDGYAFAEGARDMIAGFHQPNDLSYY
 GSSLSALTYWLYKITPFSFESIILYMSTFLSSLVVIPTILLANEYKRPLMGFVAALLASIANSYNRTMSGYYDMDLVIVL
 20 PMFILFFMVRMILKKDFFSLIALPLFIGIYLWYWPSSYTLNVALIGLFLIYTLIFHRKEKIFYIAVILSSLTLSNIAWFYQSAIIV
 LFALFALEQKRLNFMIIIGLSATLIFLILSGGVDPIYLKQFYIFRSDSANLTQGFMYFNVNQTIQEVENVDLSEFMRRR
 SGSEIVFLFSLFGFVWLLRKHKSMIMALPILVLGFLALKGGLRFTIYVPMALGFGFLLSEFKAIMVKKYSQLTSNVCIV
 25 FATILTLAPVFIHIYNYKAPT VFSQNEASLLNQLKNIANREDYVVTWWDYGYPVRYYSVDVKTLDVGGKHLGKDNFFPSF
 ALSKDEQAAANMARLSVEYTEKSFYAPQNDILKTDILQAMMKDYNQSNVDLFLASLSKPDFKIDTPKTRDIYLYMPARM
 SLIFSTVASFSFINLDTGVLDKPFSTAYPLDVKNGEIYLSNGVVLSDDFRSFKIGDNVSVNSIVEINSIKQGEYKITPID
 DKAQFYIFYLKDSAIPYAQFILMDKTMFNSAVYQMFFLGNVDKLNFLDLVINSRDAKVFKLKI

SEQ ID NO: 7 (Ejemplo de secuencia de aminoácidos de *gtrA*; *E. coli* W3110 yfdG, GenBank: BAA16209.1)

MLKLFKYTSIGVLNLIHWVVFVGVCIYVAHTNQALANFAGFVAVSFSFFANAKFTFKASTTTMRYMLYVGFMTLSA
 30 TVGWAADRCALPPMITLVTFSAISLVCGFVYSKFIVFRDAK

SEQ ID NO: 8 (Ejemplo de secuencia de aminoácidos de *gtrB* -*E. coli* W3110 yfdH, GenBank: BAA16210.1)

MKISLVVPVFNEEEAIPIFYKTVREFEELKSYEVEIVFINDGSKDATESIINALAVSDPLVVPLSFRNFGKEPALFAGLDH
 ATGDAIIPIDVDLQDPIEVPHLIEKWQAGADMVLAKRSDRSTDGRLKRKTAEWFYKLNKISNPKIEENVGDFRLMSRD
 35 VVENIKLMPERNLFMKGILSWVGGKTDIVEYVRAERIAAGDTKFNGWKLWNLALEGITSFSTFPLRIWYIIGLVVASVAFI
 YGAWMILDITIFGNAVRGYPSLLVSILFLGGIQMIGIVLGEYIGRTYIETKKRPKYIIRVKK

SEQ ID NO: 9 (Ejemplo de secuencia de nucleótidos de locus O4 *rfb* - O4-producción EPA cepa BVEC-L-00684f)

ATGACGAATTTAAAAGCAGTTATTCTGTAGCGGGTCTCGGGATGCATATGTTGCTGCCACTAAGGCGATACCC
 AAAGAGATGCTACCAATCGTCGACAAGCCAATGATTACAGTACATTGTTGACGAGATTGTGGCTGCAGGGATCAAA
 40 GAAATCCTCCTGGTAACTCACGCGTCCAAGAACGCGGTGAAAACCACTTCGACACCTCTTATGAGTTAGAATCA
 CTCCTTGAGCAGCGCGTGAAGCGTCAACTGCTGGCGGAAGTACAGTCCATCTGTCCGCCGGGCGTGACCATTA
 TGAACGTGCGTCAGGGCGAACCTTTAGGTTTAGGCCACTCCATTTTGTGTGCGCGACCTGCCATTGGTGACAAC
 CCATTTGTCTGGTACTGCCAGACGTTGTGATCGACGATGCCAGCGCCGACCCGCTACGTTACAACCTTGCTGC
 45 CATGATTGCACGTTTCAACGAAACGGCCGACCCAGGTGCTGGCAAAACGATGCCGGGTGACCTCTCTGAAT
 ACTCCGTCATCCAGACTAAAGAGCCGCTGGACCGTAGGGTAAAGTACGCCGATTGTTGAATTTATCGAAAAA
 CCGGATCAGCCGACGCTGGACTCAGACATCATGGCCGTAGGTCGCTATGTGCTTTCTGCCGATATTTGGCC
 GGAACCTGGAACGTAAGCTGCTGATGACCGGCGACAGTTACGACTGCGGCAAAAAAATGGGCTATATGCAGGC
 50 GTTTGTGAAGTATGGCCTACGCAACCTGAAAGAAGGGGCGAAGTCCGTAAGGATTGAGAAGCTGTTAAGCG
 AATAATGAAAATCTGACCGGATGTAACGGTTGATAAGAAAATATAACGGCAGTGAAAATTCGCAGCAAAAAGTAA
 TTTGTTGCGAATCTTCCGCTGTTGTTTTATATAAACCAATCAGATAACAACGAGTTAGCAGTAGGGTTTTATTCA
 55 AAGTTTTCCAGGATTTTCTTTGTTCCAGAGCGGATTGGTAAGACAATTAGCGTTTGAATTTTCCGGTTTAGCGC
 GAGTGGGTAACGCTCGTCACATCATAGGCATGCATGCAGTGCTCTGGTAGCTGTAAAGCCAGGGGCGGTAGCG
 TGCATTAATACCTCTATTAATCAAACCTGAGAGCCGCTTATTTACAGCATGCTCTGAAGTAATATGGAATAAATTA
 AGTGAATAACTTGTACTGGTGGCGCAGGATTTATTGGTTCAGCTGTAGTTCGTCACATTATAAATAATACGCAG
 GATAGTGTGTTAATGTGATAAATTAACGTACGCCGAAACCGGGAATCACTTGCTGATGTTTCTGATTCTGAA
 CGCTATGTTTTTGAACATGCGGATATTTGCGATGCACCTGCAATGGCACGGATTTTTGCTCAGCATCAGCCGGAT

GCAGTGATGCACCTGGCTGCTGAAAGCCATGTTGACCGTTCAATTACAGGCCCTGCGGCATTTATTGAAACCAA
TATTGTTGGTACTTATGTCCTTTTGAAGCCGCTCGCAATTACTGGTCTGCTCTTGATAGCGACAAGAAAAATAG
CTTCCGTTTTTCATCATATTTCTACTGACGAAGTATATGGTGATTGCTCCTCATCCTGACGAGGTAAATAACAGAA
5 GAATTACCCTTATTTACTGAGACAACAGCTTACGCGCCAAGCAGCCCTTATCCGCATCCAAAGCATCCAGCGAT
CATTTAGTCCGCGCGTGGAAACGTACCTATGGTTTACCACCATTGTGACTAATTGCTCTAACAATTATGGTCTCT
ATCATTTCGGGAAAAAATTGATTCCATTGGTTATTCTCAATGCTCTGGAAGGTAAAGCATTACCTATTTATGGTAA
AGGGGATCAAATTCGCGACTGGCTGTATGTTGAAGATCATGCGCGTGCCTGTTATATACCGTCTGTAACCGAAGGTA
AAGCGGGTGAACCTTATAACATTGGTGGGCACAACGAAAAGAAAAACATAGATGTAGTGCTCACTATTTGTGATT
10 TGCTGGATGAGATTGTACCGAAAGAGAAAATCTTATCGTGAGCAAATCACTTATGTTGCCGATCGTCCGGGACAC
GATCGCCGTTATGCGATTGATGCTGAGAATATTGGTCGCGAATTGGGATGAAACCACAGGAAACGTTTGAGAG
CGGGATTCCGAAGACAGTGGAATGGTATCTGTCCAATACAAAATGGGTTGATAATGTGAAAAGTGGTGCCATCA
ATCGTGGATTGAAGAGAACTATGAGGGCCGCCAGTAATGAATATCCTCCTTTTTGGCAAAACAGGGCAGGTAGG
TTGGGACTACAGCCGTGCTGTCGACCTCTGGGTAACCTTGGTCTTGTGATGTTTCACTCCACTGATTATTGTGG
CGATTTACAGTAAACCGAAGGTGGCTGAAACCGTCAAAAAAATTCGCCAGATGTTATTGTTAATGTTCTGTGC
15 TCATACCGCGGTAGATAAGGCTGAGTGCAGAACCGAATTTGCACAATTACTCAATGCGACCAGCGTTGAAGCAAT
TGCAAAAGCGGCTAATGAAGTTGGGGCTTGGGTAATTCATTACTCAACTGACTACGCTTCCCTGGAAATGGCGA
CATGCCATGGCTCGAGACTGATGTAACCGCTCCGCTCAATGTTTATGGCAAAACCAAATTGGCTGGAGAAAAGAG
CATTACAAGAACATTGCGCAAAGCATCTTATTTCCGTACCAGCTGGGTATATGCAGGTAAAGGAAATAACTTTG
CCAAAACAATGTTACGTCTGGCAAAAGAGCGCGAAGAAGTGGCTGTGATAAACGATCAGTTTGGCGCACCAACA
20 GGTGCTGAATTGCTGGCTGATTGCACCGCTCATGCCATTCCGCTGGCATTAAAAAACCAAGATTGCTGGCTT
GTACCATTGGTAGCAAAATGGCACAACCACTGGCAGGATTACGCGCGCTAGTATTTCGAAGAAGCCCGTAAAG
CAGGGATTGACCTTGCACCTTAACAACTCAACGCGTACCAACAACGGCTTATCCTACTCCAGCCCGCCGCTCCT
CATAATTCTCGCCTCAATACCGAAAAGTTTCAGCAGAAGTTTGGCCTTGTCTTGCCTGACTGGCAGGTGGGCGT
GAAACGTATGCTCAACGAATTTTACGACTACGGCAATTTAACAAATTTTTGCATCTCGCTCATGATGCCAGAGC
25 GGGATGAATTAAGGAATGGTGAATGAAAACGCGTAAAGGTATTATTCTGGCTGGTGGTTCCGGCACTCGTC
TTTATCCTGTGACGATGGCAGTGAGTAAACAAGTCTGCCGATTTATGATAAGCCGATGATTTATTATCCGCTTTC
AACGCTTATGTTAGCGGGTATTCGCGATATTCTTATTATCAGTACGCCACAGGATACACCGCGTTTCCAACAATT
GTTGGGGACGGGAGTCAGTGGGGGCTTAATCTACAGTATAAAGTACAACCGAGTCCGGATGGCCTGGCGCAA
CGTTTTATTATTGGTGAAGACTTTATTGGTGGTGATGATTGTGCACTCGTACTTGGCGATAATATCTTATGGAC
30 ACGACTTGCCGAAATTAATGGAAGCTGCTGTTAACAAAGAAATCGGTGCAACGGTATTTGCTTATCACGTCAATG
ATCCTGAACGTTATGGTGTCTGGAGTTTATAATAACGGTACTGCAATTAGCCTGGAAGAAAAACCGCTGGAAC
CAAAAAGTAAGTATGCGGTTACTGGGCTTTATTTCTATGACAATGATGTTGTAGAAATGGCGAAAAACCTTAAGCC
TTCTGCCCGTGGCGAACTGGAAATTACCGATATTAACCGTATTTATATGGAGCAGGGACGTTTGTCTGTGCTAT
GATGGGGCGTGGTTATGCCTGGTTGGATACTGGTACACATCAAAGTCTTATTGAAGCAAGTAACTTCAATGCCAC
35 CATTGAAGAGCGTCAGGGATTAAGGTATCTTGCCCGAAGAGATTGCTTACCCTAAAGGGTTTATTGATGCTGA
GCAGTGAAGTATTAGCCGAACCGTGAAGAAAAATGATTATGGTCAAGTATCTGCTAAAAATGATTAAGGTTA
TTAATAAAATGAACGTAATTAAGAACTGAAATTCCTGATGTGCTGATTTTTGAACCAAAAGTTTTTGGTGAACG
GGCTTCTTTTTGAGAGTTTTAACCGAAAAGTATTTGAAGAAGCTGTAGGACGGAAGGTTGAATTTGTTGAGGATA
ACCATTCTAAGTCTAAAAATAATGATTGCGTGGGATGCATTATCAAACACAAAATACTCAAGGAAAACCTGGTTCCG
40 GGTAAATTTCTGGTTCAGTATATGATGTTGCCGTAGATTTAAGAGAAAAATCAAAGACATTTGGCAAATGGGTGGG
TGTAAGAATTACTGGGAATAATAAAAGACAATTGTGGATCCCCGAAGGTTTTGCCCATGGTTTTATGTGTTGGAG
GAGAATACCGAATTTGTTTATAAATGTACCGATACTTAACCCGTCTCATGAACACACATTGCTATGGAATGATC
CAACTATCAATATAAGTTGGCAAATCATACAAACTGCAAGCCAATTTTCTGAAAAAGATGCTAATGGACATCT
TTTTTCACATAAAACCTATTTCTGAAATGCAATATTATGAGTTTAAATTAGAAACAGTTTCTATAATATTGCTGGTTTT
45 GCTGTGCCGACATTAGTTGCAGTCCCTGCTTTGGGGATTCTTGCCAGGCTGCTTGGACCGGAGAAATTTGGACT
TTTCACTAGCATTTCGCTTTGATAGGATATGCAAGTATTTTCGACGCCGGGATTAGTCGAGCTGTAATCAGAGA
AATCGCTCTTTATCGAGAAAGTAAAAAGAGCAAATACAAATTTTCGACAGCAAGTGAATCGTACTATTCTTA
GGGGTGGTTGCAGCTTTGTTACTTTATTTAGTAGTAATAAAGTTGTTGAGTTATTGAATGTTAGTTCCGTTTATAT
TGAACAGCAGTGCCTGCATTCTGTTATTTTCAATTTATAACTGTGTATCTGATTAACAGATTTGGCTTGGT
50 TATCTGGAAGGGCTAGAAAAATTTGCAAATATAAATGTTTCAGAGAATGATTTCTAGCACAAAGCTTGGCTATATTAC
CAGTGATATTTTTGTTATTACAATCCCTCGTTTGTCTTTATGCTATGATGGGTTGGTGGTTGGGCGTGTGATTTTCATT
TTTGATTAGCGCAATAATTTGTCGAGATATTATTCTTAAAAGTAAACTTTACTTTAATGTGGCAACTTGAATCGTC
TTATCTCTTTTGGTGGATGGATAACAGTTAGTAATATCATAAGCCCAATCATGGCATAATTTGACCGCTTTATCAT
CTCTCATATTATGGGGCTTCGAGAATTGCATTTTATACAGCGCCCTCAGAGGGTGTATCAAGGTTAATTAATATC
55 CCATATGCTTTGGCAAGAGCTCTATTTCCATAATTGGCATATGAATAATGATGATGAACGAAAAAATTACAAC
TACAGAGCTACGCAATTAAGCATTGTATGCTACCCATAGTTGTTATTGGTGTCAATTTTGCCTCATTCCATAATG
ACAACATGGATGGGAGATTATGCCTTAGAAGCAGCAACTATGAAATACTCTTGTGCTTTTTTTTACTTTA
ACTCTTTAGCGCAAAATACCTTATGCATACTTGCAATCTATCGGAAAGTCAAAAAATTACCGCATTTGTGCATCTCAT
AGAACTTGCGCCATACTTATTATTATTGATTACTTACAATGCATTTCCGCATAATTGGCACGGCAATCGCTTGG
60 TCACTTAGAACATTTTTGTGATTTTTGTTATACTACTTTTCGATATCGAGAAGAAAAATGATTGCGGTTGATATTGCGCTT
GCAACCTACAATGGTGTCAATTTTATTCGGCAACAGATTGAATCTATCCAGAAACAACTTATAGAAATTTGGCGTC
TTATAATAAGTATGATAACTCGAGTGATGATACTGTTGATATTATTAAGGATATGATGTCTAACGACAGTCGTAT
CTATTTGGTAGGAAATAAAGACAAGGAGGGTTATTCAGAACTTTAAATATGCTCTTTACAAAACCTACATCTGAA
ATTGTGTTACTATGTGACCAGGATGACATTTGGCCGGAGGAGCGTCTGAAAATCTTATAGATAAAATTAAGGCC

TTGCAGCGTAATGATTTTTGTTCCGGCAATGATGTTTACTGATTTGAAATTAGTAGACGAAAATAATTGTTTGATTG
CAGAAAGTTTTTATCGAACGAATAATATTAATCCACAAGATAATCTGAAAAATAATAATCTTCTCTGGCGTTCAACG
GTATATGGCTGTACTTGCATCATGAATAAGAACTTGTGATATTGCATTGCCTATACCTACATATGCACATATGC
ATGATCAATGTTGGCATTATTAGCGAAGCAATATGGTAACATTTTTTATTTCCGACTATGCGTCTGTTTCGTTATAG
5 GCAACATTCTACAAATGTTGTTGGTGGTAGAAATAAAACGCCATTTCAAAAAATTAATCCATACAAAAAACCTAA
AAAGGATTAATTTGCTAGTGGATAGAAGTGGCTTTAATTAATCAAATAACGATTTCTATCCAGGGAATAAAATG
GAAAAATAAATTGATTACTTAAAATTTGGAGTGAATGAAGTATTACCTTATCTTTTTAAAGGAAACAAGAAAGTTTT
TTCACCTTGTGATTAATTAGTTTTGGCATTACAAAATGATATATTTATTTTTTTTTGCACTGTTTATGATCTGTA
10 CGTTTTTAACACACAGGCGACAGGCATTATATGTTGTATCTGCGTTAGTATTTCTTTTTTTGGCTTAACTATCCA
TCAGGAGGGGACTGGATAGGTTATTTCTCCATTATGACTGCATGGTTAATGAGCAGTGAATAATGTTTATAA
TGTTTGAACCTGGATATGAATTAATTGTTTCTTATTTGGATATTTGGGATTTCCAGACAATTATTTTTTATAGCC
GCTGTAAATGTAATTAATATTAATTTTGCAAAGCATTGTTGAAAACGGAAGTTTTGTTATTGTTGCGATAATGTG
CATGTTCTTTGGAGTGTATGTTGAGGCCATTAGACAGGCTCTGGCCTTATCTATAGTTAATTTGGGATTCAT
15 TCTCTTTTTTTGGAGTAGAAAAAGGAAATTTATAACATTAGTATTTACCTTTTTCGTCACACTTTCCATAATACTGTTGATT
TGTTTTCTTCTAATGACTCCTCTATTTTCAAAGAAATTAAGCAAGATAATAAGTTATAGCCTATTAATTTTTAGTAGC
TTCTTTTTCGCTTTTTCTGAAACCATATTAAGTGCCTCTTGAATTTTCCAGAAGGATCCATTGCCAGTGAAA
AATTAAGTTTTTACTTAGCAACCGAGCAATACAGGCCACAGTTATCTATTGGGAGTGGCACTATTCTTGACATTAT
ACTTATTTTTCTGATATGTGAAGTTTTAAACGAATAAAGAAATATATGCTCGCTAATTATAATGCTGCAAATGAGA
TATTGCTTATTGGTTGCTGTCTTTATATTTCTTTCGGTATTTTTATCGGAAAATGATGCCAGTTATGACTCGCATT
20 GGTGGTATGTTTTCCATTTGTTATAGTACTCTTTATATTAACCTGGGTTATTGAGAATATTTAAGAGGTATATA
AATAAAAGAGGTTGGGTATAGCAAATTTAATTTGCTTTTTATTTTTGCTACAAAATTTGCGACCATTAAACATA
TGATTATAGCTATTATAATATAATGCACCAGGATACTTTGCTGAATAGGTTTGATGCATTAGATGATGCATCATTAA
GACAATCAGCGAAGAGAAAATGTTTCGATTTGGGAAAGATAGGATATGGTTTCTTATGTAGTATATAATATCCTGC
ATTCATTCCGATAATTTCTATGGAAGTGTCTTTGCTCTGTCTGTCTCATTGTTGAAATTTTATGTTAATAAGA
25 AGCTTTAGATAACCCTTAGGAAGTGTATGTTGATCTGTCCAAAAATTATATTGTAAGTGCAGCGCGCTGG
CTTCCGGAGGTGCATTAACCTATATTAAGCAATTTATAAAACATGCATCACAAAATTCAAATGACTATATTATGTTT
GTATCTGCGGGATTGGAGTTGCCGCTGTGATAACATCATTTACATAGAAAACACACCAAAAAGGATGGTTGAAA
AGAATATTTGGGATTGGTTCGGTTGTCGGAAGTTTATCTCGGAACATAAGATTAACGTTAAGAAAGTAATTTCTC
TACAAAATCCAGTTTGAATGTTTCTTACGAACAGATTTTACTTGCACCAGCCAATTCCTTTTTAGTAAAGTTGAT
30 TCTTTTTTAAAAAATATCACATCCGATAACGTAAGCTTTTTTATATAAAAAAGTTTTATTCTTATTTTATATTTAAT
ATGTGAATGCCAATACAACCATCGTAGTGCAAACGAATTGGATGAAAAAAGGAGTGTGAGCAATGTGATAAAA
TTAGTACCGAAAGGTCCTTGTATATAAACCTGATATCAAAGCATTTAATAATACTAATTTTATGATGTAGATATGGAT
GTATCTGCAAAAACACTCTTATATCCAGCGACACCCTTACCTATAAAAAATCATTGTTGTCATTCTGAAGGCGTTGG
TTATTTTAAAGAAAAAGTATTTTATAGATGATCTGAAATTCGAAGTACTTTTAAAAGAAATAGGTACAAAAATTT
35 GATAAGTTTGTGCAATTAATAACTTAAGCAAAAACGTTGATTATCTCGCGTTCTTTTACATACTCGAACTTGCAAA
AAAAATATAGCGGCATCTTTAATCGTTTTTCTAGCTATATCGAATCATATGGGTTACCACTCATCGAAGCTGC
TAGTTTAGGAAAAAATCATTAGTAGTGATCTTCTTATGCCCGGGATGTTTTAAAGGATTATAGCGGCGTAGAT
TTTGTAAATTTACAATAATGAAGATGGCTGGGCTAAGGCGTTGTTAATGTTTTAAATGGCAATTCGAAGCTCAAT
TTAGGCCTTATGAAAAAGATAGTCGTTTCTTGGCCACAGTTCTTCTCTATTTTGAATAAGGTGATTTATGTTTA
40 ATGGTAAAATATTGTTAATTACTGGTGGTACGGGGTCTTTCCGTAATGCTGTTCTAAGACGTTTTCTTGACACTGA
TATCAAAGAAATACGATTTTTTTCCCGGGATGAAAAAAAACAAGATGACATGAGGAAAAAATAATAATCCGAAA
CTTAAGTTCTATATAGGTATTTCCGCACTATTCGAGTATTCGATCTCAATGCTTCTCGAGGTGTTGATTTTATTA
TGCTGCAGCTCTGAAGCAAGTACCTTCTGCGAATTCACCCAATGGAAGCTGTAAAAACGAATGTTTTAGGTAC
GGAAAACGTAAGGAGCGCAATAGCTAATGGAGTTAGGCGAATTTGATGTTTGGAGTACAGATAAAGCTGTATA
45 TCCTATCAATGCAATGGGTATTTCCAAAGCGATGATGGAAAAAGTAATGGTAGCAAAAATCGCGCAATGTTGACTG
CTCTAAAACGGTTATTTGCGGTACACGTTATGGCAATGTAATGGCATCTCGTGGTTCCAGTTATCCCATTATTTGTC
GATCTGATTAATCAGGTAGACCAATGACGATAACAGACCCTAATATGACTCGTTTCATGATGACTCTCGAAGAC
GCTGTTGATTTGGTTCTTTACGCATTTGAACATGGCAATAATGGTATATTTTTGTCCAAAAGGCACCTGCGGCTA
CCATCGAAACGTTGGCTATTGCACTCAAAGAATTACTTAATGTAACCAACACCCTGTAATAATAATCGGCACCC
50 GACACGGGGAAAAACTGTACGAAGCGTTATTGAGCCGAGGAAATGATTGCAGCGGAGGATATGGGTGATTAT
TATCGTGTTCACCAGATCTCCGCGATTTGAACTATGGAAAAATGTGGAACATGGTGACCGTCTGATCTCGGAA
GTGGAAGATTATAACTCTCATAAATACTGATAGGTTAGATGTTGAGGGAATGAAAAAATTAAGTCTAAAACCTTCTT
TTATCCGGGCACTTCGGTCTGGTGAAGATTATGAGTTGATTTCATAATATGAAAATTTTAGTTACTGGCGCTGCA
GGGTTTATCGGTGCAAAATTTGGTATTCGGCTTAAGGAAGCTGGATATAACGAACCTACATACGATAGATCGTAAC
55 TCTTCTTTGGCGGATTTAGAGCAGGGACTTAAGCAGGCGAGATTTTATTTTTTACCTTGTCTGGGGTAAATCGTCCC
GTAAGGAGTGTGAATTTGAAGAGGGAAATAGTAATCTAACTCAACAGATTGTTGATACCTGAAAAAACCAATA
AAAATCTCTATCTGAGTTCTTCCATCCAGGCTGAATGTGATAACGCTTATGGAAAGTAAAGCAAGTA
CGGAAAAAATCATTACGAGTATGGGGAACGACAAACGCTAAATATTATTTTTATCGCTTGGCGAATGTATTCCG
GTAAGTGGTGTGACCAAATTATAACTCTTTTATAGCAACTTTCTGCCATCGCATTGCAAATGATGAAGCTATTAC
60 AATTAATGATCTTCAGCAGTTGTAATCTGGTGTATATAGATGACTTTTTGTTCTGACATATTAAGCTATTAGAAG
GAGCGAACGAACTGGTTACAGGACATTTGGTCCAATTTATCTGTTACTGTTGGTGAAGTGGCACAATTAATTTA
CCGGTTTAAAGAAAGTCGCCAAACATTAATCACCGAAGATGTAGGTAATGGATTTACACGTGCATTGTACTCAAC
ATGGTTAAGTTACCTGTCTCCTGAACAGTTTGGGTATACGGTTCCTTCTTATAGTGATGACAGAGGGGTATTCTGT
GAAGTATTGAAAACGAAAAACGCGGGCCAGTTTTCTTCTTACTGCGCATCCAGGAATTAAGTCTGGGGTGGTCAAT

TATCATCATTCCAAAAATGAGAAATTTATTGTCATCCGAGGAAGTGCTTGTTCCTTCAAATTTGAAAAATTGTCACGAG
 TGAACGATATGAACTTAATGTTTCCTCTGATGATTTTAAAATTGTTGAAACAGTTCGGGGATGGACGCATAACATT
 ACTAATAATGGCTCGGATGAGCTAGTTGTTATGCTTTGGCAAATGAAATATTTAATCGTTCTGAACCAGATACTA
 TAGCGAGAGTTTTATCGTGAAAAAATTGAAAGTCATGTCCGTTGTTGGGACTCGTCCAGAAATTTATTCGACTCTC
 5 GCGTGTCCCTTGCAAAATTAGATGAATATTGTGACCACCTTATTGTTTCATACCGGGCAAAACTACGATTATGAACTG
 AATGAAGTTTTTTCAAAGATTTGGGTGTTGCAAACTGATTAATTTCTTAATGCCGCAGGTAAAAATGCAGCAG
 AGACTATTGGACAAGTTATCATTAAAGTTGATGAGGTCTTTGAACAGGAAAAACCAGAAGCCATGTTAGTACTTG
 GCGATACTAACTCCTGTATTTTACGCAATACCAGCAAAGCGTCAAGAATTCCGATCTTCCATATGGAGGCTGGGA
 ATCGTTGTTTTGACCAACGCGTACCAGGAAAGAACTAACAGAAAAATAGTTGATCATAACCGCTGATATCAATATGA
 10 CATATAGTGATATCGCGCGTGAATATCTTCTGGCTGAAGGTGTACCAGCCGATAGAATTATTTAAAACCGGTAGCC
 CAATGTTTGAAGTACTCACTCATTATATGCCGCAGATTGATGGTCCGATGACTTTCTCGCCTGAATTTAACACC
 TGGGAATTTCTTTGTGGTAAGTGCCACAGAGAAGAAAAATGTTGATACCCCTAAACAACCTTGTGAAACTGGCGAA
 TATACTTAATACCGTGGCTGAAAAATATGATGTCCCGGTAGTTGTTTCTACTCATCTCGACTCGTAAACCGCATC
 AACGAAAACGGTATTCATTTCCATAAAAAATCTTGCTTTAAGCCATTAGGATTTTACAGTTTACAACCATGTTAG
 15 AAAAAATGCAGTGTCTGTTTTATCGGATAGTGGGACTATTACAGAAGAGTCTCCATTATGAACCTTCCCTGCAC
 TCAATATACGAGAAGCGCACGAACGCCGGAAGGCTTCAAGAAGGGCAGTAATGATGGTCCGCTTGAATCT
 GATCGCGTTTTACAGGCATTAGAAATATTGCAACACAGCCTCGTGGAGAAGTACGCTTACTTCGTCAGGTTAGT
 GACTATAGCATGCCAAATGTTTACAGATAAAGTCTGCGTATTATCCATTCATATACTGACTACGTTAAACGGGTTG
 TCTGGAAGCAATACTAATGAACTTGCATTAATCATTGATGATTTTGCCTCATAGCACACGCGTTGGGGCTAAA
 ATGTTTCATGAGTTAGGCTTGAATTACTGAGCAGAGCCATGATGTAACGTAACTGACCTGACATCTCATTAC
 AAGCAATTTATTCTATTAGTATGATTGATGGTATAAAGTTTGGCGTTTTCAAAGTGGACCTTTAAAGCATGTAGG
 20 TAAGGCTAAACGTGCCATAAATGAACTCTTTTTATCTTTTCGCGCATGGCGCGATTTAAGCACCTCATTCAACAT
 GATACATTTGATGGTATCGTTTATTATCCCCCTCTATTTTTGGGGCGACTTGGTTAAAAAATAAAACAACGAT
 GCCAGTGCCCAAGCTATCTGATCCTAAGGGATATGTTCCACAGTGGGTGATTGATGCAGGTATGTTGAAAGCC
 GGTTACCAATTGAAAAATATTTTAGGTATTTGAAAAAAGTCATATCAGCAGGCTGGCCGGATAGGGGTAATG
 25 TCTGATAAGAATCTTGAGATATTTGCGCCAGACCAATAAAGGTTATCCGTGTGAAGTTTTACGTAATTGGGCTCAA
 TGACTCCTGTGCTGCCAGCGATGATTATCATTCACTTCGTCAAAAAACGATCTAAAAGATAAAGTCAATTTTTTTC
 TATGGCGTAATATTGGGCATGCTCAGGATATGGCAAACCTTAATGCGCCTTGCAGTAAATGATGCGTTATCAT
 GATGCTCATTCTGTTTATAGGCGAGGGTATGAAGTTGAGCTGATAAAATCTCTTGCTGCAGAATGGAATTTA
 30 ACTAATTTCACTCATCTACCTTCAGTGAACCAGGAAGAGTTAAATTAATTTTATCTGAAGTTGATGTCGGCCTGT
 TCTCCCTTTTCACTCGCCATTCTTACATAATTTCCCGGAAAATTACTAGGGTATATGGTTCAATCAATCCCGAT
 CCTTGGGAGTGTGAATGGCGGCAATGATTTAATGGATGTAATTAATAAGCACAGAGCCGGTTTTCAATCATGTTAA
 TGGTGAAGATGATAAACTGTTTGAATCTGCACAATTGCTTCTTAGTGATTACAGTTTTAAGAAAAACAGCTAGGTCAG
 AACGCTAATGTGTTGTTAAAGTCTCAATTTTTCGGTTGAATCGGCGGCACATACTATCGAAGTCCGACTGGAGGCT
 35 GGAAATTTTATGTTGATGACAATATTCTGGATGAACCTTTTCGCACTGCAGCAAATTTGAACGTTTGCAGC
 CTGCAATTTTATGACGCATCTCATCAGGAGAAGGTTCAACGTTTACTTATTGCATTTGACGCGACAGCATGT
 TGAACCCCATTTGGCATGAGTTACCGCATCAGTGGGAAATGTTTGTGTCATGCAAGGGCAATTAAGAAGTTTGT
 GTATGAGCAAAAATGGTGAGATCCAAAAACAGTTTGTGTTGGAGACGGTACGGGAATAAGCGTGTGGAATTTTC
 CCCAGGAGATATACATAGTGTCAAATGCCTGTACCAAAAAGCCCTTATGTTGGAGATAAAGGAGGGGCCATTTG
 40 ACCCACTCAAAGCTAAGGCTTTTTCTAAGTGGTTATAGGGCGATACACCACCGTTTTATTCTTCTATCTTATTCTAT
 ACATGCTGGGTTACCATCTTAGCTTCTTCAAGCCGCGCAACCCCGCGGTGACCACCCCTGACAGGAGTAGCTAG
 CATTTGACCACCCCTGACAGGATTAGCTAGCATATGAGCTCGAGGATATCTACTGTGGGTACCCGGGATCCGTG
 TAGGCTGGAGCTGCTTCAAGTTCCTATACTTTCTAGAGAATAGGAACTTCGGAATAGGAACTAAGGAGGATATT
 CATAT

45 **SEQ ID NO: 10** (Ejemplo de secuencia de señal para proteína portadora EPA)

MKKIWLALAG LVLAFSASA

SEQ ID NO: 11 (Ejemplo de secuencia de nucleótidos de locus O1A *rfb* – O1A–producción EPA cepa stGVXN4411 y stLMTB10217)

ATGACGAATTTAAAAGCAGTTATTCTGTAGCGGGTCTCGGGATGCATATGTTGCCTGCCACTAAGGCGATACCC
 50 AAAGAGATGCTACCAATCGTCGACAAGCCAATGATTACAGTACATTGTTGACGAGATTGTGGCTGCAGGGATCAAA
 GAAATCCTCCTGGTAACTCACGCGTCCAAGAACGCGGTGCAAAACCACTTCGACACCTCTTATGAGTTAGAATCA
 CTCCTTGAGCAGCGCGTGAAGCGTCAACTGCTGGCGGAAGTACAGTCCATCTGTCCGCCGGGCGTGACCATTA
 TGAACGTGCGTCAGGGCGAACCTTTAGGTTTAGGCCACTCCATTTTGTGTGCGCGACCTGCCATTGGTGACAAC
 CCATTTGTGCTGGTACTGCCAGACGTTGTGATCGACGATGCCAGCGCCGACCCGCTACGTTACAACCTTGTGCTG
 55 CATGATTGCAGTTCACGAAACGGGCCGAGCCAGGTGCTGGCAAAACGATGCCGGGTGACCTCTCTGAAT
 ACTCCGTCATCCAGACTAAAGAGCCGCTGGACCGTGAGGGTAAAGTCAGCCGCATTGTTGAATTTATCGAAAAA
 CCGGATCAGCCGACGCTGGACTCAGACATCATGGCCGTAGGTCGCTATGTGCTTTCTGCCGATATTTGGCC
 GGAACGAACTGACTCAGCCTGGTGCATGGGGACGTATTACGCTGACTGATGCTATTGCCGAGCTGGCGAAA
 AAACAATCCGTTGATGCAATGCTGATGACCGGCGACAGTTACGACTGCGGCAAAAAAATGGGCTATATGCAGGC
 60 GTTTGTGAAGTATGGCCTACGCAACCTGAAAGAAGGGGCGAAGTCCGTAAGGATTTGAGAAGCTGTTAAGCG
 AATAATGAAAATCTGACCGGATGTAACGGTTGATAAGAAAAATTATAACGGCAGTAAAAATTCGCAGCAAAAAGTAA

TTTGTTCGAATCTTCCTGCCGTTGTTTTATATAAACCATCAGAATAACAACGAGTTAGCAGTAGGGTTTTATTCA
AAGTTTTCCAGGATTTTCCTTGTTCAGAGCGGATTGGTAAGACAATTAGCGTTTGAATTTTTCGGGTTTAGCGC
GAGTGGGTAACGCTCGTCACATCATAGGCATGCATGCAGTGTCTGGTAGCTGTAAAGCCAGGGGCGGTAGCG
5 TGCATTAATACCTCTATTAATCAAACCTGAGAGCCGCTATTTACAGCATGCTCTGAAGTAATATGGAATAAATTA
AGCTAGCGTGAAGATACTTGTACTAGGGGCGCAGGATTTATTGGTTCTGCTGTAGTTCGTCACATTATAAATAAT
ACGCAGGATAGTGTGTTAATGTCGATAAATTAACGTACGCCGGAACCTGGAATCACTTGTGATGTTTCTGAC
10 TCTGAAACGCTATGTTTTGAACATGCGGATTTTGCATGTCTGCAATGGCGCGGATTTTTGCTCAGCATCAG
CCGGATGCAGTGTGCACCTGGCTGCTGAAAGCCATGTGGATCGTTCAATTACAGGCCCTGCGGCATTTATTGA
AACCAATATTGTTGGTACTTATGTCTTTTTGGAAGCGGCTCGCAATTACTGGTCTGCTCTTGATGGCGACAAGAA
AAATAGCTTCCGTTTTTCATCATATTTCTACTGACGAAGTCTATGGTGATTTGCCTCATCCTGACGAAGTAAATAAT
AAAGAACAATTACCCCTCTTTACTGAGACGACAGCTTACGCGCCTAGTAGTCCTTATCCGCATCAAAAGCATCC
AGCGATCATTTAGTCCGCGCGTGGAAACGTACCTATGGTTTACCGACTATTGTGACTAACTGTTTCAATAACTAC
GGTCTTATCACTTTCCGAAAAAATTGATTCACCTAGTAATTTCTAATGCTCTGGAAGGTAAGGCATTACCTATTT
15 ATGCAACAGGGGTCAAAATTCGTGACTGGTGTATGTTGAAGATGCTGCAATGGCGCTGCTTATGCTGACTG
AAGGTCAAGCGGGTGAAACCTATAACATTGGCGGACACAACGAAAAAGAAAAACATCGATGTTGTGCTGACTATTT
GTGATTTGTTGGACGAGATAGTCCCGAAAGAGAAATCTTATCGTGAGCAAATTAATTGTTGCTGATCGCCCAG
GGCATGATCGCCGTTATGCGATTGATGCTGAGAAGATTGGTCGCAATTGGGATGGAACCACAGGAAACGTTT
GAGAGTGGGATTTCGTA AACCGGTGGAATGGTATTTGGCTAATGCAAAATGGGTTGATAATGTGAAAAGTGGTGC
20 CTATCAATCGTGGATTGAACAGA ACTATGAGGGCCGCCAGTAATGAATATCCTCCTTTTTGGCAAAACAGGGCAG
GTAGGTTGGGA ACTACAGCGTGTCTGGCACCTCTGGGTAATTTGATTGCTCTTGATGTTCACTCCACTGATTAC
TGTGGTGATTTTAGTAACTTGAAGCTGAAGTGTGGCTGAAACAGTCAAAAGAAATTCGACCTGATTGTTAATGTAATGCT
GCGGCTCACACCGCAGTAGATAAGGCTGAGTCAGAACCCGAATTTGCACAATTA CTCAATGCGACTACGCTTGA
ATCAATTGCAAAAGCGGCAATGAAGTTGGGGCTTGGGTAATTCATTA CTCACTGACTACGTATTCCCTGGAAA
25 TGCGGACACGCCATGGCTGGAGATGGATGCAACCCGACCCGCTAAATGTTTACGGTGAACCAAGTTAGCTGGA
GAAAAAGCATTACAAGAGCATTGTGCGAAGCACCTAATTTCCGTACCAGCTGGGTCTATGCAGGTAAAGGAAAT
AATTTCCGCAAAACGATGTTGCGTCTGGCAAAAGAGCGTGAAGAACTAGCCGTTATTAATGATCAGTTTGGTGGC
CCAACAGGTGCTGAACTGCTGGCTGATTGTACGGCACATGCCATTCGTGTGCGACTGAATAAACCCGGATGTCCG
30 AGGCTTGTACCATTTGGTAGCCAGTGGTACCACAACCTGGTACGATTATGCTGCGTGGTTTTGAAGAGGCGC
GCAATGCAGGCATTCCTCTTGCACTCAACAAGCTCAACGCAGTACCAACA ACTGCCTATCCTACACCAGCTCGTC
GTCCACATAACTCTCGCCTTAATACAGAAAAATTTACGACAGAAATTTGCGCTTGTATTGCCTGACTGGCAGGTTG
GTGTGAAACGCATGCTCAACGAATATTTACGACTACAGCAATTTAATAGTTTTTGCATCTTGTTCGTGATGGTGG
AGCAAGATGAATTAAGGAATGATGAAATGAAAACGCGTAAAGGTATTATTTAGCGGGTGGTTCGGTACTCG
TCTTTATCCTGTGACTATGGTCGTCAGTAAACAGCTATTACCTATATATGATAAACCGATGATCTATTATCCGCTTT
35 CTACACTGATGTTAGCGGGTATTCGCGATATTCTGATTATTAGTACGCCACAGGATACTCCTCGTTTTCAACAAC
GCTGGGTGACGGTAGCCAGTGGGCGCTGAATCTTCAGTACAAAGTCAACCGAGTCCGGATGGTCTTGCAG
GCATTTATTATCGGTGAAGAGTTTTATTGGTGGTGTAGTTGCTTTGGTACTTGGTGATAATATCTTCTACGGTC
ACGACCTGCCTAAGTTAATGGATGCCGCTGTTAACAAAGAAAGTGGTGCAACGGTATTTGCCTATCACGTTAATG
ATCCTGAACGCTATGGTGTGCTTGTGATTTGATAAAAACGGTACGGCGATCAGCCTGGAAGAAAAACCGCTACAA
40 CCAAAAAGTAATTATGCGGTAACCGGGCTTTATTTTTATGATAACGACGTTGTGCAAAATGGCGAAAAATCTTAAGC
CTTCTGCCCGCGGTGAACTGGAAATTAACCGATATTAACCGTATCTATATGGAACAAGGGCGTTTATCTGTTGCCA
TGATGGGGCGTGGTTATGCGTGGTTAGACACGGGGACACATCAGAGCCTGATTGAGGCAAGCAACTTTATTGCA
ACAATTGAAGACGCTGAGGGCTGAAAGTTTCTGCCCGGAAGAAATGCTTACCCTAAAGGGTTTTGTTGATGC
TGAGCAGGTGAAAGTATTAGCTGAACTCTGAAAAAAATGCTTATGGTCAAGTATGCTGAAAAATGATTAAAGG
45 TTATTAATAAAATGAACGTAATTAACAGAAATTCCTGATGTA CTGATTTTTGAACCGAAAGTTTTTGGTGTAG
CGTGGTTTTCTTTTTGAGAGCTTTAACAGAAAGTTTTTGGGAAGCTGTAGGCCGCAAAAGTTGAATTTGTTTCA
GATAACCATTGAAAGTCTAGTAAAGGTGTTTTACGCGGGCTGCATTATCAGTTGGAACCTTATGCACAAGGAAAA
TTGGTGCCTTGCCTTGTGCGGTGAAGTTTTTACGCTAGCTGTTGATATTCGTAATCGTCATCGACTTTTTGGCAAAT
50 GGGTTGGGGTGAATTTATCTGCTGAGAATAAGCGGCAATTTGTGGATTCCTGAGGGATTTGCACATGGTTTTTTAG
TGCTGAGTGAGACGGCGGAGTTTTTGTATAAGACGACAAATTTATCATCCTCAGAGTGATAGAGGAATAAAAT
GGGATGATCCAAGCATCAATATTTTATGGCCAGTCTGATTACAAAGTGTGCTATCAGCTAAAGATAATAAGCATC
CTCCATTAACAAAGATTGAAATGTATAGTTAAGATCACGATAAATCTTGAAGGGTTGCAAAATTAATAAAATAG
TGAGCAAAAGTGAATAAAGGAACGTAATCCACAATGCTGGCTATATGATGATTACTCAGATAGCTTTATATGTTGC
ACCATTTTATACTGAGTTATCTGTTAAAAACACTGGGGTTGCACAGTTTGGTAATTATGCCTTAATACTATCAA
55 TCGTTGCATATTTACAGATTATAACGGATTATGGTTTTCTTTTATGCAAGTCGTGCGATCTCACAGAATAGAGA
GGACAAAGAATATATACAAAAATTTATCTGTCAACTATGACTATCAAGTTGGCGATATGCGCTTTCTTATTCTTAT
TGCTGATGCTATTTTTAAATCTTTTGCCTGTGCAAGCTGAATTA AAAACAAGGAATATTATGGATATCTTCTTGT
ATAGGAAATACCTTCCAACCAACTGGTTTTTCCAAGTGTAAAAAATTA AAAATCATAGCCCTTTCTAATGTTAT
ATCAAGATGCCCGCGTGTTTACTTGTATTTATCTATGTGAGGAATAGCGAGGATTTACAAAAAGCATTTTTAGTA
60 CAGTCACTTCCATTAGTAATTTCTGCGATTGGATTAATATATTTATATTGAAATATATCAATATTATTTTTCCGGAA
AAAAAATTTAAGGTAATTTTAAAAGAAGGTAAGGATTTTTTCTTGCATCACTTTATTCTGTTATTCTCAATAAT
AGTGGCATTTTTCTATTAGGGATTTTACTAATCCTGTTATTGTTGGTGTATATGCCGCGCTGAAAAGATAGTCA
AGGCCGATTGTGCTATTTACACCACTGACGCAAGCTATATATCCTTATAATTGTCGTAAGTTTTCACTATCCGT
ATTTGACGGCATTGAGGCAGCAAAAAAACTGGTATACCAATTATAATTTTAGCATTTATAGCTGCTGTTATCGTT
GCAATTACCTTACCTGTTGCAATCGACTATCTTAATTTTCAAAAAGAAACAATTTTTGTAGGTCAAATATTAAGTGC

ATGGATCTTTTTGGTGTCTTAATAATGTATTCCGGCATTGAGATATTGAGTGCATCAGGAAGAAGTAAAAATATATA
 GTAGGATGGTATTCGTATCAGCGCTTATAACATTACTTTTGATTACTCTATTATTGCAGTTTTGTAACGCCACTGG
 AGTGGCATGTGCAATATTATTGGGTGAAATGTTCTTATCAATATTGTTACTTAAGCGATATAAAAAAATAATTTAAG
 GAATAGTTATGAAGAAGTTATTATTAGTGTTCGGTACTAGGCCTGAAGCAATAAAGATGGCCTCTATCATTGAATT
 5 ATAAAAAAGATTGTAGATTGCAATATAAAAATATGTGTGACAGGCCAACATAAAGAGATGCTTGATCAAGTTATG
 CAAGTATTTGATGTTAAACCTGATTATAATTTACGGATTATGCAGCCTGGGCAACATTAAGTATCTATAGCAACAA
 ATATACTCTCACGGTTAAGTGAAGTTTTAATTATAGAAAAGCCAGATATTATACTTTGTCATGGGATACAACGAC
 TACCCTTGCTGCTACTTTAGCTGGGTATTACCACCAAATAAAGTTTTGTCATGTGGAAGCAGGATTAAGAACAGG
 10 GGATATTTACTCTCCTTGGCCTGAAGAGGGCAATCGTAAAGTTACAGGGGCATTAGCATGTATTTCATTTGCCCC
 AACAGAGAGATCAAAGATAATCTCCTGAGGGAGGGGTCAAAGTAAATAATATATTTGTAACGGGTAATACCGT
 CATCGACTCTTTATTTATTGCAAAGATATCATAGATAATGACCCTAATAAAGAACGCTTTACATAATAAATTTA
 ATTTCTTGATAAAAAGCCGACGAGTAGTACTTATAACAGGTCATCGAAGAGAAAATTCGGGAAAGTTTTGAA
 ATATAGCTTTGCAATAAAGGAATTAGCTTTTCATTTATCCTAATGTAGATTTTATTTATCCGGTGCATCTTAATCC
 15 ATGTAATGGAACAGTACATCGTATTAGTAAATATGTAATATTTACCTTATTGAGCCTGGATTATTGGCC
 TTTTGTATTTAATGAATGAGTCATATTTAATATTGACTGATTCAGGGGGATACAAGAAGAAGCCTTCGTTA
 GGTAAACCGTTTTGGTTATGCGTGATACTACTGAACGCCCTGAGGCGTTGAGGCTGGTACTGTTGTATTAGT
 GGGGACTTCTAAGATAAAAATAGTAAATAAAGTAACGGAGCTATTAACAATGCTGATATCTACAATGCTATGTCT
 CTGTTACATAATCCATATGGCGATGGAACAGCTGCTCAAAAAATTCCTAATGTGCTCGCCCAAGAGCTAATTTAAT
 20 TTAAGCTAAAAATATGTTATTAATTATTGCTGATTATCCAAACGAAATGAATATGCGCGAGGGAGCTATGCAACGA
 ATAGATGCGATAGACTCTCTCATTGAGATCGCAAGCGAGTGTATTTGAATATTTTCATTTCAAAAAGCATCTAGTTC
 GCTCAAATAGTTCCTTTAATAATGTTATAGTTGAAAATCTAAATGCAATTTACAGAAACATCAAAAACAGTAC
 ATGCAAAAATCAACAACATATATATGTTTCATTCTGTTTATAATTTATTAAGGTTATAACGCTCATTGATCTAAAAA
 ACAATTCCTTGATATACATGGTGTGTACCGGAAGAAGTTCCTGGCAGATAATAAAAAATTAAGTATATAA
 25 CATGGTGGAAAAAAGGTGTCCTTGGATGCAAAAAATTAATACACGTCAGTACAGAAATGCAAAAAACTATGA
 AGCAAAATATGGAGTAACTTGGCTGAAAGGTCAATAGTGTCTCCCGATTTTTGAATATAAAAAATATAACCCAATCG
 CAAAACAAATGGACAGAAAATAAATACGAAGTATCTATCTTGGAGGATTACAAACATGGCAAAATATTGATAAAA
 TGATTCAGTTTGTGATGACACAGTGATAAACAATGAAGCAGGTAAGTATGAATTCACCTTTTTCATCCCACAGAG
 TAACTTGGAAGGTTTATAGATAAAATTCGTTAAAATACATAATCAATGCTAATGCATCTACGCTATCACGCTG
 ATGAAGTAATCCCTTTCTAAAAGAATGTCATATTGGTTTTGTATTGCGCGATGATATAATAGTAAACAGAGTTGC
 30 GTGCCCTACAAAATTGGTTGAATATTTAGAGTGTGGTGTGCTTCCAGTTGTGCTCTCCCACTTATAGGTGATTTT
 TATTCGATGGGATATCAATACATTACTACAGAGGAAATGGCTAACAGAAGTATAAGTTTGTGGATCTTGAAAAA
 TGGCTGCACATAATTTACAAATTTGACTTCTTATCAGAAGAGAACCTACAAGGCACAGAAAGAAGTATTGCTCA
 ACTGTGCTGAATTTTTTACATATATAAAATTATGTAAGCATATCGCGGGTCAAGGTAATTGTATGCGTATCAAAATA
 AAGATAACGGTTATATATTATGTTTTCTATTAATGTTTCATTTTGAGCTACTTAGTTTTACTCAAATCTGACTACTTTC
 35 ATATTTAACCGTTGTTTTCATTAATTTAAATTTCCCTATATTTTTGCAATGTTAGTTTGTCCCTTATGTTTAAAG
 TTGGAATAAATAATGCAAGAAAAATAATTAAGATAGTTATATATATTTGTTCTTGTATGTATATGTATCATTTTA
 TGTGTTTTTGCATGAAATGACTCAATTGCGCATAGCAATTGCAGTCACTATGTGCTATGTGCTCGGTTTTATTATTAC
 TTTTATAAAAAATTGATTAACATGCACTGCCATGGATGGTGTGGCTATTTTGTTCATTACAGCGCCTTGCTTTT
 40 ATTTATGTCATTATTTATACAGTTATAGGAGGTTAATAGTAATTATAGGGTTTGAATATGTATGAGCTTTTT
 AAACGTGATGCAGATACAATTGCACTATATTTGCCAAATGAAAAAATAGTAAATTTTATATAGTATTTTCATCAT
 CATTAGCAATAGAAATGATTTGGCAATATCAACATATAAATATTTTATCAATTTTATTTTATTTTATTTTATTTT
 ATCTTAGCCGATATATAAATAAATGATAATGAGCGAAGTTTATTAAGTATGTGCAATGTTGAGGAAATTTAGC
 CTTTTGATTTTCTTCTGGCTAGTGGAGTCCCGGTCATTGCTTATCGAACTGCAGAGTTGCTGCGAATATTTTAT
 45 CCGATGGCTTTAGTATTAATCCTTTGCGATATAAAAAATAAATATGCGTTATTTTATTGCAAGTATTATAGTTATC
 CTTTCAGGCTTAATGTTGTTTATAACACTAAGGGCTGTATCAATAGTTGGTCAAGGATTATAAAATGAATGTTGCT
 ATTTTGTGCTACGTATAATGGCGAAAAATTTAGAGGAACAACCTGGATTCAATTGCTGCTTCAAAGTTATCAGG
 ATTTTGTAGTGTATATCCGTGATGACGGATCATCTGATAGAAGTGAATATAATAAACAATACGTAATGAAAGA
 TAACAGATTTATTAACGTGGGTAATTCAGAAAATCTTGGTTGTGCTGCTTCTGTTTATTAATTTAAGAAATGCTT
 50 CAGCCGATATTTATATGTTTTGTGACCAAGATGATTATTGGCTTCCGAATAAATTACAGGGCTGCTGTGGATTATTT
 TTCGGCTATTGATCCTTTACAACCTACCTTGTATCATTGCGATCTAAGCGTTGTTGATGAAAAACTTAATATTATAC
 AAAATTCATTTTTGCAGCATCAGAAAATGTCAGCGTATGATTCAATGAGAAAAATAATCTTTTCATACAAAATTTT
 GTTGTGGTTGTTTCATGTGCTGTTAATGCTTCACTTGCAGGAAATTTGTTCTTTGCGAATTGGAGAGCAGCATGTAA
 AAATGATAGCTATGCATGACTGGTGGTTAGCCGTGACTGCAAACTTTTTGGTGAATCCATTTTGATAAATACTCA
 55 AACGATTCTTTATCGACAACATCAGGGCAATGTATTAGGTGCAAAATCATCAGGTATGATGCGTTTTTATTCGATTA
 GGATTAATGGGCAAGGGATTTGCGGAGTAGTACTTTTAGAAAAAAGTTTGTGCGCAAAAATAAGCTTCTTTTAG
 ATGTCTATGATAAAGATTTAAATCTTGAGCAAAAAAATCTAAGGCTTGAATTGAGGGCCTTAAAGAACTTAAT
 TTCAATTGCTGACCTTTTTAAATGTTTCTATCATGGTAGCTATATGCAAGGTTTTAAACGTAATCTTGCCTTAATAT
 ATTCAGTTCTTTACAAAAAAGAAGATAGTGTATCCTTATGAAAAAATTGCTATTATCGGACTGTTGGCATA
 60 CCAGCATCATATGGCGGATTTGAAACATTAGTTGAAAATTTAAACAAGATACAATTCCTCGGGAGTTGAATATAATG
 TTTTTGTTTCATCGTTTCACTACAAATCCCACAAAAAACAATAATGGGGCCGTTTAAATTTATATTCCGCTTAAA
 GCCAATGGATGGCAGAGCATTGCGTATGACATAATTTGTTAGCATATTCTATTTTTTTGAAGCCTGATGTGATTC
 TGATTTTAGGGGTTTCTGGTTGTTCAATTTTGCCTTTCTTCAAACCTTAACACGCGCTAAGTTTATTTACTAATATT
 GATGGCCTGGAATGGCGAAGAGATAAATGGAATTCAAAAGTGAACGTTTCTTAAAAATTTTACAAAAAATCGCA

GTTCAATATTCGGATGTCGTTATTACGGATAATGAGGCAATTTCTGAGTACGTTTTTAACGAGTATAATAAAGATA
 GCCGAGTTATTGCCTATGGAGGGGATCATGCATGGTTAAATACTGAGGATGTATTTACAACAAGAAATTATAAAA
 GCGATTACTACCTTTCTGTATGTCGTATCGAACCCGAAAACAATGTAGAATTAATTTAAAAACATTTTCAAAGCTA
 5 AAATATAAAATAAAATTTATTGGAAATTGGAATGGCAGCGAGTTTGGAAAGAAACTTAGGCTGCATTATTCTAACT
 ATCCAAATATTGAAATGATTGATCCGATTTATGATCTTCAACAATTTTCACTTACGAAATAATTGCATAGGATAT
 ATACATGGTCATTTCGGCTGGAGGAACAAACCCTTCTTAGTCGAGGCAATGCATTTTAGTAAACCTATATTTGCAT
 ATGATTGTAAGTTTAAATAGGTACACTACTGAAAATGAAGCATGTTATTTTCTAATGAATCTGACCTCGCAGAGAA
 AATCATAATGCATTGTGAGCTATCATTAGGTGTCTCTGGCAGGAAAATGAAAGAAATTGCTAACCAGAAATACACT
 TGGAGACGAATAGCAGAAATGTATGAGGATTGCTATTAACCTCTGTTAAACTTCAAATCTTTTACAATATATGGCAT
 10 GACTATAAGCGCATTAAATTGTTTTTCAAGCCGCTCTCGCGGTGACCACCCCTGACAGGGGATCCGTGTAGGCT
 GGAGCTGCTTCAAGTTCCTATACTTTCTAGAGAATAGGAACTTCGGAATAGGAACTAAGGAGGATATTCATATG
 GATAAAGCCGTAAGCATATAAGCATGGATAAGCTATTTATACTTTAATAAGTACTTTGTATACTTATTTGCCAACAT
 TCCAGGCCGCGAGCATTACGCGCGGTGATCACACCTGCAGGAGTATGTAATGTCCAAGCAACAGATCGGCGT
 AGTCGGTATGGCAGTGTGGGACGCAACCTTGCCTCAACATCGAAAGCCGTGTTATACCGTCTCTATTTTCA
 15 ACCGTTCCCGTGAGAAGACGGAAGAAGTGATTGCCGAAAATCCAGGCAAGAAACTGGTTCCTTACTATACGGTG
 AAAGAGTTTGTGAATCTCTGAAACGCCTCGTCGCATCCTGTTAATGGTGAAGCAGGTGCAGGCACGGATGC
 TGCTATTGATTCCCTCAAACCATATCTCGATAAAGGAGACATCATCATTGATGGTGGTAACACCTTCTTCCAGGAC
 ACTATTCGTGTAATCGTGAGCTTTCAGCAGAGGGCTTAACTTCATCGGTACCGGTGTTTCTGCGCGTGAAGA
 GGGGGCGCTGAAAGTCCCTTCTATTATGCCTGGTGGCCAGAAAGAAGCCTATGAATTGGTAGCACCGATCCTGA
 20 CCAAAATCGCCGCTAGCTGAAGACGGTGAACCATGCGTTACCTATATTGGTGCCGATGGCCGAGGTCACTAT
 GTAGAATGGTTTCAACACGGTATTGAATACGCGGATATGCAGCTGATTGCTGAAGCCTATTCTCTGTTAAAGGT
 GGCTGAACTTACCAACGAAGAAGTGGCGCAGACCTTTACCGAGTGAATAACCGTGAACCTGAGCAGTTACCT
 GATCGACATCACCAAGATATCTTACCAAAAAAGATGAAGACGGTAACTACCTGGTTGATGTGATCCTGGATGA
 AGCGGTAACAAAGGTACCGGTAAATGGACCAGCCAGAGCGCGCTGGATCTCGGCGAACCGCTGTCGCTGATT
 25 ACCGAGTCTGTGTTTGCACGTTATATCTTCTCTGAAAGATCAGCGTGTGCGCATCTAAAGTCTCTCTGCT
 CCGCAAGCACAGCCAGCAGGCGACAAGGCTGAGTTCATCGAAAAAGTTGCTGCTGCGCTGTATCTGGGCAAAA
 TCGTTTTCTTACGCCAGGGCTTCTCTCAGCTGCGTGTGCTGCTGAAGAGTACAACGGGATCGAACTACGGC
 GAAATCGCGAAGATTTTCCGTGCTGGCTGCATCCGTGCGCAGTTCCCTGCAGAAAATCACCGATGCTTATGC
 CGAAAATCCACAGATCGCTAACCTGTTGCTGCTCCGTACTTCAAGCAAATTGCCGATGACTACCAGCAGGCGC
 30 TGCGTGATGTCGTTGCTTATGCAGTACAGAACGGTATTCCGGTCCGACCTTCTCCGCAGCGGTTGCCTATTAC
 GACAGCTACCGTGTGCTGTTCTGCCTGCGAACCTGATCCAGGCACAGCGTGAATTTTGGTGCGCATACTTA
 TAAGCGTATCGATAAAGAAGGTGTGTTCCATACCGAATGGCTGGATTA

SEQ ID NO: 12 (Ejemplo de secuencia de nucleótidos de locus *O2 rfb* – *O2*–producción EPA cepa stGVXN4906)

ATGACGAATTTAAAAGCAGTTATTCCTGTAGCGGGTCTCGGGATGCATATGTTGCTGCCACTAAGGCGATACCC
 35 AAAGAGATGCTACCAATCGTCGACAAGCCAAATGATTACAGTACATTGTTGACGAGATTGTGGCTGCAGGGATCAAA
 GAAATCCTCCTGGTAACTCACGCGTCCAAGAACGCGGTGAAAACCACTTCGACACCTTATGAGTTAGAATCA
 CTCCTTGAGCAGCGCGTGAAGCGTCAACTGCTGGCGGAAGTACAGTCCATCTGTCCGCCGGGCGTGACCATTA
 TGAACGTGCGTCAGGGCGAACCTTTAGGTTTAGGCCACTCCATTTTGTGCGCGACCTGCCATTGGTGACAAC
 40 CCATTTGTCTGGTACTGCCAGACGTTGTGATCGACGATGCCAGCGCCGACCCGCTACGTTTACCAACCTTGCTGC
 CATGATTGCACGTTTCAACGAAACGGGCCGACGAGGTGCTGGCAAAACGTATGCCGGGTGACCTCTCTGAAT
 ACTCCGTATCCAGACTAAAGAGCCGCTGGACCGTGAGGGTAAAGTCAGCCGCATTGTTGAATTTATCGAAAA
 CCGGATCAGCCGACGACGCTGGACTCAGACATCATGGCCGTAGGTCGCTATGTGCTTTCTGCCGATATTTGGCC
 GGAACGAAACGTAACGCTCAGCCTGGTGCATGGGGACGTATTACGCTGACTGATGCTATTGCCGAGCTGGCGAAA
 45 AAACAATCCGTTGATGCAATGCTGATGACCGGCGACAGTTACGACTGCGGCAAAAAAATGGGCTATATGCAGGC
 GTTTGTGAAGTATGGCTACGCAACCTGAAAGAAGGGGCGAAGTCCGTAAGGTATTGAGAAGCTGTTAAGCG
 AATAATGAAAATCTGACCGGATGTAACCGTTGATAAGAAAATTAACCGCAGTGAATAATTCGAGCAAAAAGTAA
 TTTGTTGCGAATCTTCTGCGGTTGTTTTATATAAACCATCAGAATAACAACGAGTTAGCAGTAGGGTTTTATTCA
 AAGTTTTCCAGGATTTTCTTGTTCAGAGCGGATTGGTAAGACAATTAGCGTTTGAATTTTTCGGGTTTAGCGC
 50 GAGTGGGTAACGCTCGTCACATCATAGGCATGCATGCAGTGTCTGGTAGCTGTAAAGCCAGGGGCGGTAGCG
 TGCATTAATACCTCTATTAATCAAACCTGAGAGCCGCTATTTTACAGCATGCTCTGAAGTAATATGGAATAAATTA
 AGTGAATAACTTGTACTGGTGGCGCAGGATTTATTTGGTTACGCTGTAGTTCGTCACATTATAAATAATACGCAG
 GATAGTGTGTTAATGTCGATAAAATTAACGTACGCGGAAACCGGAACTCACTTGTCTGATGTTTCTGATTCTGAA
 55 CGCTATGTTTTTGAACATCGGATATTTGCGATGACCTGCAATGGCAGCGGATTTTGTGCTCAGCATAGCCGGAT
 GCAGTGATGCACCTGGCTGCTGAAAGCCATGTTGACCGTTCAATTACAGGCCCTGCGGCATTTATTGAAACCAA
 TATTGTTGGTACTTATGTCCTTTTGAAGCCGCTCGCAATTACTGGTCTGCTCTTGATAGCGACAAGAAAAATAG
 CTTCCGTTTTTATCATATTTTCTACTGACGAAGTCTATGGTATTGCTCATCCAGATGAAGTAAATAATACAGAA
 GAATTACCCTTATTTACTGAGACGACAGCTTACGCGCAAGCAGCCCTTATTCGCGATCCAAAGCATCCAGCGAT
 60 CATTTAGTCCGCGCATGGAACGTACGTATGGTTTACCAGCATTGTGACTAATTGCTCGAACAACTATGGTCCG
 TATCACTTCCCGGAAAAGCTTATTCCATTGGTTATTTAATGCACTGGAAGGTAAGGCTTACCTATTTATGGCA
 AAGGGGATCAAATTCGCGACTGGTTGTATGTAGAGGATCATGCTCGTGTGCTGTTATATACCGTACCTAACCCTA
 AAAGCGGGTGAACCTTATAACATTGGCGGACACAAGAAAAGAAAAACATCGATGTTGTGCTGACTATTTGTGAT
 TTGTTGGATGAGATTGTACCGAAAGAGAAATCTTATCGTGAGCAAATTAATGTTGCTGATCGCCCAGGGCAT

5 GATCGCCGTTATGCAATTGATGCCGATAAAAATTAGCCGCGAATTGGGCTGGAACCACAGGAAACGTTTGAGAG
 CGGGATTTCGAAAACGGTGGAATGGTATCTGGCTAATACAAATTGGGTTGAGAATGTGAAAAGCGGTGCTTATC
 AGTCATGGATCGAACAAAACACTATGAGGGCCGTGAGTAATGAATATCCTGCTTTTCGGCAAAACAGGGCAGGTGG
 GTTGGGAACTGCAGCGTGTCTGGCGCCGCTGGGTAATCTGATCGCTCTTGATGTTCACTCCACTAATTATTGT
 10 GGAGATTTTCAGCAACCCCGAAGGTGTGGCAGAAACCGTCAAAAAAATTCGTCCTGACGTTATTGTTAATGCTGCT
 GCTCACACTGCAGTAGATAAAGCAGAATCAGAACCAGGATTTTCGCACAATTACTTAACGCGACAAGCGTGAAGC
 GATTGCAAAAAGCTGCTAATGAAGTCGGGGCCTGGGTTATACACTACTCTACTGATTATGTTTTCCAGGCAGTGG
 TGACGCGCCATGGCTGGAACGGATGCAACAGCACCGCTAAATGTTTACGGTGAAACAAAATTAGCTGGGGAAA
 AGGCATTACAAGAACATTGCGCAAAGCATCTTATTTCCGTACCAGCTGGGTATACGCTGGTAAAGGAAATAACT
 15 TTGCTAAAACGATGTTGCGTTTGGCAAAAAGAACGCGAAGAAGTGGCTGTGATAAACGATCAGTTTGGCGCACCA
 ACAGGTGCTGAATTGCTGGCTGATTGCACCGCTCATGCCATTCGCGTGGCATTAAAAAACAGAAAGTCGCTGG
 CTTGTACCATCTGGTAGCAAGTGGCACAACAACCTGGCAGCATTATGCTGCGCTGGTTTTTGAAGAGGCGCGCA
 AAGCAGGGATTAATCTTGCACTTAACAAAACCTAACGCGGTGCCAACACGGCCTATCCACACCAGCCCGTGA
 CCCCAACTCTGCGCTAATACAGAAAAGTTTACGAACTTTCGCGCTTGTCTTGCTGCTGACCTGCAAGTGG
 20 CGTGAACCGTATGCTCAACGAATTTTACGACTACGGCAATTTAACAAAATTTTGCATCTCGCTCATGATGCCAG
 AGCGGGATGAATTAAGGAATGGTGAATGAAAACGCGTAAAGGATATTATTCTGGCTGGTGGTTCCGGCACTC
 GTCTTTATCCTGTGACGATGGCAGTGAGTAAACAATTGCTGCCGATTTATGATAAGCCGATGATTTATTATCCGCT
 TTCAACGCTTATGTTAGCGGGTATTCGCGATATTCTTATTATTAGTACGCCACAGGATACACCGCGTTTCCAACAA
 TTATTGGGGGACGGGAGCCAGTGGGGTCTTAATCTACAGTATAAAGTACAACCGAGTCCGGATGGCCTGGCGC
 25 AAGCGTTTATTATTGGCGAAGACTTTATTGGTGGTGATGATTGTGCACTCGTACTTGGCGATAATATCTTCTATGG
 ACACGATTTGCCGAACTGGAAGTGTGAAGTCTGTTAACAAAGAAAGCGGTGCAACGGTATTTGCTTATCAGTTAA
 TGATCCTGAACGCTATGGTGTGCTGGAGTTTGAATAACGGTACGGCAATTAGCCTGGAAGAAAAACCGCTGG
 AGCCAAAAAGCAACTATGCGGTTACTGGGCTTATTTCTATGACAATGACGTTGTGGAATGGCTAAAAACCTTA
 AGCCTTCTGCCCGTGGCGAACTGGAATTAACCGATTAACCGTATTTATATGGAACAAGGACGTTTGTCTGTAG
 30 CCATGATGGGGCGTGGCTATGCATGGTTGGATACAGGGACGCATCAAAGCCTTATTGAAGCAAGTAACCTCATT
 GCAACAATTGAAGAGCGTCAGGGATTAAGGTATCTTGCCCGAAGAGATTGCTTACCGTAAAGGGTTTATTGAT
 GCCGAGCAGGTGAAAGTATTAGCCGAACCGCTTATCAAGAATCAATATGGTCAATATTTGCTGAAAATGATCAGC
 GAATAGTATATGGAACTCAATGATGGATATAAATTAATCTCTTTGCAAAAACATGGGGATGAGCGCGGTGCAT
 TAATTGCTCTTGAAGAGCAACGAAATATACCTTTTCAAGTCAAAAAGAAATATATTACATACTTGAGACTCTTAATGG
 35 AGTAAGACGCGGATTTTCATGCGCACAAGGTTACTCGTCAGTTAGCTATTGTAGTCAAGGGAGCTTGTAATTTCA
 TCTGGATAATGGTAAAGAAACAAGCAGGTGGAACCTAATGATCCAACAATTGCGTTGCTGATAGAACCCTATAT
 ATGGCATGAAATGTATGATTTTATGATGATTGTGTGCTGCTTGTAAATTGCGGATGATTTCTATAAAGAGTCTGAT
 TATATCCGCAATTATGATGATTTTATTAGAAGAGTAAATCAATTGAGAAATCATAAGCTAAGTGACGTCCAGACA
 ACATCAATTGGTATGGAACAACCTATCTGGCAGTTTGTGTGTAACATAAAAAGGTGCTGTAATTGGTAATAATTGCA
 40 ACATCTGTGCAAACTAATTAATGAAAATAACGTTGTAATTGGTAACAATGTCACAGTCAAAAAGCGGTGTGTATAT
 TTGGGATGGCGTTAAAAATAGAGGATAATGTTTTATTGGTCTTTGTAGCATTACAAAATGATAAGTATCCCTCGC
 TCTAAAGTCTATCCTGATGAATTTTTGCAAAACAATAACGCAAAGGAGCATCAATAGGTGCTAACGCAACCATCC
 TGCCAGGAATTGAAATGGTGAAAAAGCAATCGTTGGTGGCGGGGAGTGTGTAACCAAAAATGTACCGCCATGC
 GCAATAGTAGTAGTAATCCAGCTCGATTTATTAATGGGTAGAGGATAATGAATAAAATTGATTTTTTATGATCTTT
 45 TTGCAATTAACCAGCGACAGCACAAAGAATTAGTCTCTGCGTTTATGATAGGGTGTAGATTCTGGTTGGTATATCA
 TGGGCGAAGAAGTTCGAGCAGTTTCGAGAAAGAGTTCGAGAAACTGTGGAGTTAAGTATTGCATTTGGTGTAGCA
 AATGGCCTTGATGCGTTGATACTAGTATTGAGGGCATGGAAGAAGTTCGCTATCTTGAAGACGGTGACGAGGT
 ATTAGTACCGGCAAAATACATATATTGCTTCTATTCTGCTATAACAGAGAAACAACCTTTGCTGTTCTTTGTTGAAC
 CAGATATAGAAACTTATAATATTAATCCTGCTTTAATTGAAAATTACATTACGGAAAAAACTAAAGCAATATTACCG
 50 GTTCACTTATATGGTCTATTGTGCAATATGCCAGAAATTAGTGAATCGCCAGAAAATATAATCTGTTGATTCTTG
 AAGATTGTGCACAAGCACATGGTGAATACGTGATGGTGCAAAAGCTGGAGCTTGGGGGGATGCTGCAGGATTT
 AGTTTTTATCCAGGAAAAAACCTTGGAGCTTTGGGGGATGCGGGAGCTGTTACTACAAATAATGCAGAATTATCC
 TCAACTATAAAAGCTTTGCGAAATATGGGTACATAAGAAATATGAAAATATTTATCAGGGATTGAATAGTCGAT
 TTGCTGAAAAATATATTGTAATAAATAAACCCTGCGGATTACGTTACCAGTGTACGAAGGCCAAGGTGCGCATG
 55 TTTGGCATTATTTGTAGTAAGAATCGCTAATCGTGAAAAATTCAGTACATACTTATTAGAGAAGGGTATCAAAAC
 CTTAATTCATATCCATTACCACCCCATAGCAGCAAGCATATCAAAATATGTCTAGCCTTAGCCTTCCAATTACT
 GAGCAAATTCATGATGAAGTCAATTTCTTTACCTATAAGTCCGGTAAATGAGTGAAGATGATGTCAATTATGTAATCA
 AAATGGTCAATGATTACAAGTAAATGAAAAATTTCTTCAGGTAACCTATATTTACCGCTATCTATACATTCATTA
 60 TGATTGCGGGTTTTATCATCGGTAAGGTAGTAGCAATTTATACAGGGCCATCAGGGGTAGCAATGCTTGGCCAA
 GTGCAAAAGTTAATCACAATAGTTGCAGGTACTACCTCTGCACCTGAAGCACAGGCCTTGTTCGATACTACGCG
 GAAAATTGGCAAGGACAAGGACATGCGCCCATGTTGGCGCGCATGCTTAAGGGTACTCTGTTTTTTATT
 CTTGCTTATTATCCCGTTGTTATTATATTGTCGAAAAATATTAGTGAGTTACTTTTTAGCGATGACAATACACAT
 GGTAAATCATTTTCGCATGTTGTATATTGCCATTCTCCATTATAAATACATTGATCGCTTCAGTTTTAAATGGTCAA
 CAATTTTATAAGCAATATATATTGGTTGGGATGTTTTCTGTATTCAATTTCTACTATGTTTATGATTTTGTGATTGTA
 GCTTATAATCTTAAAGGTGCATTGATTGCCACAGCTATAAATAGTGCTATTGCTGGTCTTGATTGGTTTTATTTG
 TCTCAATAAATCTTGGTTTAGATTTAAATATTGGTGGGGTAAAACGGATAAAGACAAAATTATAAAAAATTATTCATT
 AACTGTTGGGAGGATGCAGGGCAATGGCAGGCCGATGGAAGATATCTGAGGTTTATCTTGGTGTGTGACAA

TTGCTTTGTCAACATATTTCTTACCAAGATTGACAATTATAAAAAACAAGTTTCCCTTATAAAAAAAGAAGTAAATAGT
ACTATATTATACATAATATCTATTACTTCATTCATGGCGTTGAGTATCTATTTATCCGCGATTTGGTAATAACAGTT
TTATTTACTGAACAGTTTCGCTCAGCTCGTGAATTATTTTATTACAACCTTATAGGGGATGTAATAAAAAATTGCTGG
GTTTCTTTATGCATACCCTCTTCAAAGTCAGGGGCATACTAAACTATTCATCAGTTTCAGAAAGTATTTTTTCTATGC
5 TCTTTATCATTACCACCTATATTTTTGTTGTA AATTATGGAGTACATGGTGCTAACATAAGTTATGTCATTACATATA
GTTTATATTTGTGTTTGCATTTGTGTTACTAATTTTATTAATGTTAGAAGAAATAATAAAAACAGAGGTTGAATT
TTGAAAATAATTATACCTGTCTTAGGATTTGGCAGGGCTGGTGGTGAAGAGTTCTTTCTAAGCTGGCAACTGAA
TTGATGAATTATGGACATGATGTAAGTTTTGTTGTTCCAGATAATAGA ACTAATCCATATTATGCTACCACAGCAAA
AATTGTCACGAGTAAATCTAGTCAAAACCGTGTA AAAAATATTGAGAATCATTAAAAATTACTATAATCTGTGGCGTA
10 AATGCATAGAGTTAAATCCTGATGCTGTAGTTGCTAGTTTTCATTTGACTGCCTATCTTGTGCGCATTATTACCAATC
ACCCGTCGTAAAGAAATATTTATATTACAGGCGTATGAAGTTAATTTTTTTGATAATAAATATGAAATTAATAGC
GGGTTTAAACATATTTTACCGCTTAAAAAAATACTAAATAGTCCTAATTTGCTTCCCTCATAAACATGATGATTTTAT
AGGAGTAGTTCCTGCAGGAGTAGATTTAAACGTTTTCTATCCGAAACCATCAAATAGGTTATTAATGGTCACACA
15 TCAATAGGTTATTTGAGAAAAGAGACAAAAGGAACGAAATTTTCAAGTATTTCAGTATTGTTGTTCACTGAA
AATAAAGCTGGAATTATAATCAATATTGCGATCTACTTTGAAGAAGTTGATAAGCAGCTTTAATCGCTGCCGGT
TTCAGGTTAATTTTTTCCGATTACTTCTGATTTAGAATTGGCATCCTTTTATCGAAGCAATGACATCATGATTGCT
GTTGGGTTAATTGAAGATGGCGCTTTCCATTATCCTTGTGCTGAATCAATGGCTTGTGGTTGCTTGTATTTC
ATTATGCGCCACTTACTGAACTAACAGTGTACTTAAATTAGTCAAGTTTGATGCTTGCAAACTTGGTGAAGCAAT
20 TAATCTTTGTCTCAATCTTGACCTAGAAGAAAAAAGCAAAGAAATCCAATCTAATATTTCTGTGTTGAATAAATATG
ACTGGAAAAATTGTTGGTGAACCTTCAATAGTTTATTGTTAGATGCAATAAATAGTATACGTTGAGGGGAAAAAT
ATGAATATTGTTAAAACGATTTCCAGATCTGATCGTTCTTGAACCAAAAAGTGTTTAGTGTGAACGCGGCTTTT
TTATGGAGAGTTATAATCAGATTGAATTTGAGAAGGCAATAGGAAGGCACGTAAATTTTGTTCAGGATAATCATT
AAAATCTAGTAAAGGCGTACTACGTGGGTTGCATTATCAATTAGCACCGTATGCACAGGCTAAATTAGTTGATG
25 TGTTGTAGGTCAGGTATTTGATGTTGCTGTTGATCTTAGAAAAAATTCACCAACGTTCAAAAAATGTTTGAATA
ACCCTTTCCGCGAGAAAATAAACGACAATTATGGATACCCGAAGGATTTGCTCATGGTTTCTTGGTGACCAGTGAT
GAAGCTGAGTTCATTTATAAGACA ACTA ACTATGCTCCTGGTCATCAGCAAGCAATTTTACAATGATCCTA
TTTTAAACATCGATTGGCCTTTCTGCAGTAGTCTGTCTGATTATCACAAAAAGATCAAGAAGCAAAATTTATTTTCA
GAATTATTGGACAGTGAACCTTCTAATAAAGTGTGCCACCTTATCCGCTGAAGGATAGGTGGTTGCTTATATTT
30 TTTTGTAGTATGTTTGTATAATGACAGAAAAATAGTCCGAAATATAAACACGATAAAAAGCTTAATAAGTTTTTACT
TATTTTTTATATTTACACTTATTGTAGGCTTTATTATCGCAAAATACCAGTTTTTGGGGCGAAGTAGAGACTATGAT
AATTATACAGATCTTTTCTGGTAAAGAAGGGGAGGGGTTCTTGAATTATTTTATCGCGGATTGATGTTAATAA
CGACCAGCTATGAACTATCATTTTTATAATTTAACATGTTCTTTTTTATAAAGGCAAGGTTTCTCGCTAACTATT
CGCGTAATTTTTTCAAGCTTGACCTTATTTCTTTATTTATTATGCAAGCGTTGCACCTTTGGGTTTTAGATTACTCAA
35 TTCAGAAATGGTCTATGATTTCCATTTAATGTTTTCCGTATACTATTTATTTATAAATAAACCGACTTATTTTTATT
TCTCGTATTATGTCAATTGCAACTCATTGGTCTGTTTGCCTTTTTGCTTTTTATATCCTTTTGTCTATTCAACA
AAAATAAGACGCTTGGTTATTTTTGTTTTCAGTATTTCTTTTGTATTTGATTTGCGATCTCAGGAGAAGGAAAAGAGATCA
TATCTTTTATAAGAAATTTTGGAGTGGGACAAAAAATAGGAAATGAAGCTGGTGTAAATTTAATAAATTCATTATCC
CTTACCCTATTTTCTGGTTTATTATTAGTTACATATCAAGCATTGAAAATGAAAGGAGAAATTTAAGGCTTTTCTT
40 TTGTTATGGTGTGATGCAATACGTGACTTTTAGCCTTTTCTCTACCTGTTATGGCTTTCCGTATTTTGGAAATGT
ATTTTTTCCTTATGCTAACCATTTGGGGTGTTTATTAAGCAAAAAAGAATTATTTATTTTTTGCAAAGTGTTAA
TTTTATTGTATCTAACATACTATTATCATATGGTCTTTGGAGTGATTAATGTGTAAGGCTAAGGTGTTGGCTATAAT
TGTTACTTACAACCCGAAATTAATCGATTGACGGAATGTATTAACCTTTAGCCCCACAAGTTGAGAGAATAAT
CTTGTAGATAATGGCTCAAATAATAGTGATTTGATAAAAAAATCAGTATTAATAACCTTGAATTTTACTTTTCCG
45 GAAAACAAAGGCATTGCAATTTGCTCAGAACCATGGTGTAAAGAAGGGCCTGGAAGCAAAAGAGTTTACTATTTA
TTTTTCTCAGATCAGGATACTTGTCTTCTAGCGATGTTATTGAAAACTTAAAGAGTACATTTACGAAAAATAATAA
AAAAGGTAAAAATGTTGCTTGTGCTTCTCCTTTTTTAAAGACCATCGTTCAAATTATATGCATCCGTCAGTCAGC
CTAAATATTTTTACGAGTACAAAAGTTATATGTAGTGAAGTAGACGATGATCTTTATCCCTCGCATGTTATTGCTTC
TGGGATGTTAATGTCTCGTGAAGCATGGCGCGTCTCGGACCATTTTGTGAAAACTCTTTATAGACTGGGTTGA
50 TACAGAATGGTGTGGCGTGCATTAGCTAATAATATGATTATTGTTGAGACACCATCAGTCATCATTTCTCATGAA
CTTGGGTATGGGCAGAAAAATTTTGTGCTGATCTGTTACAAATACATAATTCTTTTCAGAAATTTTTATAAAAATACG
CAATGCAATATACTTAATGCTGCATTCAAATTATAGCTTCAAGTATCGTTATCATGCTTTTTTTTTCATGCGACAAAGA
ATGTTGATTTGAAATTTTATTTTCGAAAGAAAAATTAATTCACTGAAGGTTTGTTTTAAAGCTGTACGTGATGGT
ATGTTCAATAATTTTTAATACGAAAATAGTTAGGCTCAAGGTGTTTAAATGGAAGAAAATAATATGAAGACGGTCG
CTGTAGTTGGCACAGTGGGTGTTCTGCTTGTATGGTGGGTTTCAATCACTTGTTCAGAAATCTAATTGATTATCA
55 ATCTGATGGTATAACAATATCAGATATTTTGTCTTCAAAAAAATGATAAAAAAATTTAAAAATTATAAAAAATGCAGA
ATTAATCTAATTTGCCGATAAATGCCAATGGCGTCTAGCATAATTTATGATATTATGTGTTAAATTTATTTGTTATT
CAAAAGGCCAGATGTTGTTTTAATATTGGGGTGTCTGTTGTTTATTCTACCAATTTATAAACTATTTTCAAAT
CAAAGATTATTGTCAATATTGATGGGCTTGAATGGCGTAGAAAATAAATGGGGAACGTTTGTAAAGAAATTTCTTAA
AATATCTGAGGCGATATCTATTAGAATAGCTGATATTATCATTTTCAGATAATCAAGCAATAGCTGATTATGTGGAA
60 AATAAGTACAAGAAAAAAGTGTAGTTATAGCTTATGGCGGAGATCATGCCACTAATCTTAGTACACCGATAGAC
AATGATCAAAAAAAGAAGGTTATTTTGGGGCTTTGTAGGATAGAGCCTGAGAATAATATAGAAATGATTCTGA
ATGCCTTCATTAATACAGATAAAAAAATTAATTTATGGGTAATTTGGGATAACAGCGAGTATGGACGCCAGCTAAA
AAAATATTTTCAAATATCCAATATCACCTACTAGAACCTAACTATAATATTGAAGAGCTTTATAAACTAAGAA
AAAATTGCTTGCATACATTCATGGACACTCGGCTGGTGAACAAACCCCTTCTTAGTTGAAGCGATGCATTTTAA

TATTCCTATTTTTGCTTTTCGATTGTGACTTTAATCGTTACACAACATAACAATTTAGCTCATTACTTTAATGATTCTGA
 ACAACTTAGCTTATTAGCAGAAAAGTTTGTCTTTTGGAAATCTTAAATGTCGAGTATTAGATTTAAAAAATTATGCTG
 AAGATATGTATAACTGGAGGCATATAGCTGCTATGTATGAATCTATTTATTAACGCATTAACAATAATATAATTGA
 CTTTATATAGCAGGGAAAAGATCACGTAACGCTGCGGGCGCGCCGATCCCATATGAATATCCTCCTTAGTTCCCTAT
 5 TCCGAAGTTCCCTATTCTTTCTAGAGAATAGGAACCTCGGAATAGGAACTAAGGAGGATATTCATATGGATAAAGC
 CGTAAGCATATAAGCATGGATAAGCTATTTATACTTTAATAAGTACTTTGTATACTTATTTGCGAACATCCAGGCC
 GCGAGCATTACGCGCGGTGATCACACCTGCACAGGAGTATGTAATGTCCAAGCAACAGATCGGCGTAGTCGGTAT
 GGCAGTGATGGGACGCAACCTTGCCTCAACATCGAAAAGCCGTGGTTATACCGTCTCTATTTTTCAACCGTTCCC
 GTGAGAAGACGGAAGAAGTGATTGCCGAAAATCCAGGCAAGAACTGGTTCTTACTATACGGTGAAAGAGTTT
 10 GTCGAATCTCTGAAAACGCCTCGTCGCATCCTGTTAATGGTGAAGCAGGTGCAGGCACGGATGCTGCTATTGA
 TTCCCTCAAACCATATCTCGATAAAGGAGACATCATCATTGATGGTGGTAACACCTTCTTCCAGGACACTATTCCG
 CGTAATCGTGAGCTTTCAGCAGAGGGCTTTAACTTCATCGGTACCGGTGTTTCTGGCGGTGAAGAGGGGGCGCT
 GAAAGGTCTTCTATTATGCCTGGTGGCCAGAAAAGCCTATGAATTGGTAGCACCGATCTGACCAAAATCG
 CCGCCGTACTGAAGACGTGAACCATCGTTACCTATATTGGTCCGATGGCGCAGGTCACTATGTGAAGATG
 15 GTTCACAACGGTATTGAATACGGCGATATGCAGCTGATTGCTGAAGCCTATTCTCTGCTTAAAGGTGGCCTGAAC
 CTCACCAACGAAGAAGTGGCGCAGACCTTACCGAGTGAATAACGGTGAAGTGAAGCAGTTACCTGATCGACAT
 CACCAAAGATATCTTACCAAAAAAGATGAAGACGGTAACTACCTGGTTGATGTGATCCTGGATGAAGCGGCTAA
 CAAAGGTACCGTAAATGGACCAGCCAGAGCGCGCTGGATCTCGGCAACCGCTGTCGCTGATTACCGAGTCT
 GTGTTTGCACGTTATATCTTCTCTGAAAAGATCAGCGTGTGGCCGATCTAAAGTTCTCTCTGGTCCGCAAGCA
 20 CAGCCAGCAGGCGCAAGGCTGAGTTCATCGAAAAAGTTCGTCGCTGCGCTGATCTGGGCAAAATCGTTTCTTA
 CGCCACGGCTTCTCAGCTGCTGCTGCTGCTGAAGATACAACCTGGGATCTGAACCTCGGCAAAATCGGCG
 AAGATTTTCCGTGCTGGCTGCATCCTCGTGCAGTTCCTGCAGAAAATCACCGATGCTTATGCCGAAAATCCA
 CAGATCGTAACTGTTGCTGGCTCCGTACTTCAAGCAAATTGCCGATGACTACCAGCAGGCGCTGCGTGATGT
 CGTTGCTTATGCAGTACAGAACGGTATTCCGGTTCGACCTTCTCCGCAGCGGTTGCCTATTACGACAGCTACC
 25 GTGCTGCTGTTCTGCCTGCGAACCTGATCCAGGCACAGCGTACTATTTGGTGGCATACTTATAAGCGTATC
 GATAAAGAAGGTGTGTTCCATACCGAATGGCTGGATTA

SEQ ID NO: 13 (Ejemplo de secuencia de nucleótidos de locus *O6A rfb* – *O6A*–producción EPA cepa stGVXN4112 y stLMTB10923)

ATGACGAATTTAAAAGCAGTTATTCCCTGTAGCGGGTCTCGGGATGCATATGTTGCCTGCCACTAAGGCGATACCC
 30 AAAGAGATGCTACCAATCGTCGACAAGCCAATGATTCAGTACATTGTTGACGAGATTGTGGCTGCAGGGGCAAAA
 GAAATCCTCGTGGTAACTCACGCGTCCAAGAACCGGTCGAAAACCACTTCGACACCTTATGAGTTAGAATCA
 CTCTTTGAGCAGCGCGTGAAGCGTCAACTGCTGGCGGAAGTACAGTCCATTTGCCCGCCGGGCGTGACAATTA
 TGAACGTGCGTCAGGGCGAACCTTTAGGTTTGGGCCACTCCATTTTATGTGCACGACCTGCCATTGGTGACAAT
 CCATTTGTCGTGGTGTGCCAGACGTTGTGATCGACGACGCCAGCGCCGACCCGCTGCGCTACAACCTTGCTG
 35 CCATGATTGCGCGCTTCAACGAAACGGGCCGAGCCAGGTGCTGGCAAAACGTATGCCGGGTGACCTCTCTGA
 TACTCTGTATCCAGACCAAAGAGCCGCTGGACCAGGTAAGTAAAGTCAAGCCGATTGTTGAATTCATCGAAA
 AACCGGATCAGCCGACGCTGGACTCAGACATCATGCGCGTTGGTTCGCTATGTGCTTTCTGCCGATATTTGG
 CCGAACTTGAACGCACTCAGCCTGGTGCATGGGGCGTATTCAGCTGACTGATGCCATTGCCGAATGGCGA
 AAAAACAGTCCGTTGATGCCATGCTGATGACCGGCCGACAGCTGCACTGCGGTAAAAAATGGGTTATAGCAA
 40 GCGTTGCTGAAGTATGGACTACGCAACCTCAAAGAAGGGGCGAAGTTCGTAAGGGGATTGAGAAGCTGTTAAG
 CGAATAATGAAAATCTGACCGGATGTAACGGTTGATAAGAAAATTATAACGGCAGTGAAGATTAGCGGCGAAAGT
 AATTTGTTGCAATTTTCTGCGGTTGTTTTATATAACAATCAGAATAACAACGACTTAGCAATAGGATTTTCTGTC
 AAAGTTTTCCAGGATTTTCTTGTTCAGAGCGGATTGGTAAGACAATTAGCATTGAAATTTACGGGTTTAGCG
 CGAGTGGGTAACGCTCGTCACATCGTAGACATGCATGCAGTGCTCTGGTAGCTGTAAGCCAGGGGCGGTAGC
 45 GTGCTGAAATTATAAAGTCACTTCTTATAGAATCGCATTCAATAATATAATTACACCTAAATGAATAGGATACAA
 CGTGTGCACAATTTAAAGGCTTAAAGATAAAAATAAAAAACGTATTTTAAAGGTTGATATATTGCAGTTATTTAA
 TTATATCGCGCCATTGGTAATTATCCCTATCCTGATAAAAATATATTGGGTTGGGGGAATATGGGGGAATTTGGTCTAT
 ATTACATCTATTTATCAAATAGTGGCTTTGATTATTGATTTTGGCTTTACTTACACAGGACCTGTGGTTGCTGCGA
 GACATAGATGTGAGACCCAAAATTTACAGCGCTATTACTCAATAGTTGTTCTTTTAAAATCATTGCTTTTTATAATT
 50 GCATTAACATGTGTATTTTTATTGTGCAGATTAATATAGTCCACTTGTCAATTTTTGGGTTTTTGTCAATTTTTCTA
 TGCATATTGGTAATATATTATCGCCCAATTGGTTTTTGCAGGGGATTGGTGATTTTAAAAAATTTTCATACTCACA
 AGTAATAGTGAGAATAACATTGTTTATCATACTTCTGTTTATGTCTGATGGCGGAGATAATGTTTTATCCTAA
 GTTTTTGCAAAATGCAACATTACTCATATGCTGTATATACTTATGCGCAAAATTTCAATAGCCATGTTGTTTCA
 CTTAAACCTAATGAATGCATTGTGGAATTTAAGAAGGCAGGAAATGTTTTTATTGGCGTAATAGGTACGATTGGTT
 55 ACAATGGTCTAATTCCTGTGTTAATTGAAAACCTTTGCGGTAATACGAGTCTTGGTGTTTTTTCAATCGTTCAAAA
 AATGACAACAGCATGTCAAAGTCTAATTAATCCAATATCACAGTATATGTTATCTCAAGTTTCAGAAATTAACCTC
 AAGATAAATGTTTTATTATAGAATTAATAAAGTTTTTTTTGTGCATTTAACAATTAGCATAATTGCATGTTTATGTT
 ATATGGGGTTAGGGCAATATGTGGCGACTTTTATAGGTAAGTTGACGTTTCAATTTGTTATTTTATTTTGCCTC
 AATAATTACCAATTTTTTCACTTTTAAATAATGCTTGGTATACAGTTTCTTATACCGACAGATAATGTAAAATACT
 60 ACGAAGTATAAATGTTATGGCGGAAATTATTGTTGTTAGTTTGTCTGGCTGTTAATATCACGCTTTTACAGATTTG
 GGGGGGTTTTTATAAACCTAATTGGTGAAGTTTCTTGTATTAGTATGCTAGCTTTTATTGCCATCGAAAGTGGG
 GAGCGAGAGTATAATGAAAGTGAAGGCGGTTCTGCTATTACATTCTATTTAAGTTAATGCTGACAATTTTAGTG

T T A C T G T T T G G T A A T G A A C C A A A T A A A T C A C A A T A T A T C C T T G T T A T A G C A A C G A T A A C A G T T T T T T A T A T C G C A T A
 T A T C A C T A A T A A A A T A A C T T C T C C G G C C A G C C T T C T C G T T A T A T C A T C T T T T G T G T T T T A G G T T G T C G C C C T T T A T
 T A T C T T T G T T T G C A A A C T A T G A T T A T A G G A T T G C C G A T T G G T T A T T G A A G G A T A T A T G G A T G A C G A T G T G A T T T
 5 G G C T A A C T A T G C T A T A A C A C T A A T G T A T T A T G G T T A T A C A T T G G G A C T A A T T C T A T G C A A A A A T A C T G A A A A T T T T
 A T C C G C A T G G T C C T T A T C C T G A A A A A C A A T T G C T A A A A A T A A A G T T T C T T T T G A C T T T A T T T T T C T G G G T T C G A T A
 G G T A T G G T T G T A A A A G G G A T A T T C T T T T T A A C T T T A T A G A A T C T A A T A G T T A T G T T G A T A T T T A T C A A T C A A A T A T A
 A C A A C G C C A A T A G G T T A T G A T T T T C T A C T T A T T T A T T T A T T T A T T G T T C T T T T T C C T T A T A T G T G C G T T T C A T A T A C A G
 T T C A G A A C A A A T A A A A A T T T C T T T T A T T G C G A T A T G C A T T G C T G C A T T T A G C A C C T T G A A G G G T A G T C G T A G T G
 A A G C T A T A A C G T T T C T T T T A A C G G T T A C A T G T A T A T A T T T T A A T G A A G T A A A G A C A A G A A A C T T A C G T C T G C T G A T T
 10 A C A A T G A T T T T T G T T T T A G C G T C A T T T T T G T G A T T A G T G A A T T T A T C T C A A T G T G G C G C A C T G G A G G G A G T T T T T
 T C A A T T A A T G C A G G G T A A T A A T C C T G T T A T A A A C T T T G T A T A C G G C A T G G G A G T A T C A T A T C T T T C A T T T A T C A A T
 C A G T A A A A C T A C A A C T A T T G T C A G G G G G A T A A T G T T A C C T A T C T A T T C A G C C A G T T A A T A A A C T T G C T C G T C
 A A T A T T A A T G T C A A A T T G A G C T T G C C G G A A A T A A G C T A T A G C C A T T T G C C T C A T A C A C A G C A A A C C C A G A A C T A
 T A A A T C T T G G G T T C G G A C T T G G G G G A G T T A T T A G C A G A A T C G T T T T A G C A T T T G G T A G C A T T T G G A T T G A T T C A
 15 T T A T A C C C T T T T T A C T T T T A C T T A A T T T A A A T G T A T T G G A A A A T A T A C A A A A A C A A A C C A A T T A T A T T T T G T T T A
 T T A T A G T G T G T T G C C A C C T A T A T T A T T C A C A C C A A G A G A G A C T T T G T T C T A T T T C T C C C C T A T C T T G T C A A A A G T A
 T A T T T G T T G C T T T T T A G T T A C A T T A T A C A T C C A G T A T A A A A A G G A T T G A C C A A A A T G T C A G A A A A A A T G T C A G C A
 T A A T A A T C C C A A G T T A T A A C A G G G C T C A T A T T C T T A A G G A G G T C A T A C C A A G T T A T T T T C A G G A T G A G A C T T T A G A
 G G T T A T A G T T A T C A A T G A T G G A T C A A C A G A T A A T A C A A A T A G T G T A T T A G C T G A A C T G A A G G A A A A A T A T T C T C A G
 20 T T A G T T A T T T T A G A A A A T G A A C G A A T A A A A A C A G A T G T A T T C T A A A A A C C G A G G G A A T T G A A A T A G C C A A A G G G A
 A A T A T A T T T T T T G T G T A T G A T G A C T C T T A C C T C T A C C C G G T T A T A T A T C T C G G T T A T T G G C T A C A A A A T A T T G A G
 A C A G G C G T G A T G T A A T C G G C G A A G A A T A C T T T A T A T G A A T A A A C G A G A A A C A A T T G A A G A T T G C A T A A A T C
 G A C A T A A A A A G A G G G G C G T T T T G T T A G T G A T C T A A A T A G A T T G G A T T T T A G T T A T A C A T G T G A T T T G G A C C A T C C
 G A T T G A A T G T T T T A T G C A C A G C C T T T T G T T C T A G C T G A A A G G G A A C T A A T A T C G A A A T A T C G A T T T G A T A T A T C T T
 25 A T A C G G G A A A C T G C T A T C G T G A G G A A A C T G A T T T C A T G C T A T C T C T A T T T A T T A A A A A T A A A A A T T T A T A T A T G A T
 T C A A A G G C T T T G T T A A A A T T T A C C T C C A A G A A A A G C G A C G G G A G G G G C A A G A A C A G C T A A T C G A T T A A A A T A T
 C A T T A C G A A A G T T G C A T A A A A T A A T T A T A G A T T T T A A A A A A T A T A A T G A T A A T T T G A A T C T T C T T T C A G G A C A A A A G
 C A T G C T A T A T T T T A C C G A C A G T G T C A A T T C G T C T G C T A A A A A T G A A G T C G T T T A T C G G G A A G T T T T A A A A T G A T T
 A T A T A T A T C G C C G C T A T A A T G G T T C A G G A G G G C A A G G T G G G G T G G A A A G G T T G T T G C C C A A C A A T G T A A C A T
 30 T C T T A A A A A T T T G G G G T T A A A G T C A T T A T A C T T G A T A A A A C A T A C T T C A A A A T T T C T A A C A A A A T T C G T A A C A A A A
 A A A T A C A A G T A G C A C T T T A T C C A A T A T T A G T T T C T T T A T T T A A C C T T A C A A A A A T T A C G T G G C G T G A C G T T T A A A
 G T T A T T G C A C A T G G C T A T T G T T C C T T T T A T A G G A A T G A C A T C T T A A T A G C T A T G G C A A T A T G A A A T G T T A T T T
 T C A A A C A G T C A T G A A T A A A A A C C T A A T C G G T T G T C T G G C A G T G G T C T T T T A T C T T T C T A T G A G C G T T G G G C T G G
 A G C A T T T T C A A A A A A T A C T G G G C T G T T C A A A T A A G G T T A A A A G T G A A T G G A A T G A G C T T T A C A A T A T T A A T T C A C
 35 A T A A A A T C A A A G T T T C G A A A T T T A A A A T C T T G C A C A A T T T G A T T A C A C T G A T G T T A A T G A A G C A G A A T A T G T G
 A C A T T T G T C G G G C G T A T T G G A A A A A G G A A A A G G A A T A G A T A G A T C T G T A T T A C A T A T G T A A A A A T C T G C C A G A T A C T T
 C C T T C C A T T T A G T T T C A A G T A T T C C C G C C C A C A A A A T T T T G C T T C G C T A A A T A A T G T T C T G A C C A G C A T T G C T G T
 C C C C T A T G C G A A A A T G C C A G A A A T A T T T A A G A A A T C C A G A G T A C T T A T T T T A C C G T C C T A T T A T G A A G G A T A T G A G
 C T G G T T A C T A T T G A A G C G C T A T G C T G T G G T T G C C C T G T G A T A G G C T A T A A T G T T G G T G C A A T T A G A G A G T T G T A T
 40 G C A G A A A G T T T T C C T G G C G T A T T T A T T G C C A A T A A T A A G A A G A T T T A G C A C A A G T A G C C T A C A A A T T A A T T A G T C
 T T G A T A A T G A A A A T A T T A T C A T T T G A G A C A A A C T A T T T A T A G C A A G C G T G A G C T T T T T T C T G A A G A G A G A T A T G C
 G G A A T T T T A C A C G G C G C A T T T A A T G A A A A A A A A A A A A A A C T G T C T C A T T T C A A T T A A C T A T A T A A T G A A C T T
 A C C G A G G A G A G A T A T A T T A C G T A C G C T T G T A G T T T T C T A C A A A A C A G A A T G T T A A T T T A A C A C T T A T T G A T A
 A A A A T C C T C A G G T A A A C T A T T C G A A G A C A A T A C T T T T C A A C A T A T A T C A T T T A T T A A A G G T A A A C G T C A G G A T A T A
 45 A T A T C C A G G C T T T T T T A T A C C A T C A T T T T A T G T C C T T A T A T T T T C T C A A T A A T T A A A A T T T A C G G A A G C A A G A T
 A T T C T T G C T T T T C A C A A C T C T C G G C T T G G A T T G T A T G T C T G C T T T T A G A A T A C T C A T G C C C C A C A A A A A G A T C A T
 A T T G T T T A C G G A T A A C T T C G A A T A T G A C T T A A T A A G A C A A A A G A T A A A A C A T A A C T A C T T T T A T T G A A A A T T A A T
 T G T T T A T C T C A A T G A A T T T A T C G G G C T T A A G A A T T C A G A T T T A G T T A G C T A T A T T A C C C G G C A A G A T A A A A A T G C A A
 T G G A T A A A T T T A T G G G A T T A A A A A A A G C A G A A A T T A A T T C C C C T G T G A T T T A G T A G A G A A A A C C A A C T G A T
 50 G T A T T G T C A G C T C A C T T T A T T A A T G A G T A T A A T C G A T T G A A T A A T G A T A A T A A T A A T A A G A A A A A A G T A G T A T T T A C T G C A T C
 T T T T G A T T T T T T C C A A A T A T A G A T G C T G C C A A C T A T G T T T T A A T G C A G C A A A G T C T A A A T A A T G A T T A T T G C T A T A T
 T T T G G C A G G T A G G A A A G T A C T A C T T T G A A T C T T C C T G A T T T G G A T A A T T A T T T T T T T C G A T A A T C T A T C T A A T A
 G T G A A A T G T C A T A T T T A T T A T C T G C T T G T G A T G T T T T T A T T C T C C T A T A G T T T T A G G A A G T G G A A T G A A A C A A A A
 A T T G C A G A A G C A C T A T C A T A T G G A T T A T A T T T A T G C G A C A G A G C A T T C C T T A A T C G G C T A T G A T G A A A T T A T A C
 55 A C A A T A A G G A G T G T G T T A A A A A A T C T C A C A T T T G G A T G A G G A A T T T C C T A A A G A T T T C A A G A T G A A A A G T A T C A A
 T A A C A G C T A A T A A T G T C T T A T C A G C A A A A A T A T T A C A T T A C A T T A C G G T T A A T G C C A T G A A C T T G A T A T A A T A A
 A T T T T G A C G A A T T A G T G G A G A T A T A A T A A C A T A T T A G T A A C T A C T G G T G G T G C T G G A T A T C G G A T C T C A A T A A
 C G G C T A T T G A A T T A C T G A A T G C A G G T C A T G A G A T T A C G T T C T G G A C A A T T T C A G T A A T G C T T C A T A C A A G T G A T
 C G A A A A A T A A A A G A A A T T A C T C G A C G T G A T T T T A T A C A A T T A C T G G A G A T G C T G G G T G T A G G A A G A C A C T C T C
 60 C G C T A T T T T C G A G A A A C A C G C C A T A G A T A T A G T T A T T C A T T T T G C T G G C T T T A A A T C T G T T T C A G A G T C T A A A A G T
 G A A C C C T T A A A G T A T T A C C A G A A T A A T G T T G G A G T G A C C A T T A C T T T A T T A C A G G T A A T G G A A G A G T A C A G A A T T A
 A A A A T T T A T C T T T A G T T C A T C T G C G A C A G T C A T G G T G A A C C A G A G A T A A T T C C A A T T C C A G A A A C A G C T A A A A T
 T G G A G G A A C T A C G A A T C C A T A T G G C A C A T C G A A G T A T T T T G T T G A A A A A T T C T A G A G G A T G T T A G T T C C A C G G G
 A A A A C T G G A T A T A A T T T G C T T G A G A T A T T T A A T C C T G T C G G T G C T A T T C T A G T G G T A A A A T A G G T G A G G C T C C A

TCTGGTATCCCTAATAATCTTGTTCCTTATTTATTGGATGTTGCGAGTGGTAAACGTGATAAATTATTTATTTATGG
CAATGATTACCCTACTAATGATGGAACAGGTGTAAGGGATTTTATTCATGTTGTTGACTTAGCGAAAGGTCATTTG
GCTGCAATGAATTAATTAAGTATCAATTCGGGATATAATATCTTAAATCTTGGTACAGGAAAAGGTTATTCGGTACT
TGAATTAATCACTACATTTGAAAAATTAACAAACATTAAGGTCAATAAATCTTTTATAGAGAGAAGGGCAGGGGAT
5 GTTGGCTCTTGTGGGCTGATGCAGATAAAGCTAATTCCTTTATTGGACTGGCAAGCCGAACAAACTCTAGAACAG
ATGTTATTGGACTCGTGGCGTTGGAAAAAATTTATCCAGACGGATTCTGAATATAAAAAGGTTTCAGTTTTATGAA
TCAATCAGAGCAGAGAAAAAATACTGGTTCTTACACCTCGCTTCCCTACCCTGTCTATTGGAGGGGATAGATT
AAGAGTCTATATGTTATGTAAAGAACTTTCCAAAAATATGATCCTATTCTTCTGAGCTTATGTGATCAACCACTAG
AACTTGAATAAATAAATGACTCGGTCTTCAAAGAAATTCATCGTGTCTATCTACCAAATAAATAATCATATTAT
10 AATGTATTAAGAAGCTTTGGTTACGCAAAAACCGTTGCAAATTGCTTATTATCAATCGGACACATTTAAGAATAAATA
CAATAAATTAATTAACAATGCGATGCAGATTTTTGTCATCTGATAAGAGTTGCTGATTATGTTAAGGATACAGAC
AAGTTCAAAATTCCTGATATGACAGATGCAATATCTTTGAATTACAGTCGCGTTAAAAAATTAGCAAGTAAAAAAG
TTTGGCGTGAATTAATTTCTCTGGAACAAAAAAGATTAGAATCATATGAACGTTCTGTGGCGAATCTTTTTGATT
TGACCCTTTTATTCCTAGACCTGACTATCTACCCTAATCTGGGCAGTAATTCATATGTAATAA
15 TGGGGTTGATACATCAGCCTTGAGATATAAAAAAGAGAAAATAAAAAATCGATAAGCCTGTGGAACCTTATATTTATC
GGAAATATGTATTCTTTACAAAATATGGATGCTGCAAAACATTTTGCTAAGAATATTTTACCTTGCTTGATGATGA
GTTTAATATTATTTTAAAGTGATTGGTAAGATCTCAGAACTAATAAAAAATATATTAATTCATTTAAAAATACAAT
TGCTTTAGGTAAGTGTGATGATCAATTCCTCCGCTTCTACAGGGCATATAGGTATATGTCCTGTTCTGCTTGGGA
GCAGGCGTACAAAATAAAATTCCTGAATACATGGCTTTAGGTTTACCATGTATTACATCTAGCATTGGTTATGAAG
20 GTATTAATGCAAAAACAGGTAGCGAAATTTTTGTTGCAGATACAGTACAGCAATATAAAAAACGTACTAAGAGAAAT
AATTTACGATAAATCTTACTGAAGTGGCTGAAAATGCCGCTAGTTTTGTAGAAAATAATTTTTCTTGGGAAT
CAAAAGTTGCCAATTAATGAATACATTTAGATGAGAAATTAATGAACAATAAATAAATTTTACACCTATCATTAT
GGCTGGTGGTTCAGGCAGTCGGTTGTGGCCACTATCAAGAATTCCTATCCGAAACAATTTCTTAGCCTAATCGG
TAGTCATACCATGCTTCAAACAACGGCTAATCGTCTGGATGGTTGGATTGTACCAACCCTTATGTCATTTGTAAT
25 GAACAATACCGCTTTATAGTTGCTGAACAGCTTAGAAAAATCGATAGATTGACTTCAAAGAATATCATCCTTGAGC
CTGTTGGGCGTAACACTGCCCTGCAATTGCATTAGCGGCGTTGCTGATGTCTAAGTCTGATAAAAGTGCAGAT
GATCTTATGCTCGTACTGGCTGCAGATCACGTTATACACGATGAAGAAAAATTTGTAACGCTGTTAGATCGGCA
ATCCATACGCTGCTGATGGGAAATTTGGTAACATTTGGTATAATTCAGACAAAGCAGAAACTGGTTATGTTATA
TACATCGAGGACAATAATTAATCAGGAAGATTCGGATGCATTTATAGTGTATCATCATTTGTTGAAAAGCCAAATCA
30 TGAGACAGCCACTAAATATCTTGTCTCCGGTGAAGTATTATTGGAATAGCGGTATGTTTTGTTTGTAGTGCAAAATCGT
TATATAGAGGAACTTAAACAATTTCCGCTGATATTTATCCGCTTGTGAAAAGCAATTTGCTTCAGCGAACTTTG
ACCTTGATTTTGTGCGTTTAGATGAAAGTTCTTCTCTAAGTGCCTGAAGAATCAATTGATTACGCTGTAATGGA
AAAAACAAAAGACGCAATTTGTTATTCCAATGGATGCTGGCTGGAGTGATGTCGGTTCATGGTCTTCTCTTTGGGA
AATTAATGATAAAGACTCAGACGGCAACGTAATAGTTGGGGATATTTTCTCTCATGAAACAAGAATTCCTTTTATA
35 TATCCCGAATCGGAAATTTGCTACAGTTGGAGTGGAAATTTAGTTGTTGTTCCAAACAAGGATGCTGTTCTT
GTCTCAGAGAGAAATAAAGTTTCCAGATGTAAGAAAAATAGTAGAACAATAAATAAATTTAGGTCGTAGCGAGCAT
TATGTTTATCGCGAAGTATATCGTCTTGGGGTAAATATGATTCCATTGACACAGGGGAGCGTTATCAGGTCAA
CGTATAACAGTAAATCCTGGTGAAGGACTTTCTTTACAAATGCACCATCATAGGGCAGAACATTGGATCATAGTTT
CTGGAAGTCAAGGGTACTATAGGTTCTGAAACTAAGATTCTTAGCGAAAATGAATCTGTTTACATACCTCTTG
40 GTGTAATACACTGCTTGGAAAATCCAGGGAAAATTCCTCTTGATTTAATTGAAGTTTCGTTCTGGATCTTATTTAGA
AGAAGACGATGTTATCCGTTTTTCCAGGACCGATATGGTCGTAGCTAAATTTTTGATAATGTAACGTTAGTAGAAGAG
CGTAAATTTTTAGTTAATCTGTAATAAGTATTTGTTAAGGTATATCATGTCGAGTTTACCCTGCTTTAAAG
CCTATGATATTTCCGGGAAATTAGGCGAAGAACTGAAATGAAGATATTGCCTGGCGCATTGGTCCGCTTATGGC
GAATTTCTCAAACCGAAAACCTTGTGTTAGGCGGTGACGTCCGACTACCAGCGAAACCTTAAAACCTGGCGCT
45 GCGAAGGGGTTACAGGATGCGGGCGTCGATGTGCTGGATATTGGCATGTCCGGCACCGAAGAGATCTATTTT
GCCACGTTCCATCTCGGCGTGGATGGCGGCATCGAAGTTACCGCCAGCCATAACCCGATGGATTACAACGGCA
TGAAACTGGTGCGCGAAGGGGCTCGCCCGATCAGCGGTGATACCGGACTGCGCGACATCCAGCGTCTGGCAG
AAGCCAACGACTTTCTCCCGTTGATGAAACCAAACGCGGTGCTATCAGCAAATCAATCTGCGTGACGCTTAC
GTTGATCACCTGTTCCGTTATATCAACGTCAAAAACCTCACGCCGCTCAAGCTGGTGATTAACCTCCGGAAACGG
50 CCGCGCGGGTCCGTTGGTGACGCCATTGAAGCCCGCTTTAAAGCCCTCGGCGCACCCCGTGAATTAATCAAAA
GTGCACAACACGCGGACGGCAATTTCCCAACCGGTATTCTAACCCTGCTACTGCCGGAATGTGCGGACGACA
CCCGCAATGCGGTATCAAAACACGGCGCGGATATGGGCATTGCCTTTGATGGCGATTTTACCCTGTTTCTG
TTTGACGAAAAAGGGCAGTTTATTGAGGGCTACTACATTGTGCGCCTGCTGGCAGAAGCGTTTCTCGAAAAAAT
CCCCGGCGGAAGATCATCCACGATCCACGTCTCTCTGGAACACCGTTGATGTGGTGACTGCCGAGGGCGGCA
55 CCCCAGTAAATGTGCAAAACCGGACACGCCTTTATTAAGAACGATGCGCAAGGAAGACGCTATCTACGGTGGC
GAAATGAGCGCCACCTACTTCCGTGATTTCCGTTACTGCGACAGCGGATGATCCCGTGGCTGCTGGTCCG
CGAAGTGGTGTGCTGAAAGGAAAAACGCTGGGCAACTGGTGGCGGACCGGATGCGACGCTTTCCGGCAAG
CGGTGAGATCAACAGCAAACCTGGCACACCCCGTTGAGGCGATTAACCGCGTGAACAGCACTTTAGCCGCGAG
GCGCTGGCGGTGGATCGCACCGATGGCATCAGCATGACCTTTGCCGACTGGCGCTTTAACCTGCGCTCCTCTA
60 ACACCGAACCGGTGGTGCGGTTGAATGTGGAATCGCGCGGCGATGACCCTGATGGAAGAAAAGACAAAAT
TATCCTTGAGTTACTGAACAAGTAATTCAGTAATTTTATATAAATGGGTTTTAAAAACGGAAAAGATGAGATATCC
GGTGTGGTATATCCAAGGTAATGCTATTCAGTATCTCTATGAGTGAGTTAACATCTATACCACATTTAAGCCGCAC
ACTTCGGGATCCCCATATGAATATCCTCCTTAGTTCCTATTCCGAAGTTCTATTCTTTCTAGAGAATAGGAACTT
CGGAATAGGAACTAAGGAGGATATTCATATGGATAAAGCCGTAAGCATATAAGCATGGATAAGCTATTTACTTT

AATAAGTACTTTGTATACTTATTTGCGAACATTCCAGGCCGCGAGCATTACGCGCGGTGATCACACCTGACAGGA
 GTATGTAATGTCCAAGCAACAGATCGGCGTAGTCGGTATGGCAGTGATGGGACGCAACCTTGCCTCAACATCG
 AAAGCCGTGGTTATACCGTCTCTATTTTCAACCGTTCCTCGTGAGAAGACGGAAGAAAGTATTGCCGAAAATCCAG
 5 GCAAGAAACTGGTTCCCTTACTATACGGTGAAGAGTTTGTGCAATCTCTGGAAACGCCCTCGTCGCATCCTGTTAA
 TGGTGAAGCAGGTGCAGGCACGGATGCTGCTATTGATCCCTCAAACCATATCTCGATAAAGGAGACATCATC
 ATTGATGGTGGTAACACCTTCTCCAGGACACTATTCGCTGTAATCGTGAGCTTTTACGAGAGGGCTTTAACTTC
 ATCGGTACCCGGTGTCTTGGCGGTGAAGAGGGGGCGCTGAAAGGTCCCTTCTATTATGCCTGGTGGCCAGAAAG
 AAGCCTATGAATTGGTAGCACCGATCCTGACCAAAATCGCCGCCGTAGCTGAAGACGGTGAACCATGCGTTACC
 10 TATATTGGTGCCGATGGCGCAGGTCATATGTGAAGATGGTTCACAACGGTATTGAATACGGCGATATGCAGCT
 GATTGCTGAAGCCTATTCTCTGCTTAAAGGTGGCTGAACCTACCAACGAAGAACTGGCGCAGACCTTTACCG
 AGTGAATAACGGTGAACCTGAGCAGTTACCTGATCGACATCACCAGAAATATCTTACCAAAAAAGATGAAGACG
 GTAACCTACCTGGTTGATGTGATCCTGGATGAAGCGGCTAACAAAGGTACCGGTAATGGACCAGCCAGAGCGC
 GCTGGATCTCGCGAACCGCTGTCGCTGATTACCGAGTCTGTGTTTGCACGTTATATCTTCTCTGAAAGATCA
 15 GCGTGTTCGCGCATCTAAAGTTCTCTCTGGTCCGCAAGCACAGCCAGCGCAGGCGACAAGGCTGATTCATCGAAA
 AAGTTCGTCGTGCGCTGTATCTGGGCAAAATCGTTTTCTTACGCCAGGGCTTCTCTAGCTGCGTGTGCGTCT
 GAAGAGTACAACCTGGGATCTGAACTACGGCGAAATCGCGAAGATTTTCCGTGCTGGCTGCATCATCCGTGCGCA
 GTTCTGCGAGAAAATCACCGATGCTTATGCCGAAAATCCACAGATCGCTAACCTGTTGCTGGCTCCGTACTTCAA
 GCAAATTGCCGATGACTACCAGCAGCGCTGCGTGATGTCGTTGCTTATGCAGTACAGAACGGTATTCCGGTTC
 20 CGACCTTCTCCGCAGCGGTTGCCATTACGACAGCTACCGTGCTGCTGTTCTGCCTGCGAACCTGATCCAGGCA
 CAGCGTACTATTTTGGTGGCGATACTTATAAGCGTATCGATAAAGAAGGTGTGTTCCATACCGAATGGCTGGAT
 TAA

SEQ ID NO: 14 (Ejemplo de secuencia de nucleótidos de locus O8 *rfb* – O8–producción EPA cepa stLMTB11734)

ATGACGAATTTAAAAGCAGTTATTCCTGTAGCGGGTCTCGGGATGCATATGTTGCCTGCCACTAAGGCGATACCC
 AAAGAGATGCTACCAATCGTCGACAAGCCAAATGATTCAGTACATTGTTGACGAGATTGTGGCTGCAGGGATCAAA
 25 GAAATCCTCCTGGTAACTCACGCGTCCAAGAACGCGGTGAAAACCACTTCGACACCTCTTATGAGTTAGAATCA
 CTCCTTGAGCAGCGCGTGAAGCGTCAACTGCTGGCGGAAGTACAGTCCATCTGTCCGCCGGGCGTGACCATTA
 TGAACGTGCGTCAGGGCGAACCTTTAGGTTTAGGCCACTCCATTTTGTGTGCGCGACCTGCCATTGGTGACAAC
 CCATTTGTCGTGGTACTGCCAGACGTTGTGATCGACGATGCCAGCGCCGACCCGCTACGTTACAACCTTGCTGC
 30 CATGATTGCACGTTTCAACGAAACGGGCCGACCCAGGTGCTGGCAAAACGTATGCCGGGTGACCTCTCTGAAT
 ACTCCGTCATCCAGACTAAAGAGCCGCTGGACCGTGAGGGTAAAGTCAGCCGCATTGTTGAATTTATCGAAAA
 CCGGATCAGCGCAGCGCTGGACTCAGACATCATGCCCGTAGTCTGCTATGTCTTTCTGCCGATATTTGGCC
 GGAACCTGGAACGTACTCAGCCTGGTGCATGGGGACGTATTCAGCTGACTGATGCTATTGCCGAGCTGGCGAAA
 AAACAATCCGTTGATGCAATGCTGATGACCGGCGACAGTTACGACTGCGGCAAAAAAATGGGCTATATGCAGGC
 35 GTTTGTGAAGTATGGCCTACGCAACCTGAAAGAAGGGGCGAAGTCCGTAAGGTATTGAGAAGCTGTTAAGCG
 AATAATGAAAATCTGACCGGATGTAACGGTTGATAAGAAAATTATAACGGCAGTAAAATTCGAGCAAAAGTAA
 TTTGTTGCGAATCTTCCCTGCCGTTGTTTTATATAAACCATCAGAATAACAACGAGTTAGCAGTAGGGTTTTATTCA
 AAGTTTTCCAGGATTTTCTTGTTCAGAGCGGATTGGTAAGACAATTAGCGTTTGAATTTTTCCGGTTTTAGCGC
 40 GAGTGGTAAACGCTCGTACATCAGGATGATGAGTCTGTTAGCTGTAAGCCAGGATGGGATAGCGTACGCG
 TGTGTAATACTCTATTAATCAAACGAGAGCCGCTTATTTACAGCATGCTCTGAAAGTAAATGGGAATAAATTA
 AGCTAGCGATCGCTTAAAGATCTAGGATTTCAATTATGTTACTTCTGTAATTATGGCTGGTGGTACCGGCAGTCGT
 CTCTGGCCGATGTCACGCGAGCTTTATCCGAAACAGTTCCTCCGCTGTTCCGGCAGAACTCCATGCTGCAGGA
 45 AACCATCACCCGACTCTCGGGCCTTGAATCCATGAACCGATGGTCATCTGTAACGAAGAGCACCGCTTCTGG
 TGGCTGAACAGCTACGCCAGCTCAATAAGCTGTGAATAATATTCTTGAGCCGGTCCGGGCGCAACACCGCC
 CCGGCCATCGCCCTGGCAGCCCTTCAGGCCACCCGCGACGGCGACGCCGCTGATGCTGGTTCTCGCCGCT
 GACCATATCATCAATAACCGTCCGCTTCCAGCAGCCATCCGGTCCGCGAGCAGTATGCTGATGAAGGTCA
 50 TCTGGTACCTTCGGTATCGTGCCGAATGCCCGGAAACTGGCTACGTTACATTCAGCGCGCGCTGGCGCTC
 ACCGATAGTGCCCATTCGCGTACCAGGTGGCCCGCTTTGTGGAGAAGCCGGATCGCGAGCGCGCCGAGGCT
 TACCTCGCCTCCGGGGAGTACTACTGGAACAGCGGCATGTTTATGTTCCGCGCCAAGAAATACCTCATCGAGCT
 GGCCAAATACCGTCCGGATATCCTGGAAGCCTGCCAGGCTGCGGTGAATGCCGCCGATAATGGCAGCGATTTT
 55 ATCAATATCCCGCATGATATTTTCTGCGAGTGCCCGGATGAGTCCGTGGACTATGCCGTTATGGAGAAAACCGC
 CGATGCGGTGGTGGTCCGTCTCGATGCTGACTGGAGCGACGTCGGCTCCTGGTCCGCACTATGGGAGGTGAG
 CCGGAAAGACGAGCAGGGCAATGTCCTCAGCGGTGACCGTGGGTACACAACAGCGAAAACTGCTACATCAAC
 AGCGACGAGAAGCTAGTGGCGGCCATTGGCGTAGAGAATCTGGTGAATTGTCAGCTAAAGGACGCCGTGCTGG
 60 TGATGAATCGCGAGCGTTCCAGGACGTGAAGAAGGCGGTGAGTTTCTCAAGCAGAACCAGCGCAGCGAGTA
 CAAGCGCCACCGTGAGATTTACCGCCCTGGGGCCGTTGCGACGTAGTGGTCCAGACCCCGCGCTTCAACGTC
 AACCGCATCACGGTGAACCCAGGCGGTGCCCTTCTCGATGCAGATGCACCACCATCGCGCCGAGCATTGGGTTA
 TTCTCGCCGGCACCGGTGAGTGACTGTCAACGGTAAGCAGTTCCCTGTTGTCGAGAACCAAGTCCACCTTTATT
 CCGATTGGCGCCGAGCAGCTGCCTGGAAAACCTGGCTGTATTCCGCTGGAAGTGTGGAGATCCAGTCCGGGG
 CGTACCTTGGCGAGGACGACATTATTCGTAATAAGACCAAGTATGGTGCCTTGTCTAATTTTTCCGGACAAAGC
 CAGAATTGACAGGTTAACTTGTTTTAAAGCTTATGACATCCGTGGTGAACCTGGGTGAGGAACCTGAACGAGGAC
 CGCTACCGTATCGGTGCGCCTACGGCGAATTTCTGAAACCCGGGAAGATAGTGTGGGGGGCGATGTGCGC
 CTCACAAGCGAGTCGCTGAAGCTGGCGCTGGCCCGCGGGTAAATGGACGCCGGTACCGACGTGCTGGACATC

GGCCTGAGCGGTACCGAAGAGATTTACTTTGCCACCTTCCACCTTGGGGTAGATGGTGGCATCGAGGTGACCG
CGAGCCACAATCCTATGAACTACAACGGCATGAAGCTGGTGC GCGAGAATGCGAAGCCCATCAGCGGCGACAC
CGGCCTGCGGGATATCCAGCGCCTGGCGGAGGAAAACCAGTTCCCGCCAGTGGACCCGGCGCGTCCGCGGAC
5 GTCCACTGAAGTTGGTGGTGAACCTCCGAAACGGGGCTGCGGGGCACGTGATTGATGAGGTGGAGAAACGCTT
CGCGGCGGCTGGGGTGCCGGTAACCTTTATCAAGTGCATCACCAGCCGGATGGCCATTTCCCTAACGGTATC
CCGAATCCGCTGCTGCCGAGTGCCGCGCAGGATACCGCCGACGCGGGTGC GCGAGCATCAGGCCGACATGGGG
ATTGCCTTTGACGGCGACTTCGATCGCTGCTTCTGTTTCGATGACGAAGCTTCGTTTATCGAGGGGTATTACATT
10 GTCGGCCTGCTGGCTGAGGCGTTCCTGCAGAAGCAGCCGGGAGCGAAAATCATTACGACCCGCGCTTGACGT
GGAACACGGTAGACATCGTGACCCGCAACGGCGGCCAGCCGGTGTATGTCGAAGACGGGGCATGCGTTCATCA
AGGAGCGGATGCGTCAGGAAGACGCTATCTACGGCGGGGAGATGAGTGC GCGACCATTACTTCCGCGATTTCCG
CTACTGCGATAGCGGGATGATCCCGTGGCTGCTGGTGGCGGAGCTGCTGTGTCTGAAGAACAGCTCGCTGAAA
TCGCTGGTGGCGGACCCGAGGCGTTCCTGCGTGGGAGAGATCAACCGCAAGCTAAGTAATGCTGCTG
15 AGGCGATCCGCGATCCGCGGCGAGTACCGCGGCTGCACACATCGACACAACGGAAGGATCAGTA
TTGAATACCCTGAATGGCGCTTTAACCTGCGCACGTCTAACACCGAGCCGGTGGTGCCTGTAACGTTGAGTCC
AGAGCTGATGTGGCGCTTATGAATGAAAAACGACCGAGCTGTTACACCTGTTAAGCGGGGAATAAGGTGAGAG
ATTTACTAACGACGATTTATCGTTATCGGGGATTTATCTGGAGCAGTGTAAACGCTGATTTTCAGGCACGCTATCA
AACTAGTATGCTGGGCGCACTATGGCTCGTTTTACAACCGCTCTCTATGATTCTGGTCTATACCCTGTTTTTCC
20 GAGGTGATGAAGGCAAGAATGCCGATAATACCGGGTCTTTGCCTATAGTATTTATCTGTTCCGGGGTACT
GACCTGGGATTAATTTACTGAGATGCTGGATAAAGTGCAGAGCGTATTTATTAACAATGCTAATCTGATCAAGAA
ACTCAGTTTTCCGAAAATCTGCTGCCGATCATCGTGACGTTATCGGGCGGTGCTAAATTTCCGCGATTAATTTCA
CTGTTTTCTAATTTTTATCATTGTCACCGGTAACCTCCCGGCTGGCTCTTTCTCTCGGTGATACCGGTCTGCTTT
TGCAGATCCTGTTTCCGGTGGGCTGGGGATGATCCTTGGTGCATGAACGCTTTTTTTCAGGGATGTGGGGCAA
25 CTGGTTGGCGTTGCGCTGCAATTCTGGTTTTGGTTCACACCCATTGTTTATGACTGAATTCATTACCTGCATGG
GCAAAAAATCTGATGATGATAACCCGATGACTCGGATCATGCAATCTTATCAGTCCATCTTCGCTATCATCTGG
CCCCAACTGGTATTCGCTATGGCCAGTATTGGCTCTCGCCATTATTTCTGCGTCATCGGTTTCAGGATGTTCC
GCAAGCATGCGGCGGATATGGTGGATGAATTATAATGAGTTATATCAGAGTAAATAATGTCGGTAAGGCGTATCG
CCAGTACTACTCAAAGACCGGGAGACTGATCGAATGGTTATCCCTCTGAATACCAAACGCCATAATTTGAAATCA
GATCCTCCGCGATTAATTTTCGAAGTTCGCTCCGGGCGAGGCTGTCGGTATTATCGGTATCAACGGTGCAGGCA
30 AGAGTACCCTGCTTAAACTCATAACCGGGACGTCCAGGCCGACGACTGGAGAAATTGAAATCTCCGGACGTGTC
GCTGCATTACTCGAATTGGGGATGGGGTTTCATTCTGATTTCACTGGTCCGCGAATGTTTATATGCTGGGCAA
CTGTTGGGGTTATCGTCAGAGAAAATAACTGAACTGATGCCGCAAATTAAGAGTTTGCTGAGATTGGGGACTAT
ATCGATCAACCTGTGCGCGTCTACTCCAGTGGGATGCAAGTTCGATTAGCTTTTAGTGTAGCGACGGCTATCCG
35 TCCTGATGTGTAATTCGATGAGGCATTACTGTTGGGGATGCATATTTCCAGCATAAAAGCTTTGAGCGTATT
CGAAAATTTCTCAGGAAGGGACACGCTGTTGCTGGTATCCCATGATAAACAAGCGATCCAAAGCATTATTCGCA
CCGGCCATTTTTATTGAATAAAGGCCAAATTTGAAATGGAAGGTGAACCTGAAGCAGTGAATTTTCAATGAGTTC
TCTTCTGGCCGATAAACAATAAGTCCATTAACAAGTTGAGCATAATGGTAAAACGCAAATGTTTCAGGCACT
GGTGAGGTGACTATCTCTGAGGTTTCTCTCGATGAACAGGGCAATGTGACTGAATTTGTTTCGGTAGGGCAT
40 CGTGTGAGCTTGCAGGTCAACGTTGAGGTCAAGGACGATATTCCTGAGCTTGTGTGCGGATATATGATTAAGGAT
CGACTTGGGCGAGCCGATTTTCGGGACCAATACGTACCATCTCAATCAGACACTCACCTCCCTGAAAAAAGGAGA
AAAGCGTTCTTCTTATTTCTTTTCGATGCGAGATTGGGGGTTGGCTCCTATTCTGTGCGTGTGCGGTTGCATAC
TTCCAGTACGCACTCGGCAAAAACTATGATGCGCGATTTGGCCGTGGTATTCAACGCTGTTAACACGTTAACCGAAC
AACAAGAGTTTGTGCGCGTGTCTGTTGCTGCGCTGAACCTGGAGATTTTCTAATGAGTTTCTGCTGTTTATCGTTC
45 ATTTGAAGAACGACACAGAGGTTGCGTTGAAGAAATCAAGCGCCGTTGAGTTTTTATTTACCTTTTCTTGCAGG
TCTGAAGGACATTTATCCTGATGGCGTGATTGCGGATATTGGTTGCGGACGTGGCGAATGGTTGGAGATCCTGA
CTGAAAATGGCATTGCGAACATCGGCGTCGATCTCGATGATGGCATGCTGGCGCGCGCCAGGGAGGCCGGACT
GAATGTGCAGAAAATGGATTGCTGCAATTTTTGCAAAGTCAGGCGGATCAGAGCCTGATAGCGTTGACCGGTT
TTCATATTGCTGAGCATTGCGGTTTGAAGTCTGCGCAACTCGCCATGCATACCCTACGGGTGCTGAAACCA
50 GGTGGTTTGTGATCCTCGAAACGCCGAACCCGGAGAATGTAAGCGTCCGACCTGTTCAATTTATATGGATCC
AACGCATAATCATCCTCTGCCACCGCCACTGTTGAGTTTTTACCTATTCAATATGGTTTTACCCGAGCAATTACC
GTTGCTGTCAGGAAAAAGAGGTTCTTCAATCTCCGGATGCAGCCGTTAATTTGGTGCATGACTCAAAGGGGT
GAGCCCGGACTACAGCATCATTGCTCAGAAAAGCAGCGCCAACAGATATTCTTGAACGCTTTGACACCCTGTTTAC
CCAGCAGTACGGTCTGACGCTGGATGCTCTGAGCAACCGTTACGATGCGATTTTTCGCCAACAGTTTTCTGTCG
TTGTCTCACGGCTGGAGACGTTGAACCAAACCTATATGCAACAGATAAGCCAAATGTCAGAGACTATTCAGACGT
55 TGCAAGGTGAGGTTGACGACTGAGTCAATGTCATCGATCAGAACCATCAGCTTCATCAGCAAATGGCGGATTTAC
ATAACAGTCTGTTACGGCGTATTACTCAACCACTACGCTGGTTGCTTTGCAACGCTCAATTAACGTCAGGAAG
GGGCTAAAGTCCGAGCCCGTAGGGCTGGGAAAAAAATATTGCGCAAAGGGATGGCCTCTGCTGTTCTTTTT
CCATCGTTACCTAAGTCTAAGGTTTATCTGTTTAAAGTTCTGAGAAAAACTGGCTGCTATACATTGCTACAACGT
TTGTTCCAACGCGTAATGCTGGTGAATCTGACACGATGATGATGCAGTCCAGAAGATATGATGTGGGACTGAA
60 GAAATGACAAGTCGCGCGATGAGTATTTATAACGAATTAATAAATAAATAACGGAGAAATAACGATGCGTATTGT
CATAGATTTACAAGGCGCACAGACGGAAAGCCGCTTTCTGTCGATCCGGTCTGTTATAGTATCGCAATCGCCAGAG
GCATAATCAGAAAATAACAGCCGGCATGAGATTTTCATCGCGCTATCCGCCATGCTGGATGAGTCGATTGCAATA
TTAAGGCGCAATTTGCCGATCTCCTGCCGGCAGAAAATATAGTCGATGGCATGCCGTAGGCCCTGTTCTGTCG
ATGGACCAAGGTAATGAATGGCGTCCGGAGAGCGCAGAACTGATTCGGGAAGCGTTTCTTGAATCATTGTGTC

AGATGTCGTTTTTATTACGAGTTTTGTTTGAAGGTCATGTCGACGATGCGGGCTACATCGGTACACAAATTTAGTCG
TCAGTATAAAGTAGCCGTACTGCACCACGATCTTATCCCCCTCGTGCAGGCGGAAACCTATCTGCAGGACGATG
TATACAAACCCTACTATTTACAGAAAGTTGAGTGGTTAAAAAACGCTGACCTTTTTGTTGACTAACTCTGCTTATAC
CGCACAGGAAGCGATCGAGCATCTGCATTTACAGGGCGATCATGTGCAGAATATTGCAGCCGCAGTCGATTCTC
5 AGTTTTGTATGGCGGAGGTGGCAGCGAGCGAAAAAGAGACCGTCTTTGGCCATTACGGTATTCAGCGCGAGTT
CATGTTGTATGCGCCCGGAGGATTTGACTCAAGGAAAAACTTTAAACGGTTGATTGAGGCCTATGCCGGGCTCA
GTGATGCCCTTACGTCGCAGTCATCAACTGGTCACTCGTCAGTAAGCTTTCCATCGGTGATCGTCAGTATCTGGAAT
CCCTTGCGTCAGGTAATGGTTTACAGCAGGGCGAACTGGTACTCACTGGTTATGTGCCGGAAGATGAGCTGATC
10 CAGCTCTATCGCCTATGTAAGCTGTTTCATCTTTGCTTCACTACATGAAGTTTTGGGTTGCCGGTTCTGGAAGCA
ATGTCGTGCGGTGCGCCGGTATTGGCTCAAATGTCACCAGTATTCCTGAAGTCATCGGTAATCCTGAGGCATT
ATTCGACCCGTATTCTGTCTCTTCCATGAGGGATAAGATCGCGCAATGTTTACTGATGATACCTTCTCGCGCG
TCTGAAAGAAATGGCGCAGCAGCAAGCGCGTAATTTCTTGGGATAAAGCTGCGGTGACTGCTCTGGAAGCTT
TCGAAAGATCGCGGTAGAAGACACCGGTACTGCGCAGGTTTTGCTGAAGCTTTGATTGAGGATCCTTGTCT
15 ATCTCACAAGGCGAGTAGACCGCATCTGCGTGTGCGCAACGGCCATTGATTAACAATCGAAAACGGC
AGAACTTTTATCAAATCGACGATAAATCGCTGAACTGGCGTGTGGAAGGCCATTTCGATAGCTCATATAGTCTGGC
GTTGGTCAACCGCAATTTGCCCGGGCACTCTCAGCCGATGGTGTAGAGTTTTATTGCATTCCACTGAAGGAC
CAGGTGATTTTGGCCAGATGCCTCGTTTATGGCACAGTCGGAATAAGTATCTTCTGGCATTTTATAATCAAT
GTCAGACCCGCAAGAGTAACGAAAGATAGATATTATTAGCAGAAATATCTATCCACCGCGGGTTACCAAATGG
20 ATGCCAAAGTAAATTCCTTCATTGTTATGCTTGGGAAGAAACGGGCTTCCGCAACCGTGGATCAATGAATTA
ATCGGGAACCTGACGGAGTGTGTGATTCGGAACATGTTTCGTAATAACTGATTGATAACGGACTGAATGTGC
CCGCAATTTGTTGTTGCAATGGTGTGACCATTTGCTCAATATCCAGCCGAGACGACAAAAGATGTGATCAC
GGAACATTCGTTTTCTGACAGTCTTCTTGTTCACGCAAAAGGATACAGGCAATGCTTCAGGCTTGGGG
GAAGGCGTTCACCTCGTCGTGACAATGTTATCTTAATCATTAAAGACTTTTTACAATCCGCACAATGAAATTGACGCA
25 TGGCTGGCTCAGGCCAGGCTCAATTCATAGACTATCCCAAAGTTGAAGTGATCAAAGAGGATATGTCAGCCAC
CGAGCTTAAAGGGCTTTATGAAAGCTGTGATGTTTTGGTTGCTCCAGGTTGCGCTGAAGGCTTTGGTTTACCTAT
TGCTGAAGCAATGCTGAGTGGGCTACCGGCTATCGTCACCAATTGGAGCGGGCAACTTGATTTTGTAAATTCACA
AAATTCATGGCTGGTTGACTATCAGTTCACCTCGGGTAAAAACGCACCTTTGGTCTGTTTTCTCAGCCTGGGCCAG
TGTGGATATTGACAACTTAACAGATGCATTAAGCGGCAGCCTCAACCGATAAATCAGTGTGCTGACATGGC
30 CAATGCTGGTCGCGAGCTTCTTCTGACAGGTTTTACCTGGAAAGCGGTGGCTGATCGTTCTTCCAGGCGGTCA
AGACTCTGCGTGCGCATATTGATATTGCACAGCATCGGGCGCGCATTGGCTGGGTGACGACCTGGAACACGAA
ATGTGGGATCGCAACCTATTCCCAGCATCTGGTGGAAAGCGCACCTCATGGCGCGGATGTTGTTTTTGTCTCCC
AGGTCAGCGCTGGCGATCTTGTGTGTGACAGACGAAGAGTTTGTACTTCGCAACTGGATTGTAGGTAAGAGAGC
AACTATCTGGAACCTCCAGCCACACATTGATGCTCTGAGACTCGATGTCATTGTGATCCAATTCAACTATGGA
TTCTTTAATCATCGAGAACCCTGTCGGCGTTTATTCGTGCGCCAGCATGACGCCGGTTCAGTTGTTATGACGATG
35 CACTCAACTGTGGATCCGCTGGAAGAGCGGAGCTGGAATTTCCGTCTTGGTGAATGAAAGAGGCGCTGGC
ACTTGGCAGCGGTTGTTGGTGCATTGATTGCCGATGTAACCGCCTTAAAGATTTAGGCTTAACTGCGAATGT
TGCTTTATCCCGCACGGTGTATCAACTACTCCGACGAGCGTACACGTC AACAGCAGTCTTTACCGCTAAT
TGCGAGCTATGGCTTCTGCTTACCGCATAAGGGCCTGATGGAACAGTAGAATCCGTCCATAGACTCAAGCAAG
CCGGTAAACCGGTTGTTTACGACTGGTGAACGCAGAGTATCCTGTTGGGGAGTCACGCGATCTGGTGGCAGA
40 GCTTAAAGCTGCTGCTCAGCGTTAGGTGTTACCGATCTGATTGAGATGCATAATGATTTCTACCTGATGCGGA
GAGTCTGCGGTTGCTTTCAGAAGCCGATCTTCTGATTTTTGCTTATCAGAATACTGGGGAGTCTGCTAGCGGGG
GCGTACGTTATGGTATGGCGACTCAAAAACCTGTTGCGGTAACGCCCTGGCGATATTTGATGATTTGGACGAT
GCCGCTTTTAAATTTGATGGATGCAGCGTGCATGATATCAGTCAGGGGATTGACCGGATCTGAATTTCCATCCGT
GAACAGAACTCTTGGGCAACCAGGACTCAACAACGTGCCGATGCATGGCGGGAACAACATGATTATCAAGCTGT
45 TTCACGCCGCTGTTAATATGTGTCAAGGCTTAGCTAAAGCTAAATATTTTAAATAAAAATATCTCTTGTATTT
TTTGCTTTGAATACAAGAGGGGTTAGATAATGTGTCAATTTATTATGAAAATTATTTTGTACTGAGCCAATTTAA
TACCCATTAACGGGCATCGGTGCGTATTCCCTGGAGCTGGTTAAGCGGCTGGCGGTCGCCCGGAAATTTGAAG
AATTAAGCTATTTACGGTGCCTGCTTTATAGAACAGATCCCTTTGGTGGAGAATAAAGCGGATACCAAAGCCA
GCAATCATGGTGCCTGTGTCGGCGTTTCTACGCCGACAGACGCTGTTGATTGAGGCTTATCGCTTGTGCTGATCCG
50 CGGCGCCAGGCGTGGGCAATTGCGCGACTATAAGGATTATATCTACCATGGCCCCAATTTTTATCTGCCGCATAA
ACTGGAACCGCGCGTACCACGTTTTATGACATATCCATTTTTACCTGCCCGGAATATCATCCAAAAGATCGGGT
TCGCTATATGGAGAAGTCCCTGCATGAGAGTCTGGATTCCGGCAAAGCTGATCCTGACCGTTTCTGATTTCTCGC
GCAGTGAATTTATCCGCTTGTTCAACTATCCGGCGGAGCGGATCGTAACCACCAAGCTAGCCTGCAGCAGTGAC
TATATCCCACGCAGCCCGGAGAGTGTCTGCCGTAAGTGCAGAAATATCAGCTGGCGTGGCAGGCCACGCGC
55 TATATATCGGCACTATGGAGCCACGTAATAATCCGAGGCCTGCTGCATGCCTATCAGCTGCTACCGATGGAG
ATCCGCATCGCTATCCGCTAATCCTTAGCGGCTATCGCGGCTGGGAAGACGATGTGCTGTGGCAGTTAGTCGA
GCGCGTACTCGGGAAGGCTGGATCCGTTACCTCGGATATGTTCCGGATGAAGACGATGCCGCTATCTGTACGCA
GCGGCCAGAGTCTTTGTTTATCCCTCCTTCTACGAGGATTCCGTTTACCTATTCTTGAAGCGATGCTTGGCGGT
GTGCCGTTAGTATGCTCCAATGTCACCTCTTTGCTGAGGTTGTTGGCGATGCCGGCCTCGTTGCCGATCCTAA
60 TGATATAGACGCGATTAGCGCGCAAATTTTGCAGAGCCTGCAAGATGATAGCTGGCGGGAAATCGCCACCGCG
CGCGGTCTTGTCTAGGCGAAACAGTTTTCGTGGGAGAACTGTGCGACACAGACCAATAACGCCTATAAATTA
CTAAGGGTGTGAGTTGAGAGTTTACACGCTATAAGACTTACTATCCCGATACCTACGGCGGATTGAGCAGGT
CATTTATCAGCTAAGTCAGGGCTGCGCCCGCGGGGAATCGCAGCCGATGTTTTCACTTTTAGCCCGGACAAAG
ATACAGGTCTGTGCTTACGAAGATCATCGGGTCAATTTATAATAAACAGCTTTTTGAAATTGCCTCCACGCGCTT

TTCGCTGAAAGCGTTAAAGCGTTTTAAAGCTGATTAAGATGACTACGATATCATCAACTACCATTTTTCCGTTTTCCC
 TTTATGGATATGCTGCATCTTTCCGGCGCGCCTGACGCCAGGACTGTGGTGACCTATCACTCTGATATAGTGAAA
 CAAAAACGGTTAATGAAGCTGTACCAGCCGCTGCAGGAGCGATTCTCAGCGGCGTAGATTGCATCGTTGCCTC
 5 GTCGCCCAATTACGTGGCTTCCAGCCAGACCCTGAAAAAATATCTGGATAAAACGGTGGTGATCCCGTTTTGGTC
 TGGAGCAGCAGGACGTGCAGCACGATCCGCAGAGGGTCGCGCACTGGCGGGAACCTGTCGGCGATAAAGTTCT
 TTCTCTTCGTCCGCACTTTCCGCTACTACAAAGGGCTGCATATTCTGATGGATGCCGCTGAGCGTAGCCGACTG
 CCAGTGGTGGTTAGAGGGGCGGGCCGCTGGAATCGGAAGTGCGGCGTGAAGCGCAGCAGCGCGGGCTGAG
 CAATGTGATGTTTACCGCATGCTCAACGACGAAGATAAGTACATTCTCTCCAGCTCTGCCGGGGCGTGATT
 10 CCCCTCGCATCTGCGCTCTGAGGCGTTTGGCATTACGTTATTGGAAGGCGCACGCTTTGCAAGGCCGCTGATCT
 CTTGCGAGATCGGTACAGGTACCTCTTTCATTAACCAGGACAAAGTGAGTGGTTGCGTGATTCCGCCGAATGAT
 AGCCAGGCGCTGGTGGAGGCGATGAATGAGCTCTGGAATAACGAGGAAACCTCCAACCGCTATGGCGAAAACCT
 CGCGTCGTCGTTTTGAAGAGATGTTTACTGCCGACCATATGATTGACGCTATGTCAATCTCTACACTACATTGC
 TGGAAAGCAAACCTGAGCGGCCGCGAGCTCGTCACTCGAGGATCCGTTAGGCTGGAGTGCCTTCCGAAGTT
 CCTATATTTCTAGAGAATAGGAACCTCGGAATAGGAACCTAAGGAGGATATTCATATGGATAAAGCGTAAGCAT
 15 ATAAGCATGGATAAGCTATTTATACCTTTAATAAGTACTTTGTATACTTATTTGCGAACATTCCAGGCCGCGCAT
 TCAGCGCGGTGATCACACCTGACAGGAGTATGTAATGTCCAAGCAACAGATCGGCGTAGTCGGTATGGCAGTGA
 TGGGACGCAACCTTGCCTCAACATCGAAAGCCGTGGTTATACCGTCTCTATTTTCAACCGTCCCGTGAGAAG
 ACGGAAGAAGTGATTGCCGAAAATCCAGGCAAGAAACTGGTTCCCTACTATACGGTGAAGAGATTTGTGAATCT
 CTGGAACCGCCTCGTCGCATCCTGTTAATGGTGAAGCAGGTGCAGGCACGGATGCTGCTATTGATCCCTCAA
 20 ACCATATCTCGATAAAGGAGACATCATCTGTTGTTGTTAACACCTTCTCCAGGACACTATTCGTCGTAATCG
 TGAGCTTTCAGCAGAGGGCTTTAACCTCAGTACCGTACCGGTGTTTCTGGCGGTGAAGAGGGGGCGTGAAGGT
 CCTTCTATTATGCCTGGTGGCCAGAAAGAAGCCTATGAATTGGTAGCACCGATCCTGACCAAATCGCCCGCT
 AGCTGAAGACGGTGAACCATGCGTTACCTATATTGGTGGCGATGGCGCAGGTCACTATGTGAAGATGGTTCACA
 ACGGTATTGAATACGGCGATATGCAGCTGATTGCTGAAGCCTATTCTCTGCTTAAAGGTGGCCTGAACCTCACCA
 25 ACGAAGAAGTGGCGCAGACCTTTACCGAGTGAATAACGGTGAAGTGAAGCAGTTACCTGATCGACATCACAAA
 GATATCTTCCAAAAAAGATGAAGACGGTAACTACCTGGTTGATGTGATCCTGGATGAAGCGGCTAACAAAAGGT
 ACCGGTAAATGGACCAGCCAGAGCGCGCTGGATCTCGGCCAACCCTGTCGCTGATTACCGAGTCTGTGTTTG
 CACGTTATATCTCTTCTCTGAAAGATCAGCGTGTTCGCCGATCTAAAGTTCTCTCTGGTCCGCAAGCACAGCCAG
 CAGGCGACAAGGCTGAGTTCAATCGAAAAAGTTCGTCGCTGTATCTGGGCAAAAATCGTTTTCTTACGCCAG
 30 GGCTTCTCTCAGCTGCGTGCTGCGTCTGAAGAGTACAACCTGGGATCTGAACTACGGCGAAAATCGCGAAGATTTT
 CCGTGCTGGCTGCATCATCCGTGCGCAGTTCCTGCAGAAAATCACCGATGCTTATGCCGAAAATCCACAGATCG
 CTAACCTGTTGCTGGCTCCGTAACCAAGCAAAATTGCCGATGACTACCAGCAGGCGCTGCGTGATGTCGTTGCT
 TATGCAGTACAGAACGGTATTCCGGTCCGACCTTCTCCGCAGCGGTTGCCTATTACGACAGCTACCGTGCTGC
 TGTCTGCCTGCGAACCTGATCCAGGCACAGCGTACTATTTTGGTGGCATACTTATAAGCGTATTGATAAAGA
 35 AGGTGTGTTCCATACCGAATGGCTGGATTA

SEQ ID NO: 15 (Ejemplo de secuencia de nucleótidos de locus O15 *rfb* – O15–producción EPA cepa stLMTB11738)

ATGACGAATTTAAAAGCAGTTATTCCTGTAGCGGGTCTCGGGATGCATATGTTGCTGCCACTAAGGCGATACCC
 AAAGAGATGCTACCAATCGTCGACAAGCCAAATGATTCAGTACATTGTTGACGAGATTGTGGCTGCAGGGATCAA
 40 GAAATCCTCCTGGTAACTCACGCGTCCAAGAACGCGGTGCAAAACCACTTCGACACCTCTTATGAGTTAGAATCA
 CTCCTTGAGCAGCGCGTGAAGCGTCAACTGCTGGCGGAAGTACAGTCCATCTGTCCGCCGGGCGTGACCATTA
 TGAACGTGCGTCAGGGCGAACCTTTAGGTTTAGGCCACTCCATTTTGTGTGCGGACCTGCCATTGGTGACAAC
 CCATTTGTCGTGGTACTGCCAGACGTTGTGATCGACGATGCCAGCGCCGACCCGCTACGTTACAACCTTGCTGC
 CATGATTGCAGTTTTCAACGAAACGGGCCGACGAGGTGCTGGCAAAACGTATGCCGGGTGACCTCTCTGAAT
 ACTCCGTCATCCAGACTAAAGAGCCGCTGGACCGTGAGGGTAAAGTCAAGCCGATTGTTGAATTTATCGAAAA
 45 CCGGATCAGCCGACGCTGGACTCAGACATCATGGCCGTAGGTGCTATGTGCTTTCTGCCGATATTTGGCC
 GGAACCTGGAACGTAATCAGCCTGGTGCATGGGGACGTTTACAGTACTGATGCTATTGCCGAGCTGGCGAAA
 AAACAATCCGTTGATGCAATGCTGATGACCGGCGACAGTTACGACTGCGGCAAAAAAATGGGCTATATGCAGGC
 GTTTGTGAAGTATGGCCTACGCAACCTGAAAGAAGGGGCGAAGTCCGTAAGGTATTGAGAAGCTGTTAAGCG
 AATAATGAAAATCTGACCGGATGTAACGGTTGATAAGAAAATTATAACGGCAGTGAATAATTCGACGAAAAAT
 50 TTTGTTGCGAATCTTCCCTGCCGTTGTTTTATATAAACCATCAGAATAACAACGAGTTAGCAGTAGGGTTTTATTCA
 AAGTTTTCCAGGATTTTCTTGTGTTCCAGAGCGGATTGGTAAGACAATTAGCGTTTGAATTTTCCGGTTTTAGCGC
 GAGTGGTAAACGCTCGTACATAGGCATAGGCATGCTGCTGTTAGCTGTAAAGCCAGGGGCGGTAGCG
 TGCATTAATACCTCTATTAATCAAACCTGAGAGCCGCTTATTTTACAGCATGCTCTGAAGTAAATGGAATAAATTA
 AGCTAGCATGAGCAAACTAACTAAATGTTCTTTACCTTGCAATAAGTCAGGGTGGCAATTACCTACTGCCATTA
 55 TTAATTTTTCTTATCTTGTAGAGTCATTGGTGTATCGAATTTTGGTGTATGAGTTTTTTCATTGATAACTATAAA
 GTGTTGTTAATGGTTGTTGAATATGGTTTTGGATATAGTGGGACAAGAGAAATAGCACTAATAACGATAAAAAAT
 ACCATTCTGAATTTTTTGGCGGTGGTGGCTTGGCTCGTTTTATTAATGCTAATTGCAGCTATAATACTCATAATA
 CTCTGTTTTTTTTATGTTTTAACGACGTTAAGTCTTTGTTATGTGTTGGTTTTCTGTCCGTAATTGCAGGTGTTTT
 CAATCCAAATGGTTTTGCAAGGTAAGGAAATGATGAGTGTGATGGCTGTGCTGCTACTATTTTACGAGGCAT
 60 AGCAGCTGTGCAAGTTTCTAATTAATAAACCCGCAACGCGGATGTACATCAGTGCCTTATTATTAGCATGCC
 ATATATTTTGTATTCTTCTGTGGCGTTGCCTACTTACTTATTATCAAGGAGATTTTTTTTATGTAGGCCACCGATAA
 AGAAAATCAAGTAATTTTAAAAAATGGATTTCATTTTTTTTGTTCAACACTTGGCAGTGTGCATACACAATGTTG

ACCCCTCTTGATTGGGTGGCGTATCTGGAAAGTTTGATGTAGGCATCTTTAACTCAGCTAACATGATCAAACAA
GGTTTGGCTGGACTTGCATCACCATTAGTCCAAGCTTTTTATCCAAGAATTAACATTTTGC AAAGAGAGAATCCAT
ATATTGCAAACCTAAAATCTAGAATGATTCTTAAATACTTGCTTGTTTTTACATGGCTTTAGCAATACCATTTTAC
TTTTTGCCAACCAATTATCATTATTAATATTCGGGCATGAAAGGTGAAGTAATTGCAGGTGCAATGCAATTAATGAC
5 ATTGCTTCTATATTCATAGGTTTTAATACAGTTGTCCGGTACTTGTATTAGTACCTAATGGGATGCAAAAAACAG
TATTTCAAATCTATTTTCTAGGAACCTATTACTGTTTAAAGCATAGTTTATCCAGCATGTAATATTATGGAGCAAC
GGGTGCGATTGTGAGTCTTATTGTAGCTGAAATTTTCGTTGGCATGGGAATGCTTAAACAATTCATTAAGTAAAT
AAAACCGTATGTAGGCCTCATAAATTATGAATATCTCGGTAATAATATCTGTTTGGAAACGCCAGTTCAATTAGA
ATTGATTCTCTCTGAGCTCGATTCTCAGGCTAAAGACAATAGTCTACACCTAGAAGTAATTGTTCCGATAGTCAT
10 AGTGGTAAAGAAATGATGATGTAGTTGCTGATAATATTCATAAAAAGAAAAATATTAATATTATCCATCAACATAC
TAAAAATATACTCTCCGCTAAGCGCAATTTCCGGAGCATCCCTAGCCCATGGGGATTATTTAATATTTCTTGATGAT
GATTGTATACCCGCAAGTGGATATATATCATCGTTGCTGAACTATTTAAAAAAAATGAATAGTAAAAGCGTTTTAT
GTGGGGAAGTTAGATTGCAAAAATGAACTCATTGAGACCAGCAATTAATCTGCTACAGGAACCTTTACACCCTA
AGTTTGTAGTGCCTGATATCTCTATGAATGCCTGGACTTTTGCGAATGAATTGTCTTGATGATAAGGAGGC
15 ATTTTCATCAGGTATAGTTTTCATATAATGAAAATTTTATTGGTTATGGTTGTGAAGATCATGAGTTTGGGTGGCAAC
TTGAAAAAATGACTTCAAATTTTGTCTGATTTTAAAAATTTACATCACGAATACAGTGGCGATATAGAAGGA
TATACAAAAAATTCGTGCTACAGCACGTGATGGTATGAATGTATTAAGCAAAGTAAGGCCTGAAATGTTTTCTA
CTAATAAAAAATTTCCCTAGTTGAGAAAATATTTAGTAAACACAAAACGTTTAGTAAAAATATGCCAATCAATTTTT
TCAATAAATTTATTTTAAAAAATAATACAATTTTAAAAAAAACAGATGCAAAATAAAAACTCTATTTCCCAATTCT
20 TTACAGATATGTGTTGATTTCCGGCATATACATCGTATTGGAGAGCGTGGCACCTCAAAAACAGATGATTTGCTT
AAGAATGGTATATATAGATGATCTTATTATTAAGACATTTGTATGGAAGGTAAAAACAATGAAGTATA
ATGCATTGATGGCTTTTTTATTATTTTTTGTGTTTTTTTTAGATTGTCGCTGATAAATACCTTTCTTATATTTGGCATT
TATTCCTGCATTTTTTGGTATTATGATTTAGTGCATAATTTTATGATTACTATGGGCAATGGATTGGTATCTATAG
ATCGTAAAAATTTGTTGCTGTTATCTATATTCATAATTTTTTTTTATTTTGTGTTTTCGATTTGTTTCAAAAAAG
25 CCATTCTTTTCAAAGTTATTTTACCGTTAGATTATTTATGTTGTTTTATTTTCATTTGTTCTGCGTATTATTTAGTA
AATAGATTCAAAAGGGTGACTTGAAATTAATGGAGCGAATATTAGTGTATTCTCTCGGGTTCAAATAGTTATTTT
TTTTGGTATGTATAAAGTCCAGAGTTAAAAAGATTGTTATATACTTTCTTTGGTATGTCTGACTCTGTTAATCTTT
GGGAACAAAATGCTAAAGTAAGAGGATTTGGTTGTCGGGTGAAATAAATTTTCATGACACCATTTTTGATGATCTA
TATGCTATTTTTATGATGAAAAGCGTTATGCTTTAATTACTTTAATTTGTCTGACTCAAATCGTAAATTCATAACAT
30 GGCTGTGATTGCAGCCATTATTGGTATCGGTTGCTCTAGACTTAATATTAATAAAAAATTGCAACAGTATTGATTT
TGGGAGTTTTAGTTTATAGCTTAGGAGCGGTGTTCTTTCTCGATTTTATGATGAGTTTCGTTTCTGGAGATGGCAC
AAGAACTCTGGATATCTTATTACAGCAACATGTGTTTGTGTTAGGTAATTTAGATTTTTTAAATATTATATTTGGATT
ACAGCAAAAACATATCTTCATCAATCCCGATATTAACAAAGTTCCGGATATGGGCTGGGTTATACTGTTTAAATTAC
GGTGGTTTAAACATTTTATACACTCTTTTTATTTTTAACTTTACTATTTCTATTGCGACATTTGGAATGACATATCAA
35 GCAATTTATATGGATGTTAATTTGGGATAATTTCAATACCAAAGGTTTAGTTTTAGGATCTAACGGCTATTTCTTTCT
ATCTTTTATATATGTTTTGAATAGAGTAACACTTAGTGGACAGAGTTCAATTAATAAAGTTAGGTTCAAGTAA
GTAAATAGCTTCCAGAGTATATTTGTCAATGATTTGAGGTTCCGGTTATTATGTTTTCATCTAAAACACTGTTAATTA
CTGGTGGTACTGGCTCTTTCCGGGAATGCTGTATTAATAAGATTTCTTGATACAGATATTGCAGAAATCCGTATATT
TAGTCGTGATGAAAAAACAAGATGATATGCGGAAAAAATACAATAATCAAAAATTAAGTTCTATATTGGTGAT
40 GTCAGAGATTACCGTAGTATTTGAATGCGACTCGCGGTGTTGATTTTATATATCATGCAGCGGCCTTAAGCAA
GTTCCATCATGTGAATTTTCACTCTATGGAAGCCGTTAAACTAATATCCTTGGTACGGAAAAATGTTCTTGAAGCAG
CTATAGCGAATGAAGTGAAGAGGGTTGTATGCCTAAGTACTGATAAAGCTGTATACCCGATTAACGCAATGGGTA
TTTTAAAAGCTATGATGGAAAAAGGTCATGGTTCGCGAAATCCCGTAATGTTGATCGCAATAAAAACAGTAATATGTG
GTACCCGTTATGGGAATGTTATGGCATCTCGCGGTTCCAGTTATTTCCATTATTTGTTGATCTTATTAGAGCGGGCAA
45 GCCACTCACAACTGATCCTAATATGACCCGCTTTATGATGACTCTTGAGGATGCGGTAGATTTAGTTCTTTTAT
GCGTTTGAACATGGTAATAATGGTGATATCTTTGTGCAAAAAGCACCTGCAGCAACTATTGACACATTAGCTATTG
CTTTAAAGGAATTAATAATGTTCTGACCATCCGGTAAATGTCATTGGAACCGGTGATGGCGAGAAATTTATATGA
AGCTCTACTTAGTCGTGAGGAAATGATCGCTGCTATAGATATGGGCGATTATTACCGTGTCCCGCCAGATCTTCG
TGACCTTAATTTAGGCAAAATGTTGAGCAAGGTGATGCCGAATATCTGAAATAGAAGATTATAACTCTCATAAT
50 ACTCAACGGTTAGATGTTGAAGGCATGAAAGAGCTCTTGCTAAAATTAGCCTTTATTCCAGAGCAATTCGTGCTGGT
GAAAAATATAATCTGGATTCATGATATGAAAAATTTAGTTACTGGTGCAATGGTTTTATTGGTCTGAATTTATGTT
TGAGGCTTGAGGAATTTGGTTATAAAGATCTTATTAGAATTGATCGAGAATCAACGAAGCAAGATCTTGAACAAG
GCTTACAGGATGCCGATTTTATTTATCACTTAGCTGGTATCAATAGACCTAAGACTGATGATGAGTTTATTTCTGG
AAACAGTGATTTAACAAAGCATATAGTTGAGTATCTCCTTTCTATTGGTAAGAATACACCAATTATGCTAAGTTCTT
55 CGATAACAAGCTGAACTTAATAATGCTTTAGGGGTTAGCAAAGCTGTAGCTGAAAGCTATGTGCAAAAAATATGCTG
CTGCTAGTGGTTCTTCTGATTAATTTTTCAGATATCCAAGCTTTTTTGGTAAATGGTGAAGCCAACTATAATTTCT
TTTTAGCAACTTTTCTACAATATTTTCCAATGATATTGAGATTACTATCAATGATGCAGCAGCCAGTCAATCT
GGTCTATATTGATGATGTTTTGACTGATGCTATAGCTCTTCTCTCTGGGACGGTTGAAAGTGGATATAAAGTTGTT
GCACCAATTTATTCAACAACAGTTGGTGAAGTTGCAGAATTAATTTATAGCTTCAAAAATAGCCGTTCCACCCTGA
60 TCACAGAGGCTGTCCGGGCGGGATTTACCCGTGCATTGTATTCTACATGGCTGAGTTATTTACCAGCAGAGAAG
TTTGCGTACAAGGTACCTTTTTATGGGGATGCCCGCGGAGTCTTTTGTGAGATGTTGAAAACGCCCTTCAGCGGG
GCAGTTTTCATTTTTACTGCTCACCCCTGGTATTACCGGTGGCGGACATTACCATCACAGTAAAAATGAGAAGTTT
TTGGTCATTCCGAGGTCAGGCATGCTTTAAATTTGAACATGTGATTACCGGTGAGCGATATGAACTGAAAGTTTCA
TCGGGTGAGTTTAAAGATTGTTGAAACAGTTCTGGTTGACACATGACATTACAATATTGGAACCTGATGAATTA

TAGTCATGCTCTGGGCAAATGAAATTTTCAACCGTGATGAGCCCGATACTATTGCGAGACCTCTATAATGAAAA
 ATAAAAAGTTATGTCTGTTGTTGGAACCCGTCCTGAGATTATCCGTTTGTGCGAGGGTCTTGCTAAGTTTGATGAA
 TACTGCGAGCATATTATTGTCCATACTGGTCAAAATTATGATTACGAATTAATGAAGTGTTCCTCAATGACTTGG
 GTGTTCCGAAAACCTGATTATTTTTTAAATGCAGCGGGTAAAAATGCGGCCGGAACCATTGGTCAGGTTATTATTAA
 5 GGATAGATGAAGTATTAGAAAATCGAAAACCTGAAGCAATACTGGTATTGGGCGATACGAATTCATGTATTTCTGC
 CATTCCGGCCAAACGCCGTAAAGTGCCTATATTTTCATATGGAAGCAGGTAACCGTTGTTTCGATCAACGCGTGCC
 TGAAGAAAACCAACAGACGTATTGTTGACCATACGGCTGATATCAATATGACCTACAGTGATATTGCTCGTGAATAT
 CTCTTGGCTGAAGGTATCCAGCTGATCGGATCATAAAAACTGGTAGCCCTATGTTTGAGGTTCTTTCATATTATA
 TGCCCCAAATTGATGGTTCAGATGTGCTATCGCGTTTGAATCTACAGTCTGGTGAGTTTTTTGTAGTAAGTGCGC
 10 ATCGTGAAGAGAATGTTGATTCTCCAAAACAGCTCGTAAAGCTTGCGAACATTCTAAATACTGTTGCTGAAAAATA
 TAATCTTCCAGTTATTGTCTCCACACACCCAAAGGACACGTAACCGAATCCGTGAGCAAGGAATTGAATTTCAATC
 AAATATAAATCTACTGAAACCATTGGGTTTCCATGATTATAACCACTTGCAGAAGAACTCACGAGCTGTGCTTTCA
 GATAGCGGTACTATCACTGAAGAGTCATCCATGAATTTCCAGCGGTAACATCCGGGAAGCGCATGAGCG
 TCCGGAAGGCTTTGAGGAAGCATCCGTATGTTGAGGTTAGAGTGTGAACGCTATTACAACGCGTGGATA
 15 TTCTGGCAACACAACCCGCGAGGTGAAGTCCGCTTTTTACGTCAGGTTAGTGATTACAGCATGCCAAATGTGTCCG
 GATAAAGTTGTCAGAAATTGTTCACTCTTACACAGATTATGTTAAGAGAGTCGCTCTGGAAAGAATATTGATGAACT
 TGCTTTAATCATAGATGATTACCTGCCAACAGTACTCGTGTGGTGCAAAAATGTTTCATGAACCTGCTCAAGAA
 TTTATCCAGCGTGGGCACGATGTTACGGTAATTACTCCTGGTACGGGCATGCAAGAAGAGATTTCTTTTGATACC
 TTTACAGGGGTAAAAACATGGCGTTTTAAAGCGGGCCGCTCAAGGATGTAAGTAAAATTCAGCGAGCGGTCAA
 20 TGAACCGCTTTTGTCCATCGGCGTGGAAAGCCATCAAAAATGGGTAAAAAAGAGACCTTTGAGGGGGTGA
 TTTATTATTACCTTCCATATTCTGGGGCCTTAGTTAAAAAATTAAGCTCGTTGCCAATGCTGTTATCTT
 ATTTTAAGAGATATGTTTCCACAATGGGTAATTGATGCAAGGAATGCTTAATGCTGTTCCCAATAGAACGCTACT
 TTCGCTTTTTGAAAAATATCTTATCGTCAGGCAATCGTATTGGACTTATGTCTGATAAGAATCTTGATGTTTTT
 CGGAAAGATAATAAAGGCTATCCGTGCGAAGTTTTGCGTAATTGGGCATCCCTAACACCAACGATCATACCCAAG
 25 GATTATATACCACTACGTAAGCGACTTGGCCTAGAGGATAAAACCATTTTCTTCTATGGTGGAAACATAGGTCAT
 GCACAGGACATGACAACTTGATGCGACTTGTGAGAAACATGGCAGCATATCCTCAAGCTCATTTCCTATTTATT
 GGCAGGGGGATGAAGTTGAATTAATTAATTCATTAGCATCTGAGTGGGCATTGACGAATTCACCTATTTGCC
 TTTTAACCAAGATGAATTTAAGTTCATTTTGTGCGAAATGGATATCGGCTGTTTTCTCTTTCCGCTAGACACT
 CTTCCATAATTTTCTGGTAAGTTATTAGGCTATATGGTTCAGTCGCTACCTATTTTAGGTAGCGTAAATGCCGG
 30 AAATGATTTGCTCGACATTGTCAATCAAAAATGCGGGATTAATCCATGTCAATGGTGAGGACGATAAATTATGT
 CAATCTGCGCTATTAATGTTGCATGATATTGATGTGCGCCGGCAACTTGGTTCGGGGGCGAATATATTGTTGAAA
 GAACAATTCTCCGTTGAGTCTGCGGCACAGACGATAGAAATGAGGTTGGAGGCATGCAATGCGATTAATTGATA
 ATGACCAACTCGACGAATTATATGATCAAGCCGGGCAATCGGAACGTTTACGTTCCACCTTATGATGCACGGCT
 CGCATCAAGAAAAGGTACAGCGTTTACTTATTGCATTAGTAAGGGCAGCTATGTTGAACCGCATTATCACGAAC
 35 TTCTCATCAGTGGGAAATGTTCAATTGTTATGGAGGGGCAACTCAGGTTTTGTTGTATGGTAGAAATGGTGAGG
 TTATAAAGCAATTTATAGCAGGAGATAATACTGGAAATGAGCATTGTGGAGTTTTCTCCGGGCGATACACAGTG
 TCGAATGCCTATCTCCGCGTGTCTTATGGTGGAAAGTTAAGGAGGGGCCATTTGACCCTTCTTTTGCAAAATCGT
 TCGTGTGAGCGGCCGCGAGCTCGTCTGACTCGAGGATCCGTGTAGGCTGGAGCTGCTTGAAGTTCCTATACTTT
 CTAGAGAATAGAACTTCGGAATAGGAACTAAGGAGGATATTCATATGGATAAAGCCGTAAGCATATAAGCATGG
 40 ATAAGCTATTTATACTTTAATAAGTACTTTGTATACTTATTTGCGAACATTCCAGGCCGCGAGCATTACGCGCGGT
 GATCACACCTGACAGGAGTATGTAATGTCCAAGCAACAGATCGGCGTAGTCCGGTATGGCAGTGATGGGACGCAA
 CCTTCCGCTCAACATCGAAAGCCGTGGTTATACCGTCTCTATTTCAACCGTTCCCGTGAGAAGACGGAAGAAGT
 GATTGCCGAAAAATCCAGGCAAGAACTGGTTCCTTACTATACGGTGAAAGAGTTTTGTCGAATCTCGAAACGCC
 TCGTTCGCATCCTGTTAATGGTGAAGCAGGTGCAGGCACGGATGCTGCTATTGATTCCCTCAAACCATATCTCGA
 45 TAAAGGAGACATCATCATTGATGGTGGTAACACCTTCTTCCAGGACACTATTCGTCGTAATCGTGAGCTTTCAGC
 AGAGGGCTTTAACTTCATCGGTACCGGTGTTTCTGCGGGTGAAGAGGGGGCGCTGAAAGGTCCTTCTATTATGC
 CTGGTGGCCAGAAAGAAGCCTATGAATTGGTAGCACCGATCCTGACCAAAATCGCCGCCGTAGCTGAAGACGG
 TGAACCATGCGTTACCTATATTGGTGCCGATGGCGCAGGTCACTATGTGAAGATGGTTCACAACGGTATTGAATA
 50 CGGCGATATGCAGCTGATTGCTGAAGCCTATTTCTGTCTAAAGGTGGCTGAACCTCACCAACGAAGAAGTGG
 CGCAGACCTTTACCGAGTGAATAACGGTGAACCTGAGCAGTTACCTGATCGACATACCAAAAGATATCTTCAACA
 AAAAAGATGAAGACGGTAACTACCTGGTTGATGTGATCCTGGATGAAGCGGCTAACAAAGGTACCGGTAATGG
 ACCAGCCAGAGCGCGCTGGATCTCGGCGAACCGCTGTCGCTGATTACCGAGTCTGTGTTTGCACGTTATATCTC
 TTCTCTGAAAGATCAGCGTGTGCCGCATCTAAAGTCTCTCTGTCGCAAGCACAGCCAGCAGGCGACAAGG
 CTGAGTTCATCGAAAAAGTTCGTCGTGCGCTGTATCTGGGCAAAATCGTTTCTTACGCCAGGGCTTCTCTCAGC
 55 TGCGTCTGCGTCTGAAGAGTACAACCTGGGATCTGAACTACGGCGAAATCGCGAAGATTTTCCGTGCTGGCTGC
 ATCATCCGTGCGCAGTTCCTGCAGAAAATCACCGATGCTTATGCCGAAAATCCACAGATCGCTAACCTGTTGCTG
 GCTCCGTACTCAAGCAAAATGCCGATGACTACAGCAGCGCTGACGTGATGCTGTTGCTTGCAGTACAGAA
 CGGTATTCCGGTTCGACCTTCTCCGCAGCGGTTGCCTATTACGACAGCTACCGTGTGCTGTTCTGCTGCGA
 60 ACCTGATCCAGGCACAGCGTACTATTTTGGTGGCATACTTATAAGCGTATTGATAAAGAAGGTGTGTTCCATA
 CCGAATGGCTGGATTAA

SEQ ID NO: 16 (Ejemplo de secuencia de nucleótidos de locus O16 *rfb* – O16–producción EPA cepa stLMTB11739)

ATGACGAATTTAAAAGCAGTTATTCCTGTAGCGGGTCTCGGGATGCATATGTTGCTGCCACTAAGGCGATACCC
AAAGAGATGCTACCAATCGTCGACAAGCCAATGATTCAGTACATTGTTGACGAGATTGTGGCTGCAGGGATCAAA
GAAATCCTCCTGGTAACTCACGCGTCCAAGAACGCGGTGAAAACCACTTCGACACCTCTTATGAGTTAGAATCA
CTCCTTGAGCAGCGCGTGAAGCGTCAACTGCTGGCGGAAGTACAGTCCATCTGTCCGCCGGGCGTGACCATTA
5 TGAACGTGCGTCAGGGCGAACCTTTAGGTTTAGGCCACTCCATTTTGTGTGCGCGACCTGCCATTGGTGACAAC
CCATTTGTCGTGGTACTGCCAGACGTTGTGATCGACGATGCCAGCGCCGACCCGCTACGTTACAACCTTGTCTGC
CATGATTGCACGTTTCAACGAAACGGCCGAGCCAGTGTGGCAAACGATGCCCAGGTTGACCTCTCTGAAT
ACTCCGTCATCCAGACTAAAGAGCCGCTGGACCGTGAGGGTAAAGTACGCCGATTGTTGAATTTATCGAAAAA
10 CCGGATCAGCCGACGCTGGACTCAGACATCATGGCCGTAGGTGCTATGTGCTTTCTGCCGATATTTGGCC
GGAACGTGAACGTAAGTACTCAGCCTGGTGCATGGGGACGTATTACGCTGACTGATGCTATTGCCGAGCTGGCGAAA
AAACAATCCGTTGATGCAATGCTGATGACCGGCGACAGTTACGACTGCGGCAAAAAAATGGGCTATATGCAGGC
GTTTGTGAAGTATGGCCTACGCAACCTGAAAAGAAGGGGCGAAGTTCGGTAAAGGTATTGAGAAGCTGTTAAGCG
AATAATGAAAATCTGACCGGATGTAACGGTTGATAAGAAAATAAACGGCAGTGAAAATTCGCAGCAAAAAGTAA
15 TTTGTTGGCAATCTTCCGCTGCTGTTTTATATAAACCATCAGATAACAACGAGTTAGGTTTTTATCA
AAGTTTTCCAGGATTTTCTTTGTTTTCCAGAGCGGATTGGTAAAGACAATTAGCGTTTTGAATTTTCCGGTTTTAGCGC
GAGTGGGTAACGCTCGTCACATCATAGGCATGCATGCAGTGTCTGGTAGCTGTAAAGCCAGGGGCGGTAGCG
TGCATTAATACCTCTATTAATCAAACCTGAGAGCCGCTATTTACAGCATGCTCTGAAGTAATATGGAATAAATTA
AGTGAANAATACTTGTACTGGTGGCGCAGGATTTATTGGTTCAGCTGTAGTTGCTCACATTATAAATAATACGCAG
20 GATAGTGTGTTAATGTGATAAATAACGTACGCCGAAAACCGGGAATCACTTGTGATGTTTCTGATTCTGAA
CGCTATGTTTTGAACATGCGGATATTTGCGATGCACCTGCAATGGCACGGATTTTTGCTCAGCATTCAGCCGGAT
GCAGTATGCACCTGGCTGAAAGCCATGTTGACCGTCAATTAACAGCCCTCGCGCATTTATGAAAACCAA
TATTGTTGGTACTTATGTCTTTTTGAAAGCCGCTCGCAATTAATGCTGCTCTTGATAGCGACAAGAAAAATAG
CTTCCGTTTTCATCATATTTCTACTGACGAAGTCTATGGTGAATTTGCCTCATCCAGATGAAGTAAATAATACAGAA
25 GAATTACCCTTATTTACTGAGACGACAGCTTACGCGCAAGCAGCCCTTATTCGCATCCAAAGCATCCAGCGAT
CATTTAGTCCGCGCGTGGAAACGTACATATGGTTTACCGACAATTGTGACTAATTGCTCGAACAACATATGGTCCT
TATCATTTCGCGAAAAGCTTATCCACTGGTTATCTTAATGCACTGGAAGGTAAGGCATTACCTATTTATGGCA
AAGGAGATCAGATCCGCGACTGGTTGTATGTTGAAGATCATGCGCGTGCGTTATATACCGTCGTAACCGAAGGT
AAAGCGGGTAAAACCTTAAACATTGGTGGGCACAACGAAAAGAAAAACATCGATGTAGTGCTCACTATTTGTGAT
20 TGTCTGGATGAGATTGTACCGAAAAGAGAAAATCTTATCGTGAGCAAATCACTTATGTTGCTGATCGTCCGGGACAC
GATCGCCGCTATGCTATTGATGCTGAGAAGATTGGTCCGCGCATTGGGATGGAACCACAGGAAACGTTTGAGAG
CGGGATTGTAACCGGTGGAATGGTACCTGTCCAATACAAAATGGGTTGATAATGTAAAAGTGGTGCCTATCA
ATCGTGGATTGAACAGAATATGAGGGCCGCCAGTAATGAATATCCTCCTTTTTGGCAAACAGGGCAGGTAGG
TTGGAACTACAGCGTGTCTGCGCACCTTTGGGTAATTTGATTGCTTTTGATGTTCACTCTACTGATTATTGCGGT
35 GATTTTAGTAATCCTGAAGGTGTAGCTGAAACCGTAAGAAGCATTCCGGCCGGATATTATTGTCAATGCAGCCGCT
CACACCGCAGTAGACAAGCAGAATCAGAACCGGAGTTGCACAATTAATTAACGCAACAAGTTCGAAGCGAT
TGCGAAAGCAGCAAAATGAAGTTGGAGCCTGGGTTATCCACTGACTGATTACGCTTCCCTGGAAAATGGCG
ATATGCCATGGCTGGAGACGGATGCAACCGCACCACTAAATGTTTACGGTGAACCAAGTTAGCCGGAGAAAAA
GCGTTACAGGAATATTGCGCGAAGCATCTTATTTCCGGACCAGCTGGGTCTATGCAGGAAAAGGAAATAACTTC
40 GCCAAAACGATGTTACGTCTGGCAAAAGAGCGTGAAGAATTAGCGGTTATTAACGATCAGTTTGGTGCGCCAAC
AGGTGCTGAACTGCTGGCTGATTGTACAGCACATGCCATTCGTGTGCGACTGAATAAACCGGATGTGCGAGGCT
TGTACCATTTGGTAGCCAGTGGTACCACAACCTGGTACGATTAATGCTGCGCTGGTTTTTGAAGAGGCGCGCAA
GCAGGCATTCCCCTTGCACTCAACAAGCTCAACGCAATCAACCAACAGCCTATCCTACACCGCTCGTCCGCTC
45 ACATAACTCTCGCTTAATACAGAAAAATTTAGCAGAACTTTGCGCTTGTCTTGCCTAGCTGGCAGGTTGGCGT
GAAACGAATGCTCAATGAATTTTACGACTACAGCAATTTAATAGTTTTTGCATCTTGTTCGTGATGGTGGAGCA
AGATGAATTAAGGAATGATGAAATGAAAATGCGTAAAGGTATTATTTAGCGGGTGGTTCTGGTACACGTCTTT
ATCCTGTGACTATGGCTGTCAGTAAACAGCTATTACCTATTTATGATAAACCGATGATCTATTACCCGCTCTCTAC
ACTGATGTTGGCGGGTATTCGCGATATTTGATTATCAGTACACCTCAGGATACTCCTCGTTTTCAACAATTGCTG
GGTGACGGTAGCCAGTGGGGCCTGAATCTTCAGTACAAAGTGCAACCTAGCCCAGATGGCCTCGCGCAGGCAT
50 TTATCATCGGTGAAGAGTTTATTGGTGGTATGATTGTGCTTTGGTTCTTGGTGATAATATCTTTACGGTACCGA
TCTGCCGAAGCTAATGGAGCCGCTGTTAACAAAGAAAAGTGGTGCAACGGTATTTGCCTATCACGTTAATGATCC
AGAACGCTATGGTGTGCTTGTGTTTTGATAAAAACGGTACGGCAATCAGTCTGGAAGAAAAACCGTTAGAACCAAA
GAGTAATTACGCCGTTACAGGTCTGTACTTTTATGATAACGACGTGGTTCAGATGGCGAAAAACTTGAAGCCGTC
TGCACGTGGTGAAGTGAAGATTACAGATATTAACCGTATTTATCTTGAGCAGGGACGTCTGTCTGTGCGGATGAT
GGGGCGTGGCTACGCGTGGCTGGACACGGGGACTCATCAGAGTCTGATAGAAGCAAGTAATTTTATTGCGACA
55 ATTGAAGAGCGCCAGGGATTGAAGGTTTCCCTGTCTGAAGAGATTGCATTTTCGTAAGGTTTTATTGATGTTGAG
CAAGTAAGAAATAGCTTACCCTAATAAAGAATAATTAGGGCAGTATCTTATAAAATGACGAAGGATTCAA
ATTAATGAATGTGATTAGAATGAAATGAAGATGTGCTAATTCTGGAGCCAAGAGTATTGGTGAATAGATAGGT
TTCTTTTATGAGAGCTTAAATCAATCAGCATTTGAACATATTCTAGGCTATCCGGTACGTTTTGTTCAAGACAATCA
CTCACGTTTCATCAAAAAATGTAAGTACTCAGAGGCTTCACTTTCAACGCGGCGAGTACGCACAAGATAAATCTGTACG
60 CTGCACTCATGGAGCAGTTTTTGTGTTGCTGTTGATATTGACCCAAATTCGGTATCCTTTGGTAAATGGGTTGG
TGTTCTGCTTTACGCTGATAATAAGCAGCAGTTGTGGATACAAAAGGGTTTTGCTCATGGCTTTTTGGTTCTGTCT
GATATCGCTGAATTTCAATATAAAACTACAAACTATTATCATCTGAAAGCGATTGTGGAATATGTTGGAATGATG
AACGCATTGCAATTTGATTGGCCCCAAACATCAGGGTTAATCCTTTCCGCAAAAAGATGAAAGGCTCTTTACGTTAG
ATGAGCTTATCAGATTAATAAATGATGAATAACGAATAAATATCTTTAAGAAGAAAACGTTATATATCTGGCTG

TCGTTCAAGGTAGCAATTATCTTTTACCATTGCTTACATTTCCATATCTTGTGAAGAACACTTGGTCCTGAAAAATTC
GGTATATTCGGTTTTTGGCAAGCGACTATGCTATATATGATAATGTTTGTGAATATGGTTTCAATCTCACAGCAA
CTCAGAGTATTGCCAAAGCAGCAGATAGTAAAGATAAAGTAACGTCTATTTTTGGGCGGTGATATTTTCAAAAAT
AGTTCTTATCGTCAATTACATTGATTTTCTTAACGTCGATGACCTTGGCTTGTCTGAATATAACAAGCATGCCGTAA
5 TTATATGGTCGTTTTGTTCTGCATTAGTCGGGAATTTAATCTACCCTATCTGGCTGTTTCAGGGAAAAGAAAAAAT
GAAATGGCTGACTTTAAGTAGTATTTTATCCCGCTTGGCTATTATCCCTCTAACATTTATTTTTGTGAACACAAAGT
CAGATATAGCAATTGCCGGTTTTATTACAGTCAAGTGCAAATCTGGTTGCTGGAATTATTGCACTAGCTATCGTTGT
TCATGAAGGTTGGATTGGTAAAGTTACGCTATCATTACATAATGTGCGTCGATCTTTAGCAGACGGTTTTCATGTT
TTTTTTCCACATCTGCTATTAGTTTATATTCTACGGGAATAGTTATTATCCTGGGATTTATATCTGGACCAACGTC
10 CGTAGGGAATTTAATGCGGCCAATACTATAAGAAACGCGCTTCAAGGGCTATTAATCCTATCACCCAAGCAAT
ATACCCAAGAATATCAAGTACGCTTGTCTTAATCGTGTGAAGGGTGTGATTTTAAATAAAAAATCATTGACCTGC
TTGAGTTTGATTGGTGGTGCCTTTTTCATTAATTCTGCTCTTGGGTGCATCTATACTAGTAAAAATAAGTATAGGGC
CGGGATAGATAATGCAGTGATTGTGCTAATGATTATATCGCTGCTGCCTTTTCTTATTTTCATTAAGTAAATGTCTAT
GGCATTCAAGTCTGACCATAAATTATAAGAAAGTTCAGTAAAGATTTTAAATCGCTCGGGTTTGTGAGTT
15 TGTGTTGATTTTTCCGCTAACCAACTTTTTTAAAGAGATTGGTGCAGCAATAACATTGCTTGCAACAGAGTGCTT
AGTTACGTCACCTCATGCTGATGTTCTGAAGAAATAATAAATTACTGGTTGCTGAGGATTTTATGTACGATTATATC
ATTGTTGGTTCTGGTTGTTTGGTGCCTTTGTGCGAATGAGTTAAAAAAGCTAAACAAAAAAGTTTTAGTGATTG
AGAAAAGAAATCATATCGGTGGAATGCGTACACAGAGGACTGTGAGGGTATCCAGATTCATAAATATGGTGCAC
ATATTTTTCATACCAATGATAAATATATATGGGATTACGTTAATGATTTAGTAGAATTTAATCGTTTTACTAATTCTC
20 CACTGGCGATTTATAAGACAAATTTCAACCTTCTTTTTAATATGAATACTTTCCACCAAATGTGGGGAGTTAAA
GATCCTCAAGAAGCTCAAAATATCATTAAATGCTCAGAAAAAAGATATGGTGACAAGTCACTGAAAATTTGGAG
GAGCAGGCGATTTTCATTAGTTGGGGAGGACTTATACCAAGCATTGATAAAGGGTTATACGGAGAAGCATTGGGG
AAGAAGTGCAAAAGAATTGCCTGCATTTATTTAAGCGAATCCCAGTGAGATTTACGTTTGATAACAATTTTTTT
CCGATCGCTATCAAGGTATTCCGGTGGGAGGCTACACTAAGCTTATTGAAAAAATGCTTGAAGGTGTGGACGTA
25 AAATTAGGCATTGATTTTTGAAAGACAAAGATTCTCTAGCGAGTAAAGCCCATAGAATCATCTACACTGGACCCA
TTGATCAGTACTTCGACTATAGTTTTGGAGCGTTAGAATATCGCTCTTTAAAATTTGAGACGGAACGCCATGAATT
TCCAAACTTCCAAGGGAATGCAGTAATAAATTTCACTGATGCTAATGTACCATATACCAGAATAATTGAGCATAAA
CATTTTGACTATGTTGAGACAAAGCATACGGTTGTTACAAAAAGATATCCATTAGAGTGAAAAGTTGGCGACGAA
CCCTACTATCCAGTTAATGATAATAAAAAACATGGAGCTTTTTAAGAAATATAGAGAGTTAGCTAGCAGAGAAGACA
30 AGGTTATATTTGGCGGGCGTTTGGCCGAGTATAAATATTATGATATGCATCAAGTGATATCTGCCGCTCTTTATCA
AGTGAAAAATATAATGAGTACGGATTAATGATCTATCTTGAATTAGTGTCTTTCTCATTACAGCATTATCTGTTT
ATATCTTAAGAAGGATATATTTTATCCAGCCGTATGCGTTAATATCATCTTCGCACTGGTCTTATTGGGATATGAA
ATAACGTCAAGATATATATGCTTTTTAGTTAAATGACGCTACGTTGATTTTTCTACTTTGCAATGTTTTGACATTTAC
35 CCTGTATGTTTTATTGACGGAAAGTGATTTAGATCTAAATATCAGAAAAGTCAATAATGCTATTTATAGCATACCAT
CAGTTCCGGACTTACTTAGCTATATGAATTTGATAAGAGATGCTGATGTTGAAGACACATCAAGAAATTTCT
CAGCATAACATGCAGCCAATCATTCTAACTACTTTTTGCTTTATTTATTTGGTCTAAAAAATTTACTAATACAAAGGTA
AGTAAAACATTTACTTTACTTTGTTTTATTGTATTCTATCTTTGCAATTACTGAATACTGGTAAGCAAATTTGCTTTT
ATGGTTATCATCTCTTATGCATTCATCGTAGGTGTTAATAGAGTAAAACATTATGTTTATCTTATTACAGCTGTAGG
40 TGTTCTATTCTCCTTGTATATGCTCTTTTTACGTGGACTGCCTGGGGGGATGGCATATTATCTATCCATGATTTG
GTCAGCCATAAATCGCGTTTCAGGAGTTTTATTTTCAGCAAGTATCTAACTCTGCCAGTTCTCATGTCTTTTTGGT
TTTTGAAAAGCTGATGGGGCTATTAACAGGTGGAGTCTTATGCTGTTGCATAAAGAATTTGTGGTGGTGGTT
TGCCAACAAATGTTTATACTGCTTTTTCGGATTATGTTTTATATTTCCGCGAGCTAAGTATTTGATGATGTTTATT
CATGGCTGATTTTCAGGTGTTTTATGGAGATTGTCTCGAAATTACATATCTGTGAAAATATTTTATTCATATTTTTATT
45 TATACCTTTTCTTTCAATTTTATCATGAAAGCTTCATGACTAATATTAGCAGTTGGATACAAATAACTCTTTGTATC
ATAGTATTCTCTCAATTTCTTAAGGCCAGAAAAATAAGTGAATGTTTTTTGAATGATTTAAATTTCTCTAGA
CGCGATGCTGGATTTAAAGCAAGAAAAGATGCACTGGACATTGCTTCAGATTATGAAAACATTTCTGTTGTTAACA
TTCTCTATGGGGTGGAGTAGTCCAGAGAATTATTAGTTCTGTTAAGCTTAGTACATTTCTCTGCGGTCTTGAAAA
50 TAAAGATGTTTTAATTTCAATTTCCCGATGGCCAAACCTTTGGCATATATTGTCATTTCCACCGCCTTCTAA
AATTTAGAATAGTACCTCTGATTCATGATATTGATGAATTAAGAGGAGGAGGGGGTAGTGATTCTGTGCGGCTTG
CTACCTGTGATATGGTCATAAGTCACAATCCACAAATGACAAAAGTACCTTAGTAAATATATGTCTCAGGATAAAAT
CAAAGACATAAAAAATTTGATTACCTCGTCTCATCTGATGTGGAGCATCGAGATGTTACGGATAAGCAACGAGG
GGTCATATATGCTGGCAACCTTTCTAGGCATAAATGTTCTTTTCATATATACTGAAGGATGCGATTTTACTCTCTTT
GGTGTCAACTATGAAAATAAAGATAATCCTAAATATCTTGGAAAGTTTTGATGCTCAATCTCCGAAAAAGATTAACC
55 TCCCAGGCATGCAATTTGGACTCATTGGGATGGAGATTCTGTGCAAACTGTAGTGGTGCCTTTGGCGACTATT
TAAAGTTAATAACCCTCATAAGACATCTTTATCTTTCAATGGAACCTCCAGTATTTATGGGATAAAGCCGCGC
CTTGGCGATTTTATTGATAAATAGATAATAGGATGCAATGGGATCAATCAAGAAATGCAAGAGATTTGTTGCC
TCCATGACAATAGAAAATTTAAGCAAATTTAGTGAGAATACAAAAATTTTCTCAGAAAATTCGAACAGGAAGTT
ACTTCAGGGATGTTCTTGAAGAGGTGATCGATGATCTTAAAACTCGCTAACCGATATGGTCTCTGTGGTTTTATTC
60 GGCTTGTAGAGATGTCTTATTGACTCGTGTATTTTACCGGAAGTGTAGAATTATTGATTTCCCTGCTATATTCCG
CAATGATGGTAGCATTAAATTTGGTGAATAATTTACAAGTGGAGTCCGGTCTCAGGCTGGATGCATTTGGACGTGG
CGTGATTTTTTTTTCCGATAATGTGCAAGTTAACGACTATGTTTCATATCGCCTCAATTGAGAGCGTTACGATAGGT
CGGGATACGCTTATTGCAAGTAAAGTATTTTACCAGATCATAATCACGGTTCCTTTAAAGCACTCTGATCCAATGA
GTTCCGCAAAATATACCTCCAGACATGCGCACGTTGGAATCTTCAGCTGTTGTAATTGGCCAGAGGGTTTGGTTG

GGTGAGAAATGTGACGGTTTTGCTGGAACAATTATTGGTAATGGAGTCGTAGTCGGCGCCAATTCTGTTGTTAGA
GGTTCTATTCCCAGAAAATACTGTCATTGCGGGAGTACCAGCAAAAATCATAAAGAAATACAATCATGAGACCAAA
TTATGGGAAAAAGCATAGTCGTTGTTTCTGCGGTCAATTTTACCCTGCGGTCCATTTACCATTTTGA AAAAATT
5 TTTGGCAGCAACTAATAATAAGAAAAATGTCAGTTTTATCGCATTAGTCCATTCTGCTAAAGAGTTAAAAAGAAAGT
TATCCATGGGTTAAATTCATTGAGTTTCTGAGGTTAAAGGGTCGTGGCTAAAACGTTTGCACCTTTGAATATGTAG
TTTGTA AAAA ACTTTCAAAGAGCTGAATGCTACGCATTGGATTTGTCTGCATGATATTACGGCCAATGTCGTCAC
TAAAAAAGATATGTGATTGTCATAACCTGCCCCTTTTTATAAAGGAATTTTATCCGTAATTTAGCATTATGGAGC
CTAGCTTTTTCTTATTTAAAATGCTATACGGGCTGATATATAAAAATAAACATTA AAAA AATACTGCAGTGTGGTT
10 CAACAATTCTGGATGAAAAGAAAAATTTATCAAGAAATATTCTATAAAAATAACATCATTGTCAGTCGGCCAGAAAATA
ATTATCTGATAAAAAGCCAACCTACTGATGATGATTCTCAATTTAAGAATAACCTTCTGAGTTGACAATATTTTACC
CTGCTGTTCCACGAGTATTTAAAAATTACGAGCTTATTATTAGTGCAGCAAGGAAATGAAAAGAACAAATCCAATAT
TAAATTTCTGCTTACTATCAGTGGTACAGAAAAATGCGTATGCAAAAATATATTATCAGTCTTGCAGAAGGACTGGAT
AATGTTCAATTTCTCGGTACTTGGATAAAGAAAAATCGATCATTGTTATAATATTTAGATATAGTTTGTTTTCC
CTCTAGTTAGAAAACATGGGATTGCCGTTGCTGAGGCTAAAGAGCGAGGTAAGTGGTATTAGCATCAGATT
15 TCCCATTTACTAGAGAAAACCTTGGTAGTTATGAAAAGAAAGCTTTTTTTGATTCTAATAACGATGACATGTTAGTT
AACTTATTATTGACTTCAAAAAAGGTAACCTCAAAAAAGATATCTCTGATGCAAATTTCAATTTATCGTAATGAAAA
TGTATTAGTTGGGTTTGTGAACTAGTTAATTTTATTACTGAAGAACATTGAAATGGTATATATAAATAATCGTTTCC
CACGGACATGAAGACTACATCAAAAAATTACTCGAAAATCTTAATGCTGACGATGAGCACTACAAGATTATCGTAC
GCGACAACAAAGACTCTCTATTATTGAAACAATATGCCAGCATTATGCAGGCCTGGACTATATTAGTGGAGGTG
20 TATACGGCTTTGGTCATAATAAATAATTTGCGGTGGCGTATGTAAGGAAAAATATAGACCCGCGATGATGATTA
CATTTTGTTTTTAGATCCCGATATCATATGAAGCATGATGATTGCTGACATATATAAATATGCGAAAGTAAGC
GTTATGCTTTTAGTACATTATGCCTGTTCCGAGATGAAGCGAAATCTTTACATGATTATTTCCGTAAGAAAAATTTCT
GTGCTTTCTGATTTTATTGTGCTATTTATGTTAGGGATTAATAAAAACAAAAATTCCTAAAGAAAGTATCTATTCTGA
TACGGTTGTTGATTGGTGCAGGATCATTATGCTGGTACGTTTTTTCAGATTTTGTGCGTGTAAATGGCTTCGAT
25 CAAGTTACTTTATGTACTGTGAAGATATTGACCTGTGCTTGGAGCTTAGCCTGGCTGGTGTGACACTTCATTAT
GTTCCCGCTTTTTCATGCGATACATTATGCTCATCATGACAATCGAAGTTTTTTTTCAAAGCCTTCAGATGGCACT
TAAAAGTACTTTTAGATATTTAGCCAGAAAACGATTTTTATCAAATCGCAACTTTGATCGAATTTTCATCAGTTTTT
CACCCGTAAGAGCTCGGTACCCGGCCTAGGGTGTAGGCTGGAGCTGCTTCAAGTTCTTACTATCTTTCTAGAGA
ATAGGAACCTCGGAATAGGAACTAAGGAGGATATTCATATCCGTCGACGGCGCCCGCTGCAGGCATGCAAG
30 CTTGATCCATATGGATCGCTAGCTTAATTAATAAAGCCGTAAGCATATAAGCATGGATAAGCTATTTATACTTTAA
TAAGTACTTTGTATACTTATTTGCGAACATTCCAGGCCGCGAGCATTACGCGCGGTGATCACACCTGACAGGAGT
ATGTAATGTCCAAGCAACAGATCGGCGTAGTCGGTATGGCAGTGATGGGACGCAACCTTGCCTCAACATCGAA
AGCCGTGTTTATACCGTCTCTATTTTCAACCGTTCCCGTGAGAAGACGGAAGAGTATTGCCGAAAATCCAGG
CAAGAACTGGTTCCCTTACTATACGGTGAAGAGTTTGTGCAATCTCTGGAACGCCTCGTCGCATCCTGTTAAT
35 GGTGAAAGCAGGTGCAGGCACGGATGCTGCTATTGATCCCTCAAACCATATCTCGATAAAGGAGACATCATCA
TTGATGGTGAACACCTTCTTCCAGGACACTATTGCTGTAATCGTGAGCTTTACAGCAGAGGGCTTTAACTTCA
TCGGTACGGGTGTTTCTGGCGGTGAAGAGGGGGCGCTGAAAAGGTCCTTCTATTATGCCTGGTGGCCAGAAAAG
AGCCTATGAATTGGTAGCACCGATCCTGACCAAAAATCGCCGCGTAGCTGAAGACGGTGAACCATGCGTTACCT
ATATTGGTGGCGATGGCGCAGGTCATATGTGAAGATGGTTTACAACGGTATTGAATACGGCGATATGCAGCTG
40 ATTGCTGAAGCCTATTCTCTGCTTAAAGGTGGCCTGAACCTACCAACGAAGAACTGGCGCAGACCTTTACCGA
GTGGAATAACGGTGAAGTGAAGTACCTGATCGACATACCAAAAGATATCTTACCAAAAAAGATGAAGACGG
TAACCTGCTGTTGATGTATCCTGGATGAAGCGGTAACAAAGTACGGGTAATGACCAGCCAGAGCGCG
CTGGATCTCGGCGAAGCCTGTCGCTGATTACCGAGTCTGTGTTTTGCACGTTATATCTTTCTCTGAAAGATCAG
45 CGTGTGGCCGATCTAAAGTTCTCTCTGTTCCGCAAGCACAGCCAGCAGGCGACAAGGCTGAGTTTCATCGAAAA
AGTTGCTGCTGCGCTGTATCTGGGCAAAAATCGTTTCTTACGCCAGGGCTTCTCTCAGCTGCGTGTGCTGCTG
AAGAGTACAACCTGGGATCTGAAGTACGGCGAAAATCGCGAAGATTTTCCGTGCTGGCTGCATCATCCGTGCGCAG
TTCTGCAAAAATCACCGATGCTTATGCCGAAAATCCACAGATCGCTAACCTGTTGCTGGCTCCGTACTTCAAG
CAAATTGCCGATGACTACCAGCAGGCGCTGCGTGTGCTTATGCAGTACAGAACGGTATTCCGGTTCC
50 GACCTTCTCCGCAGCGGTTGCCTATTACGACAGCTACCGTGTGCTGTTCTGCCTGCGAACCTGATCCAGGCAC
AGCGTGACTATTTTGGTGCCGATACTTATAAGCGTATTGATAAAGAAGGTGTGTTCCATACCGAATGGCTGGATT
AA

SEQ ID NO: 17 (Ejemplo de secuencia de nucleótidos de locus O18A *rfb* – O18A–producción EPA cepa BVEC–L–00559)

ATGACGAATTTAAAAGCAGTTATTCCTGTAGCGGGTCTCGGGATGCATATGTTGCTGCCACTAAGGCGATACCC
55 AAAGAGATGCTACCAATCGTCGACAAGCCAATGATTCAGTACATTGTTGACGAGATTGTGGCTGCAGGGATCAAA
GAAATCCTCCTGGTAACTACGCGTCCAAGAAGCGGGTCGAAAACCACTTCGACACCTCTTATGAGTTAGAATCA
CTCCTTGAGCAGCGCGTGAAGCGTCAACTGCTGGCGGAAGTACAGTCCATCTGTCCGCCGGGCGTGACCATTA
TGAACGTGCGTCAGGGCGAACCTTTAGGTTTAGGCCACTCCATTTTGTGTGCGCGACCTGCCATTGGTGACAAC
CCATTTGCTGTTGATGCTGCCAGACGTTGTGATCGACGATGCCAGTCCGCGCCGACCCGCTACGTTACAACCTTGTCTG
60 CATATTGACAGTTTCAACGAAACGGGGCGAGCAGGTGCTAGGCGAAAACGATATGCCGGGTGACCTCTGTAAT
ACTCCGTATCCAGACTAAAGAGCCGCTGACCCTGAGGGTAAAGTACGCCGATTTGTTGAATTTATCGAAAA
CCGGATCAGCCGACGCTGGACTCAGACATCATGGCCGTAGGTCGCTATGTGCTTTCTGCCGATATTTGGCC

GGAACTGGAACGTA CT CAGCCTGGTGCATGGGGACGTATT CAGCTGACTGATGCTATTGCCGAGCTGGCGAAA
AAACAATCCGTTGATGCAATGCTGATGACCGGCGACAGTTACGACTGCGGCAAAAAAATGGGCTATATGCAGGC
GTTTGTGAAGTATGGCCTACGCAACCTGAAAAGAAGGGGCGAAGTCCGTAAGGTATTGAGAAGCTGTTAAGCG
AATAATGAAAATCTGACCGGATGTAACGGTTGATAAGAAAATTATAACGGCAGTGAAAATTCGCAGCAAAAGTAA
5 TTTGTTGCGAATCTTCCTGCCGTTGTTTTATATAAACCATCAGAATAACAACGAGTTAGCAGTAGGGTTTTATTCA
AAGTTTTCCAGGATTTTCCTTGTTCAGAGCGGATTGGTAAGACAATTAGCGTTTGAATTTTTCCGGTTTTAGCGC
GAGTGGGTAACGCTCGTACATCATAGGCATGCATGCAGTGCCTCTGGTAGCTGTAAGCCAGGGGCGGTAGCG
TGCATTAATACCTCTATTAATCAAACCTGAGAGCCGCTATTTTACAGCATGCTCTGAAGTAATATGGAATAAATTA
AGTGAAAATACTTGTTACTGGTGGCGCAGGATTTATTGGTTCAGCTGTAGTTCGTCACATTATAAATAATACGCAG
10 GATAGTGTGTTAATGTGCGATAAATTAACGTACGCCGGAAACCGGGAATCACTTGCTGATGTTTCTGATTCTGAA
CGCTATGTTTTTGAACATGCGGATATTTGCGATGCACCTGCAATGGCACGGATTTTTGCTCAGCATCAGCCGGAT
GCAGTGATGCACCTGGCTGCTGAAAGCCATGTTGACCGTTCAATTACAGGCCCTGCGGCATTTATTGAAACCAA
TATTTGGTACTTATGCTCTTTTGAAGCCGCTCGCAATTACTGGTCTGCTCTGTATAGCGACAAGAAAAATAG
CTTCCGTTTTACTCATATTTTACTGACGAAGCTATGGTGAATTTGCTCATCCAGATAAGTAAATAACAGAA
15 GAATTACCCTTATTTACTGAGACGACAGCTTACGCGCAAAGCAGCCCTTATTCGCGATCCAAAGCATCCAGCGAT
CATTTAGTCCGCGCGTGGAAACGTACATATGGTTTACCGACAATTGTGACTAATTGCTCGAACAACTATGGTCCT
TATCATTTCGCGAAAAGCTTATCCACTGGTTATTCTTAATGCACTGGAAGGTAAGGCATTACCTATTTATGGCA
AAGGAGATCAGATCCGCGACTGGTTGTATGTTGAAGATCATGCGCGTGGCTTATATACCGTCGTAACCGAAGGT
AAAGCGGGTGAACCTTAAACATTGGTGGGCACAACGAAAAGAAAAACATCGATGTAGTGCTCACTATTTGTGAT
20 TTGCTGGATGAGATTGTACCGAAAGAGAAATCTTATCGTGAGCAAATCACTTATGTTGCTGATCGCTCCGGACAC
GATCGCCGCTATGCTATTGATGCTGAGAAGATTGGTGGCGCATGGGATGGAAACACAGGAAACCTTTGAGAG
CGGGATTCGTAAAACGGTGAATGTTACCTGTCCAATACAAAATGGTTGATAATGTGAAAAGTGGTGCCTATCA
ATCGTGGATTGAACAGA ACTATGAGGGCCGCCAGTAATGAATATCCTCCTTTTTGGCAAAACAGGGCAGGTAGG
TTGGGAACTACAGCGTCTGCTGGCACCTTTGGGTAATTTGATTGCTTTTTGATGTTCACTCTACTGATTATTGCGGT
25 GATTTTAGTAATCCTGAAGGTGTAGCTGAAACCGTAAGAAGCATTCCGGCCGGATATTATTGTCAATGCAGCCGCT
CACACCGCAGTAGACAAAGCAGAATCAGAACC GGAGTTTGACAATTAATTAACGCAACAAGTGTGGAAGCGAT
TGCGAAAGCAGCAAATGAAGTTGGAGCCTGGGTTATCCACTACTCGACTGATTACGCTTCCCTGGAAATGGCG
ATATGCCATGGCTGGAGACGGATGCAACCGCACCACTAAATGTTTACGGTGAACCAAGTTAGCCGGAGAAAAA
CGGTTACAGGAATATTGCGCGAAGCATCTTATTTTCCGGACCAGCTGGGCTATGCAGGAAAAGGAAATAACTTC
30 GCCAAAACGATGTTACGCTCTGGCAAAAGAGCGTGAAGAATTAGCGGTTATTAACGATCAGTTTGGTGCGCCAAC
AGGTGCTGAACCTGCTGGCTGATTGTACAGCACATGCCATTCGTGTGCACTGAATAAACCGGATGTGCGAGGCT
TGTACCATTTGGTAGCCAGTGGTACCACAACCTGGTACGATTATGCTGCGCTGGTTTTTGAAGAGGCGCGCAA
GCAGGCATTTCCCTTGCCTCAACAAGCTCAACGCAGTACCAACAACAGCCTATCCTACACCAGCTCGTCTGCC
ACATAACTCTCGCTTAATACAGAAAAATTTACAGCAGA ACTTTGCGCTTGTCTTGCCTGACTGGCAGGTTGGCGT
35 GAAACGAATGCTCAATGAATTTTACGACTACAGCAATTAATAGTTTTTGCATCTTGTCTGATGGTGGAGCA
AGATGAATAAAAGGAATGATGAAATGAAAATGCGTAAAGGTATTATTTTAGCGGGTGGTTCTGGTACAGCTTTT
ATCCTGTGACTATGGCTGTCAGTAAACAGCTATTACCTATTTATGATAAACCGATGATCTATTACCCGCTCTCTAC
ACTGATGTTGGCGGGTATTCGCGATATTTGATTATCAGTACACCTCAGGATACTCCTCGTTTTCAACAATTGCTG
GGTGACGGTAGCCAGTGGGGCCTGAATCTTCAGTACAAAGTGAACCTAGCCCAGATGGCCTCGCGCAGGCAT
40 TTATCATCGGTGAAGAGTTTATTGGTGGTGTGATTGTGCTTTGGTTCTTGGTGATAATATCTTTTACGGTACCGA
TCTGCCGAAGCTAATGGAGGCCGCTGTTAACAAAGAAAGTGGTGCAACGGTATTTGCCTATCACGTTAATGATCC
AGAACGCTATGGTGTCTGTTGAGTTTGATAAAAACGGTACGGCAATCAGTCTGGAAGAAAACCGTTAGAACCAA
GAGTAAATTACGCGTTACAGGCTGTACTTTTTATGATAACGACGCTGGTTTACAGATGGCAAAAACCTGGAAGCCGTC
TGCACGTGGT GAGTTAGAAATTACAGATATTAACCGTATTTATCTTGAGCAGGGACGCTCTGTCTGTGCGGATGAT
45 GGGGCGTGGCTACGCGTGGCTGGACACGGGACTCATCAGAGTCTGATAGAAGCAAGTAATTTTTATTGCGACA
ATTGAAGAGCGCCAGGGATTGAAGTTTCTGTCTGAAGAGATTGCATTTGTAAGGTTTTATTGATGTTGAG
CAAGTAAGAAAATTAGCTGTACCACTAATAAAGAATAATTATGGGCAGTATCTTTATAAAAATGACGAAGGATTCAA
ATTAATGAATGTGATTAGA ACTGAAATTTGAAGATGTGCTAATCTGGAGCCAAGAGTATTTGGTGTGATAGAGGT
TTCTTTTATGAGAGCTTTAATCAATCAGCATTTGAACATATTTAGGCTATCCGGTCACTTTGTTCAAGACAATCA
50 CTCACGTTTCATCAAAAAATGACTCAGAGGCCTTCACTTTCAACGCGGCGAGTACGCACAAGATAAACTTGTACG
CTGCACTCATGGAGCAGTTTTTGTGTTGCTGTTGATATTCGACCCAATTCGGTATCCTTTGGTAAATGGGTTGG
TGTTCTGCTTTACGCTGATAATAAGCAGCAGTTGTGGATACCAAAAAGGGTTTGTCTATGGCTTTTTGGTTCTGTCT
GATATCGCTGAATTTCAATATAAAA ACTACAACTATTATCATCCTGAAAGCGATTGTGGAATATGTTGGAATGATG
AACGCATTGCAATTGATTGGCCCCAAACATCAGGGTTAATCCTTTCCGCAAAAAGATGAAAAGCTCTTTACGTTAG
55 ATGAGCTTATCAGATTAATAAATTGATGAGGCCGGCCTTAAGGAGGACTAGTCCCAGCGCGCCATGAGTTT
AATCAAAAACAGTTTTTGAACCTTTGCGGGTATGTACTTCCAGCTATTGTGACACTACCAGCTTTGGGATTATG
GGCCAAAATTAGGCCGAGAATTTTGGTGTATTTTGGCATTAGCTGTTGTGGGTTATGCAAGCAACTTTTT
GATGCAGGCCTTACTCGCGCAGTGATACGAGAAGTTCGCAATTGAAAAGATAATGAAGAAAATAAGTTGAAAATT
ATTTCTCAGCGACAGTTGTAATTTATTTTGGTGGCCGCTCACTCTTATTTTTTTTTTATTGTTGTCATATC
60 GCATTGCTACTGAACATTAGTGAGACTTTTTTTCATAATGTAAGTGTCTCGCTTAAAATTTCTCGCAGCATCCATAC
CATTATTTTTGATTACTCAAATATGGTTGTCAATTTTAGAAGGTGAAGAAAGATTTGGTTTACTTAATATCTACAAT
CAATTACGGGAGTGATATTAGCAATCTCACCGGCATTATTTATACTTATTAACCCTCTTTGATGTATGCGATAATA
GGCTTAGTTCTAGCAAGGTTTTTATGTTTTATTTTGGCTTTTATAATTTGTACAGATAAAGTGTAAAAGCTAACT
AACAAATCGATATACCAACAATTA AAAAGATTGTTTATGTTCCGGTGGTTGGATTACAGTAAGTAATATCATCAGCCCT

GTGCTATCATATTTTGATAGGTTTATTGTTTCAAATCAACTTGGGGCTGCTAATGTTGCTTTTTATACTGCACCATC
AGAAATTATTTCTCGGCTTAGTATAATTCCAGGTGCGTTTTCAAGAGCCTTATTTCCAAGATTAGCTAATGCAAAT
AATCCGCTGAAAGATATAAAACGAAAAGATTAATTACAATTCACTTTTAATAATCATCACCCCTATTTTTGTATT
GGCGTGTATTTTCAGAGAAGATAATGGTTTTATGGATGGGGGCATCATTTTTTGGTGAGCCTGGTTTGGTATTAT
5 CAATATACTGATTGGCTTTATTTTTAATGGATTGGCACAAGTACCATTTGCCAGTATCAATCCCGAGGTGATGC
TAAGATAACTGCATTTGTTTCATCTCTTAGAGTTGTTTCCTTATTTATTACTTTTTATTTACCTCATAAAAAGCACATGG
GGTTGTTGGCGGGTATTGCGTGGTCAGTGAGGATGATAGTAGATTATATAGCATTAAAGTCTTTTGGACGGTAA
GTATATTAATAAAATAAAATTCAAAATGCAAGTTAATAACTCATGGCTTTATTTGGGTAGGTGACAATTTATAATGAT
ATATATATTAACCTTAACTCTTCTTCTAGTTATAGCCATAATGTTTTCTTCTCGGCACAAAAGTAGGATCACAT
10 CTCCATTACCTTTGCATTTTTTACCATGGTTACTAACTTTAATTGTCGGGATAAGTAATTACGATCAATTTTACGAG
TTAATGAAAGAAGCTTTTACTCTTTGTTGATTTGGTTACAGTTATTTTTATTTTTATTTTCATAGGGGAAGTGGTT
AATTATAAACGTGAAAATATAAATGTTTATTATGGTCTTTTACATATTAATATGAATGTAAAAAATATTGGATCATT
GTCATCCCAATTTTATTATATACCATTTTCAAATATATAGTTTGGTATGGGGGGAGCAGATGGATTCTTTCTCA
ATTTACGCTGTCAAATACATGGAGGGCTATCGGGTAAAAATTTATCTTAATGCTGCTGATGCTGCTTAAT
15 GATGGCTATGTTTCGCAATTGTTTGTCTAACAAAACTTCAAATTAATAAATAACTCCATTTTATTTCTGGATGTTTT
TGTATTGATTGGCACAATGGGAAAAATTTCAATATTAACGCCAATATTGACATATTTAATTATTTATGACTTCAAA
CATAGATTAAGAAGTAAAAAAAACAATAAAGTTTACATTGTTGATAATTATATTAGCTTTAACTTTGCATTTTACACGT
ATGGCTGAGAATGACCACTCAACATTTTTATCTATTTTAGGGCTCTATATTTATTCACCAATAATTGCTTTAGGCCA
GTTGAATGAAGTAAATAGTAGTCATTTTGGTGAGTATACGTTTATAGATTCATATGCTATAACTAATAAAATTTGGCC
20 TTATTAAGAATTGCCAGTAAATACTATTCTTGACTATTCATACGTTTCTGTACCAACAATGTATATACTGCACCT
CAACCATTTTACCAGGATTTTGGTTATACTGCATCATATTTGGAGCAGTATTATACGGACTAATATATGTTGAT
TATACACGGCCGGTGTTCGTGGAATAATACACAGGCATTACTGATTTACGCATTGTTTTAGTTAGTAGCAGTGCAA
CGGCTTTTCTCGCTGAAACGCTAGTAACGAATTTAGCTGGAATGTGATGTTAGTATTATGTACCATCTTACTATG
GCGATTTACAGTAATATGCAAACAGTACAGTAACCATTCTAATGGCCACCTACAATGGCGAGGCCTTCATCAAA
25 AATCAGATTTTGTCACTACAACAACAACATTTTCTAACTGGCGGTTATTTATTCAGGATGATGGGTCTACAGACA
ATACTATATCTATAATAAAAACTTCCAAAAATCTGACTCCAGAATTCGGCTAGTTGATGATAATTTGAAAGGTCAA
GGTGCAGGAAAAATTTTTATCGCTGATAAAGTACAGCGAGACAGATTATACAATTTATTGTGACCAAGATGATA
TTTGGTTAGAAAACAAATATTTGAATTAGTAAAGTATGCAAATGAAATTAATTTGAATGTATCAGATGCGCCTTCG
CTAGTTTATGCTGATGGCTATGCTTATATGGATGGTGAGGGTACAATCGATTTTTCTGGGATATCTAACAATCATG
30 CTGATCAATTAAGGATTTTTCTTTTTTAAATGGTGGATACCAAGGATGTTCTATTATGTTCAATCGTGCAATGACC
AAATTTCTTCTGAATTATCGAGGATTTGTATATCTACATGACGATATCACACATTAGCTGCATACGCTCTTGGTAA
AGTTTATTTTCTCCGAAATACCTTATGTTATATAGACAGCACACGAATGCGGTAAGTGGTATCAAAACATTCCGC
AATGGATTGACTTCTAAATTTAAATCACCAGTAACTATCTTTTTATCACGAAAACATTATCAGGTAAAAAATCTTTT
TTTGAATGTAACAGCTCTATCTTATCAGAGACGAATAAAAAAGTTTTTTTGGATTTTATTTTCATTTTGTGAATCAAA
35 AATAAATTTACAGATTTTTTAAAGTTATGCGGAGGTGGTTTTAGATTAAATAACAGTAGAACTAAATTTATTTAAA
ATTTCTTAATACGGAGAAAAATTTAGCGAATGTTTTCAACTTTACACCTACTTTTTAATCGGCAACATACTTTATCAAG
GCTATTCAATTCTCTTATATTACAACTGATAAAGATTTTGGTGGATAATAATTGATGATGGTAGTATAGATGCAA
CAGCGGTACTTGTAGAAGATTTTAAAAAATGTGATTTTACTGATTTTATTGCTATCAGGAAAAAATGTTAA
40 GCCCATGGCTTTAAACGCTGGTGTAAAGCTTGTAGAGGCGATTATATCTTTATTGTTGACAGTGATGATGCACT
AACTCCCGATGCCATAAAATTAATTAAGAATCAATACATGATTGCTTATCTGAGAAGGAAAGTTTCAGCGGAGTC
GGTTTTAGAAAAGCATATAAAAAGGGGGGATTTTGGTAATGATTTAAATAATTTCTCAGAACATATACTATTT
AAATCGCACTGAGATTAGCAATTTAATAAATGGTGTGTTGCATTTGTTTTAAAAAAGAAAGTTTGGTAAAAAATC
CATTCCCGCTATAGAAGATGAAAAATTTGTTCCAGAATTATATATTTGGAATAAAAATAACTGACAAGGCGAAGAT
45 TCGATTTAACATAAGCAAAGTTATATATCTTTGTGAGTATCTTGTGATGATGGTCTTTCTAAAAATTTCCATAACCAGC
TTAAAAAATACCCAAAGGGGTTTAAAGTTTATTACAAAAGTCAAAGAAAACGAGAGAAAACCTTATATAAAAAAAC
AAAGATGCTAATTAGATATTTGCAATGTTGTTATTATGAGAAAAATAAATGAAAATACTATTTGTCATTACAGTTT
AGGCCTTGGAGGTGCTGAGAAGCAGGTTTGTCTTTAGCTGATAAATTAAGTTTAAAGCGGCACCATGTAAGAT
TATTTCACTTGGACATATGTCTAATAATAAAGTCTTTCTAGCGAAAATAATGTTAATGTCATTAATGTAATAATGT
CAAAAAACATTTCTGGAGTTATAAAAGGTTGTGTGAGAATTAGAGATGTTATAGCTAATTTCAAACCAGACATTTGT
50 ACACAGTCATATGTTTCATGCAAACATTATCACTAGATTGTCTGTAATTGGAATCAAAAACAGACCTGGTATTATAT
CAACTGCACATAATAAAAAATGAAGGTGGGTATTTTCAAGATGCTCACATATAGAATAACCGATTGTTTAAAGTATTG
TTGTACAAATGTTAGCAAAGAAGCAGTGGATGAGTTTTTACGGATAAAAAGCCTTTAATCCCGCTAAAGCAATTA
ATGTATAATGGGATAGATACCAATAAATTTAAATTTGATTTATTGGCAAGGAGGGAAATTCGAGACGGTATTAATA
TAAAAAATGATGATATATTACTTGTGTCAGGTCGTTTAAACGTTAGCTAAAGATTATCCTAATTTATTGAATGCA
55 ATGACTCTGCTTCTGAACTTTAACTTATTATTATTGGTGTGATGGTGAATTGCGTGACGAAATTAATATGCTTAT
AAAAAATGCAATTATCTAATAGGGTGTCTTGTGGGAGTTAAAAAATAATGCTCCTATTTTTCTGCAATGTG
ATTTTTTGTCTCTTCTCGTTGGGAAGGATTTGATTAGTCTGGCAGAAGCTATGCTCATGTGAGCGAATTTGT
TGTTGGCACGGATTACGGGGGAGTAAGAGAAGTTATTGGTGACGATGATTTTCTTGTACCCATATCTGATTCAAC
ACAACCTGCAAGCAAATTTAAAAATTTGCTTTGAGCCAGATACGTGATCACATTTGGTTTTCGGAATCGTGAGCG
60 TATTTTAAAAAATTTCTCAATAGATACTATTATTATGCAGTGGCAAGAAGTCTATGGAATATAATTTGCTCAAAAC
ATGAAAGGTAGATTTATTTGGAACGTGTCTTTTGTGAAATTAATTAATCAATCTCAATTGAGATTTTTGTATTTCAA
AAATACCATCATAGCTAACGATGATTGGTATTTATTTAAGATGCTTTCTATAAATATATTGACGTTTTTAAATGCGC
CGAAACGATTGGGCTGGGAACAGAGAAGTAAACTGTTTTGAGAATGAAGAGTTTTTGGATGTTTTATGGATATT
AAAAATTGATCCAGTGAATTAATTTATAATAAATCAAGATTTAATGTTAATAAATGATAATCTTTTCTGACACTC

ATATTAATTATGAGTGGTACGTTTTGGTAAACGGTAAACTATTATATGACAGCTAGAACAACATAAAGTTTTGCACTTA
 CAATTACTCCCCTCTAAGTGGCGTTCAAAGGGTAAACATTAACGAAATTAGTGCCTTATATACTGATTATGATT
 ATACACTAGTTTTGCTCAAAAAAAGGTCCACTAACAAAAGCATTGCTGGAATATGATGTCGATTGTCATTGTATCCC
 CGAACTTACGAGAGAAATTACCGTAAAGAATGATTTTAAAGCATTGTTCAAGCTTTATAAGTTCATAAAAAAGAA
 5 AAATTTGACATTGTGCATACACATTCTTCAAAAAACAGGTATTTTGGGGCGAGTTGCTGCCAAATTAGCACGTGTTG
 GAAAGGTGATCCACACTGTACATGGTTTTTCTTTCCAGCCGCATCTAGTAAAAAAGTTATTACCTTTATTTTTTC
 ATGGAATGGATAGCAAAGTTCTTTACGGATAAGTTAATCGTCTTTGAATGTAGATGATGAATATATAGCAATAAACA
 AATTAATAATTCAAGCGGGATAAAGTTTTTTAATTCTAATGGAGTAGACACTGATAAGTTTTCTCCTTTAGAAAAT
 AAAATTTATAGTAGCACCTTGAATCTAGTAATGGTTGGTAGATTATCCAAGCAAAAAGATCCTGAGACATTATTGC
 10 TTGCTGTTGAAAACTGCTGAATGAAAATGTTAATGTTAAGCTGACACTTGTAGGAGATGGTGAACATAAAGAACA
 GTTAGAAAAGCAGGTTCAAACGGCAAGATGGACGTATAATTTTTTCATGGATGGTCAGATAACATTGTTAATATTTTA
 AAAGTTAATGATCTTTTTATATTACCTTCTCTTTGGGAGGGTATGCCATTAGCAATTTTAGAAGCATTGAGCTGTG
 GACTTCCATGTATAGTCACTAATATTCCAGGTAATAATAGCTTAATAGAAGATGGCTATAATGGTTGTTTGTGTTGAA
 15 ATTAGAATGTCAGTTACTTCAAAAAATCATGTATGTTAAGTAAAGCCAGAAGTATTGCCAGACAGTACGAAAGTAA
 CCAATGCACGATCATTATTCTGAAAAATTTGATTGTTAAAGAAATAATAAGGTGAGACAGCTATATGATAAT
 TAAGAGCTCGGTACCCGGGCCTAGGGTGTAGGCTGGAGCTGCTTGAAGTTCCTATACTTTCTAGAGAATAGGA
 ACTTCGGAATAGGAACTAAGGAGGATATTCATATCCGTCGACGGCGGCCCTGCAGGCATGCAAGCTTGATC
 CATATGGATCGCTAGCTTAATTAATAAAGCCGTAAGCATATAAGCATGGATAAGCTATTTATACTTTAATAAGTAC
 TTTGTATACTTATTTGCGAACATTCCAGGCCGCGAGCATTACGCGCGGTGATCACACCTGACAGGAGTATGTAAT
 20 GTCCAAGCAACAGATCGGCGTAGTCGGTATGGCAGTGTGGGACGCAACCTTGCCTCAACATCGAAAGCCGT
 GGTATACCGTCTCTATTTTCAACCGTTCCCGTGAGAAGACGGAAGAAGTATTGCCGAAAAATCCAGGCAAGAAA
 CTGGTTCTTACTATACGGTGAAGAGTTTTGTCGAATCTCTGAAACGCCTCGTCGCATCCTGTTAATGGTGA
 GCAGGTGCAGGCACGGATGCTGCTATTGATTCCCTCAAACCATATCTCGATAAAGGAGACATCATCATTGATGGT
 GGTAAACACCTTCTTCCAGGACACTATTCGTCGAATCGTGAGCTTTCAGCAGAGGGCTTAACTTCATCGGTACC
 25 GGTGTTTCTGGCGGTGAAGAGGGGGCGCTGAAAGGTCTTCTATTATGCCTGGTGGCCAGAAAGAAGCCTATG
 AATTGGTAGCACCGATCCTGACCAAAATCGCCGCCGTAGCTGAAGACGGTGAACCATGCGTTACCTATATTGGT
 GCCGATGGCGCAGGTCACATATGTGAAGATGGTTCACAACGGTATTGAATACGGCGATATGCAGCTGATTGCTGA
 AGCCTATTCTCTGCTTAAAGGTGGCCTGAACCTCACCAACGAAGAAGTGGCGCAGACCTTTACCGAGTGAATA
 ACGGTGAACGAGCAGTTACCTGATCGACATCACAAAGATATCTTACCAAAAAAGATGAAGACGGTAACTACC
 30 TGGTTGATGTGATCCTGGATGAAGCGGCTAACAAAGGTACGGGTAATGGACCAGCCAGAGCGCGCTGGATCT
 CGGCGAACCGCTGTGCTGATTACCGAGTCTGTGTTTGCAGTTATATCTTCTCTGAAAGATCAGCGTGTGTC
 CGCATCTAAAGTTCTCTCTGTTCCGCAAGCACAGCCAGCAGGCGACAAGGCTGAGTTCATCGAAAAAGTTCGTC
 GTGCGCTGTATCTGGGCAAAATCGTTTTCTTACGCCAGGGCTTCTCTCAGCTGCGTGCTGCGTCTGAAGAGTAC
 AACTGGGATCTGAACCTACGGCGAAATCGCGAAGATTTTCCGTGCTGGCTGCATCATCCGTGCGCAGTTCCCTGCA
 35 AAAAACTACCGATGCTTATGCCGAAAAATCCACAGATCGTAACTGTTGCTGGCTCCGTAACGAAATTTGCTG
 CATGACTACAGCAGCGCTGCGTGATGCTGTTGCTTATGCAGTACAGAACGGTATTCCGGTTCCGACCTTCT
 CCGCAGCGGTTGCCTATTACGACAGCTACCGTGTGCTGTTCTGCCTGCCAACCTGATCCAGGCACAGCGTGA
 CTATTTTGGTGGCATACTTATAAGCGTATTGATAAAGAAGGTGTGTTCCATACCGAATGGCTGGATTAA

SEQ ID NO: 18 (Ejemplo de secuencia de nucleótidos de locus O25B rfb – O25B–producción EPA cepa stGVXN4459)

40 ATGACGAATTTAAAAGCAGTTATTCCTGTAGCGGGTCTCGGGATGCATATGTTGCCTGCCACTAAGGCGATACCC
 AAAGAGATGCTACCAATCGTCGACAAGCCAATGATTCAGTACATTGTTGACGAGATTGTGGCTGCAGGGATCAAA
 GAAATCCTCCTGGTAACTACGCGTCCAAGAACGCGGTGCGAAAACCACTTCGACACCTTATGAGTTAGAATCA
 CTCCTTGAGCAGCGCGTGAAGCGTCAACTGCTGGCGGAAGTACAGTCCATCTGTCCGCCGGGCGTGACCATTA
 45 TGAACGTGCGTCAGGGCGAACCTTTAGGTTTAGGCCACTCCATTTTGTGTGCGCGACCTGCCATTGGTGACAAC
 CCATTTGTGCTGGTACTGCCAGACGTTGTGATCGACGATGCCAGCGCCGACCCGCTACGTTACAACCTTGTGTC
 CATGATTGCAGTTTTCAACGAAACGGCCGACGAGTGTGGCAAAAACGATGCCCGGTGACCTCTCTGAAT
 ACTCCGTCATCCAGACTAAAGAGCCGCTGGACCGTGAGGGTAAAGTACGCCGATTGTTGAATTTATCGAAAAA
 CCGGATCAGCCGACAGCTGGACTCAGACATCATGGCCGTAGGTCGCTATGTGCTTTCTGCCGATATTTGGCC
 50 GGAACCTGGAACGTAACCTGAGCCTGGTGCATGGGGACGTATTACGCTGACTGATGCTATTGCCGAGCTGGCGAAA
 AAACAATCCGTTGATGCAATGCTGATGACCGGCGACAGTTACGACTGCGGCAAAAAAATGGGCTATATGCAGGC
 GTTTGTGAAGTATGGCCTACGCAACCTGAAAGAAGGGGCGAAGTTCGTAAGGTATTGAGAAGCTGTTAAGCG
 AATAATGAAAACTGACCGGATGTAACGGTTGATAAAGAAAATTAAACGGCAGTGAATAATCGCAGCAAAAGTAA
 TTTGTTGCGAATCTTCCGCTGCTGTTTTATATAAACATAAGATAAACAACGAGTTAGCAGTAGGGTTTTATTCA
 55 AAGTTTTCCAGGATTTTCTTGTGTTTCCAGAGCGGATTGGTAAGACAATTAGCGTTTTGAATTTTCCGGTTTTAGCGC
 GAGTGGGTAACGCTCGTCACATCATAGGCATGCATGCAGTGTCTGGTAGCTGTAAAGCCAGGGGCGGTAGCG
 TGCATTAATACCTCTATTAATCAAACCTGAGAGCCGCTATTTACAGCATGCTCTGAAGTAATATGGAATAAATTA
 TACGCAAGATAGTGTGTTAATGTCGATAAATTAACATACGCCGAAACCTGGAATCACTTGAGATGTTTCTGAT
 TCTGAACGCTATTTCTTGAACATGCGGATATTTGTGATGCAGCTGCAATGGCAGGTTTTTCTCAGCATCAG
 60 CCGGATGCAGTATTTCTGACCTGGCAGCTGAAAGCCATGTTGACCGTTCAATTACAGGCCCCGCGCATTTATTGA
 AACCAATATTGTGGGTACTTATGCTTTTTAGAAAGCGGCTCGGAATTTTGGTCTGGTCTGGATGATGAAAAGAA
 AAAAACTTCCGTTTTCTCATATTTCTACTGATGAGGTGTATGGTACTTACCCCATCCGGATGAAGTAAATAGC

AATGAAACGTTGCCGCTATTTACGGAAACGACAGCATACGCGCCAAGTAGTCCATATTCTGCTTCTAAAGCTTCC
AGCGATCATTGGTTCGCGCATGAAACGTACTTATGGTTTACCGACCATTGTGACTAATTGCTCGAACAACTAT
GGTCCTTATCATTTCGCGAAAAGCTTATTCCACTGGTTATTCTTAATTCAGTGAAGGTAAGGCATTACCTATTT
5 ATGGCAAAGGAGATCAGATCCGCGACTGGTTGTATGTAGAGGATCATGCTCGAGCGTTATATACCGTCGTAACC
GAAGGTAAGCGGGCGAAACTTATAACATTGGTGGACACAACGAAAAGAAAAACATCGACGTAGTGTTCACTATT
TGTGATTTGTTGGATGAGATAGTCCCGAAAAGAGAACTTACCAGCGAGCAAATTAATGTTACCGATCGTCCG
GGACACGATCGCCGTTATGCGATTGATGCTGAGAAGATTGGTCGCGAATTGGGATGGAAACACAGGAAAACGTT
TGAGAGTGGGATTCGTAACCGGTGGAATGGTACCTGTCCAATACAAAATGGGTTGATAATGTGAAAAGTGGTG
10 CCTATCAATCGTGGATTGAACAGAACTATGAGGGCCGCCAGTAATGAATATCCTCCTTTTTGGCAAACAGGGCA
GGTAGGTTGGAACTACAGCGTGTCTGGCACCTCTGGGTAATTTGATTGCTCTTGATGTTCACTCCACTGATTA
CTGTGGTATTTTATGTAATCCTGAAGGTGTAGCTGAAACCGTAAGAAGCATTCCGGCCTGATATTATTGTC AACGC
AGCCGCTCACACCGCAGTAGACAAAGCAGAATCAGAACCGAAGTTTGCACAATTACTGAACGCGACGAGTGTCCG
AAGCGATCGCAAAGCAGCCAATGAAGTCGGCGCCTGGGTTACTACTACTACTGACTACGTATTTCCGGGG
ACCGGTGAAAATCCATGGCAGGAGGATGCAACCCGACCCGCTAAAATGTTTACGGTAAACCCAGGTAAGCGT
15 GAGAAAAAGCATTACAAGAGCATTGTGCGAAGCACCTTATTTTCCGGACCAGCTGGGTCTATGCAGGTAAAGGA
AATAACTTCGCCAAAACAATGTTGCGTCTGGCAAAGAGCGTGAAGAATTAGCCGTTATTAATGATCAGTTTGGT
GCGCAACTGGCGCAGAGTACTGGCTGATTGTACGGCACATGCTATTCTGTGTGGCACTGAATAAACCGGAAGT
CGCAGGCTTGTACCATCTGGTAGCTAGTGGTACCACAACGTGGCACGATTATGCTGCGCTGGTTTTTGAAGAGG
CGCGCAAAGCAGGCATTCCCTTGCACCTCAACAAGCTCAACGCAGTACCAACAACAGCCTATCCTACACCAGCT
20 CGTGCCTCCACAATACTCGCCTTAATACAGAAAAATTTGAGCAAACTTTGCGCTTGTCTTGCCTGACTGGCAG
GTTGGCGTGAACGAATGTTAACGAATTTTACGACTACAGCAATTTAATAGTTTTTGCATCTTGTTCGTAATG
GTGGAGCAAGATGTATTAAGGAATGATGAAATGAAACCGCTAAAGGTATTATTTTGGCGGGTGGTTCTGGTA
CTCGTCTTTATCCTGTGACGATGGCCGTCAGTAAACAGCTGTTACCGATTTATGATAAACCGATGATCTATTACC
GCTCTCTACACTGATGTTAGCGGGTATTCCGCGATATTCTGATTATCAGTACACCACAGGATACTCCTCGTTTTCAA
25 CAACTGCTGGGTGACGGGAGCCAGTGGGGCCTGAATCTTCAGTACAAAGTGCAACCGAGTCCGGATGGTCTTG
CGCAGGCGTTTTATTCGGTGAAGAGTTTTATTGGTGGTGTGATTGTGCTTTGGTACTTGGTGATAATATCTTCTA
CGGCCACGACCTGCCGAAGTTAATGGACGTAGCTGTTAACAAAGAAAGTGGTGCAACCGTATTTGCCTATCACG
TTAATGATCCTGAACGTTATGGTGTCTGGAGTTTGATAATAACGGTACTGCAATTAGCCTGGAAGAAAAACCGC
TGGAACCAAAAAGTAACTATGCGGTTACTGGGCTTTATTTCTATGACAATGACGTTGTGGAATGGCGAAAAACC
30 TTAAGCCTTCTGCCGAGGTGAACTGGAAATACCAGATTAACCGTATTTATATGGAACAAGGACGTTTGTCTG
TCGCTATGATGGGGCGTGGCTATGCATGGCTGGATACAGGGACGCATCAAAGTCTTATTGAAGCAAGCAACTTC
ATTGCCACCATTGAAGAGCGCCAGGGACTAAAGGTTTCTGTCCGGAAGAAATTGCTTATCGTAAAGGGTTTATT
GATGCTGAGCAGGTAAAAGTATTAGCCGAACCGTTGAAGAAAAATGCTTATGGTCAGTATCTGCTCAAATGATT
AAAGGTTATTAATAAGATGAACGTAATTAACCTGATGTGCTGATTTTTGAACCAAAAGTTTTTGGG
35 GATGAACGTGGCTTCTTTTTGAGAGTTTTAATCAGAGGATTTTTGAAGAAGCAGTAGGTCGTAAGGTTGAGTTT
GTTACAGGATAACCTTCTAAGTCCAGTAAAGGTGTTTTACGTGGTCTTCATTATCAGTTAGAACCCTGCTCAAG
GAAAACCTGGTGCCTGTGTTGTTGGCGAGGTTTTGATGTTGCGGTTGATATTGTAATCGTCACCTACATTTG
GGAAATGGGTTGGGGTGAATTTGTCTGCTGAGAATAAGCGTCAGTTGTGGATTCTGAGGGATTTGCACATGGT
TTTTTGGTGTGAGTATTTAGCAGAAGTTTTATATAAACGAATCAATATTATGCTCCATCACATGAAAAAATAT
40 TATATGGAATGACCTCTTGTCTAATATTAATGGCCGAGCACAGCACTGATCACTCTGTCTGATAAGGATGCAAT
GGGAAAAGATTTGAACTAAGTGAATTTGAAATGCTCTCTTAAACATAGTATATGGAATGTTGCGGGCTACTTT
ATACCAACATTAATTGCAATTCGCGCCTTTGGATTAATTTGCGAGGAAATTTGGTGTAGAATTTGGTTGTATA
CGTTAGCAATGATTTTTATAGGATGCAAGTATTTGATGCTGGGTTAACAAAGAGCTGTTGTGCGTGAATAG
45 CATTACTAAAAACAGAGTGGACGATTGTAATACGATAATAGTAACCTTCTATTATCGCTGTGATTTTTAGGGTTT
ATCGGAGGCGGGGAGTGTCTGCTTAAAGGCGATATTATTGAACTGTTAAATATCTCACCAATATATTACGCC
GATTCGATAAAGTCTCTAGTATTATTATCATCTCTGATACCTGTATTCTTAGTCACGCAATACTATTAGCAGAGCT
TGAGGGTCGGGAATTTTTGGGATTCTAAATATACAAAAAGTGTAGGGAATTTTAAATTGCAGGGTTACCTGC
ATTATTTGTTTTAATTAATCAAACGCTTTTTTCTGCAATTATTGGTGTAGCGATTGCAAGAGTTATATGCTTGTGGT
50 TAAGCTACATTATGAGCAGGGAAAGAATAACTATCGATATCTCATTTTTTCAATAACTGTTTTAAAGCGGTTATTT
AGATATGGCGGGTGGGTAACATAAAGTAACATAATATCTCTATATTAGCGAGTATGGATAGATTTATTCTATCCC
ATATCCAGGGAGCATCAAAAATATCATTCTATACAGTCCCTAATGAGCTGGTAACTAGGCTTGGAAATAGTTCCAG
GCTCTCTTGGGAAAGCTGTTTTTCCAAAATTAAGTCATGCAAGGAATTTTACAGCGTCATATGCAGAGCAAAAAA
AGCTTATATATTAATGACTGTCATTGTAATGCCTTTGGTTTTATTTGTATATTATTACGCAAAGTTTATTTTAACT
GTGGATGGGGGCTGAGTATGCAGGGATTTCCGTCGAAATATTACGGATTATGCTTATAGGGTATATTTTTAACTG
55 TTATTCACAAATCTCTTTTGGCAACATACAGGCCTTTGGAAAAGCAAAATACACTGCATACATCCATATGATGGAA
TTTTATCTTATTGATAATGTTATATAATTAACAAAGGAATGCGGTTATTGGTGTTCGTTGATTTGACAAAT
TCGAGTAATAATTTGATGCTTTTTATATAGTTATGATTCGTTGTAATAATCTTATAAAAAAGGGTACCTG
ATGATATATATTGTGGTATTAATTTGGAATGGGGCTATAGATACCATTAATTGTGTTAAAAGTTTTAATGGATTTAAA
TGTTAGCGATTATAAAATTATCATTGTTGATAACTGTTCTATGGATAACTCATATGATACTATAAAAAAGAAATCTTAA
60 TTCATTATATATTGCTGATAAAAAGTATCATTGAGGTGAAGTATGAGGATAGAAATAAATATAAAACCTTAGAAAACG
ATAAAATCATATTAATACAATCTCCGCAAAAATAATGGGTACGCAAGTGGTAATAATATTGGCATAGAGTTCGCTCT
TAATCAGGAGAAATGAAATACGTCTGGGTTCTGAATAATGATACTGAAGTGGATAAAGAGGGCTTAACTCATTTA
ATTAGTAAATGTGATTCAGATAAAAAGTATAGGGATTTGCGGTTCTCGTTTAGTCTATTTTCCGACAGAGAGATGC
AGCAAGGACTAGGTGGGGTGCATAACAAATGGTTATGCACTACAAAAAATTATGAAATGGGAAGATTAGTTTCCA

AAAAATATGATGATGAAGTCATTAGTAATGATATAGATTATATAAATTGGCGCATCGATGTTTTTCTCTAGAGAATGT
TTGGAAACAGTTGGATTGATGAATGAAGAATATTTTTATACTATGAAGAGTTAGATATTTGCCTCAGAGCAAAG
CAAAGAACTTAAATTAGGTATTTGCTCAGAAAGTTGGTTTATCATAAAATAGGTGCAAGTACTGATGGGGGAAA
GAGCATGATGGCTGATCTTTGCTCAATAAAAAATAGGCTGGTCATTACAGAAAGTTTTATCCCAATATTATTGG
5 ACGGTATGGTTGTCACTTTTTGTGTAGCATTAAACCGTGCTAGAAGAGGTGAGTTTAATAAGATGAAAAGATGTT
TGAATGTTATGTTAACTTCAAACGAAACAAAGGTAGCAAATGCCATTAGAATATGCACTTAATCATGGTGTTAAT
AAATCTATAGTTTGATATGTTATTAAGGGTATTTAATGAAAGTGGCTTTTTTATCTGCTTATGATCCACTATCTAC
ATCCAGTTGGTCTGGCACACCTTATTATATGCTAAAGGCATTATCGAAGAGAAATATTTCCATTGAAATATTAGGA
10 CCGGTAATAGCTATATGATATACATGTTAAAAGTATATAAAATTAATTAAGGTGTTTCGGAAAAGAATATGATTA
TAGTCATTGCAAGTTGCTTTCCAGGTATTACGGTAGAATATTCGGTAGGAAATTAAAAAAATTGATGGTTTGGAT
TTTATTATCGCACCTGCAGGTTCCCTCAAAATTGCTTTTTTAAAAACAACCATACCAATAATATATCTATCGGATAC
AACATATGATCAATTAAGGCTATTATCCGAATTTAAATAAAAAACAATTATAAATGATGAGGATGCAAGTTTAA
TCGAACGCAAGGCTATTGAAAAGCAACAGTAGTATCTTTCCCATCTAAATGGGCAATGGATTTTTGCAGGAATT
15 ATTACAGATTAGTTTTGATAAATTAGTTGAAATACCATTGGGGGCTAAATTTATTTGATATTCCTTTGCTAAT
AAAAATATAATTCAAAAGAATAGTTATACCTTGTCTTTTTCTGGGAGTTGATTGGGAAAGAAAAGGTGGGAAAACAG
CCTTGAAAGCAATTGAATATGTAAGGCAGTTATATGGGATCGATGTTAGACTAAAAATTTGTGGATGTACTCCGAA
TCAAAAGATTTTACCTACTTGGGTGAATTAATTGATAAAGTAGATAAAAAAATACGTTGACGAATATCAGAAATTC
TCGATGTGTTATCTAACGCTGATATACTTCTTTTACCAACCATTGCTGAATGTTATGGAATGGTATTTTTGTGAAGCT
20 GCTGCTTTTGGATTGCCTGTTGTCGCTACAGATACAGGTGGAGTCAGTTCTATAGTTATCAACGAAAGGACGGG
GATATTAATTAAGACCCGTTAGACTATAAGCACTTTGGAAATGCAATTCATAAAATAATTAGTTCCGTAGAGACT
ATCAAAACTACTCCAAAACGCAAGAATTAGATATAAATAATATGTCATTGGGACAATTTGGGCAATTTAAAGATAAT
TGAGATTATGATGAGCATAAGAATAAGAATCAAATAGCACAAAAAGAATTATATGTTTATTTATACTTTTTCTT
25 GTTTTCCCTGATTTTTGTTTTATACATTAGGGGTTGATAATTTAGCATTTCACGATAATCTCAATTACATTGCTT
TTTGTTTTTTAAAGAGCTAAAAATTTGCAAAGATAATTTCTAATAATAGTAGCGTTATTCATATTGTTGTGTTTT
AACTGTTTGTAAAGTATGCTATTTAATATTGAACAGGCTTTAACATTTAAAGTTGTACTTTCAATATATAGCATCTTA
ATAATGGCATAACGCTCCTCTTGTATGCACAGACGTTGTGGTTATGTTCTGAAGAAATACCTAAGAGATCCGTCT
TTTATTTGTTCCGATTTCTTTGCCTTATTGGCATTATAAGTATTTCTTTTACAGAAGACTGAGATTATACATATAAAA
30 GATGATTTCTTTTTCTGAACCATCAGCATTGTCATTGGTTTTTATACCTATCTTTTCATTTTGTTTATACTATA
GAGGGGGGGGCTACTATTGCTCTATATATTATCTTTGGGTATTGCCGTTAGGTATCCAGAAATTTAAACAATGTTGG
TAGGCATTGTGATTAGTGTTTTTGTGATGAAAAAATAACTATAAGGCAAACTATTGTTATACTTTTGGGGCATG
GATTTTTCCATGATATTAAGTGATTTAGACATTTCTACTATACATCGCGGCTTGATTTTTAAAAATACTACGAACC
TATCAGTGCTTGTATATCTTTCAGGAATTGAAAGAGCTTCTTGAATTTTATTACAAGTTATGGTCTTGGTATTGGT
TTTCAACAATGGGAGTGAATGGGGAGATAGGAATATATCAACAATTTTAGCTGAACCTTGATGCCCTATGTTAA
35 ATATACCGATGGCTCATTTATTTCTTAAGTTAATATCTGAGTTGGGGTTATTGGTGCATTAATGTGTATTTTC
TATTTTTTTATTTTTCCCGATTTTATCTGCGTTTCAAAAAAGTAAAGAGATATTCACCCGAGTATTTTTAGCATAT
AGCTTCTACATGTGTTTTCTTCATCCCTCTTTTTATACGTGGTGGTTATATAAAACCCGATTTGTTTTATGTTATT
TTTCAATATTTTTGTGCAAAATATCACGCTAAAAATATCTTGTGAAATCTAATGTCCAGATAGCTATATAATAGT
AGATTATATTACATTACACGTAATTACATATTAATAGCATATATGATAACTAGGACATAAATAATGTGCATTTAAA
40 AAAAACTTAAGTTAATTAACGATATGGCCTTTATGGTGGTCTTAGGCTTCTTAAAGATATATTCTTAAACAAATT
TTTATTTTGTTCAAATGTTAGGATTATTAGATTTCCATGTTATATTAGAAAAGATGGAAGTGTAGTTTTGGAAAAG
GTTTTACATCAGGTGTAGGATTACGAGTTGATGCATTTATGGATGCCGTAGTTTCCATTGGAGAAAATGTTCAAAT
TAATGACTATGTTACATCGCGCTATTAATAATGTCATTATTGGTAGAGATACATTAATAGCAAGTAAAGTATTTA
TTAGTGATCATAAATCATGGTATTTTTCTAAATCCGATATCCATAGTTCCACCAACTATTATTCTTCTAGGCCCC
45 CTTGAATCTGCACCTGTGTATATTGGAGAGCGTGTGTGGATTGGCGAAAATGTGACAATATTACCAGGTGCGTGT
ATAGGTAATGGTGTAGTTATTGGCGCAAACAGTGTGTTCTGTTGGTGGAGATTCTAATAATGTGATCATTGCTGGT
GTTCCAGCTAAAATTGTTAAAAAATATAACTATGAGCGTATGCAATGGGAAAAGAATATAGTTGTAATATCGGCTGT
TAATTTTACAACCGGAGGCCCTTTACCGTACTAAAAATGTGCTTACAGCAACTAAAGATAGAGCCGAATGTAA
ATTTATTGCACTGGTTCATAGCTCTGCTGAACTAATGGAATTTTCCGTGGGTTGAATTTATAGAGTATCCAGAA
50 GTCGAAGTCTTGGTGGTTAAAAGATTATATTCGAATATAAATACTTGCATAGATTATCTAAGGTGATTAAGGCAA
CTCATTGGGTATGCTTACATGATATTACAGCAAATGTTAGTACTCCCTATAGATTTGTTTTATTGCCACAATCCTGC
ACCGTTCTATAAATATTTAAGCTATCGAGATATTATAGGAAACCTAAATTTTTATCTTTTTTATCTTTTTTATGGGCT
TTTATACAATATCAATATAAAAAAGAACACAGCAGTTTTTGTTCAGCAGCAGTGGCTAAAAAAGAATTCGAAAAA
AAATATAAGTTAAAGAATGTTGTTGTTAGTCGCCCTGAAGATATTTGCCCTTTTGAAGGTGATGGTTTGGTAAGAA
ATAATAAAAAAGGATGTGAGGATTTTTACCCAGCAGTGCCCGTATATTTAAAAACTTTGAAGTTATCATACG
55 TGCTGCACAAATATTACAAGATAAAAAATTTTATCTTACTTTTTGATGGTACTGAAAATAAGTATGCAAAAA
GAATATAAATAGCTTCCGAACGAAAAATGTTACATTTCCGTTACCTTAATGCAACCGAGATGGTTAACTT
TTATCAAGATTAGATTTATTTGTTTTCCCATCGAAACTGAAACGTTGGGGATTACCATTACAGAAAGCTAAAAACA
TACAAAAAATGGATATTGCGGCAGACTTACCTTATGCTCATGAAGTTTTTATAAATACTTCAAAAACTAGATTTT
TCCATTTGACGATGAGAAAATACTTGTTCGCTACATATTAGAGTACACAAGTAAAAATATGCATGAAGATATAAAA
60 AATAGTAGGGTGAATTTAATAATGATGCATTGACTGGTTTTGAACAGTTTTATTGAATATATCCTCAAGGGGA
GACGTGGTTTATATTATAATCGTTTACATGGCCATGATGACTATATAGAAAATCTTTTATTAATTTAAAGTTGCC
CTCTGGAAGATTTAAAAATAAGTTTCGTGATAACAAAAGTTCAATGGTTTTAAAAAACAATGCGAAAAAATTC
GTAACCTATTTGCATGGAGGGCAATATGGATTTGGACATAAATAACATAGCAGTGTATATAAATTAATAACT
TCATGATTATGAATAATGATTATTTTCTTTTAAACCCCGATGATTTCATAACCAGTGAAGTTTGATTAATTATG

TTGATTATATAATTAGTAATGATTATAAGTTTAGCACATTATGTCTTTATCGAGATTTTACTAAAAGCAAACATGATT
 ATTCAATACGGAGTTTTCCAACTTTATATGATTTTCTTTGTTCTTTTTATTGGGGGTGAATAAAAGTAAAATTAAGA
 AGGAAAATATACTTTCTGATACTGTAGTTGATTGGTGTGCTGGCTCATTATGCTTATTCATGCTTTAAGTTTCTTA
 AATGTGAATGGTTTTGATCAAAAATATTTTATGTATTGTGAAGATATTGACCTTTGTATGCGTTTTAAAATTAAGTGG
 5 AGTAGATCTTTACTATACTCCCCATTTTGTATGCTATTCATTATGCGCAGCATGAAAATAGAAGAATATTTACTAAAG
 CATTTCGATGGCATATAAGGAGTATTACGCGCTACATATTACGGAACCAATTCTTTCTTATAAAAACTATAGAAA
 AATTACATCCGAACGTGTAAGTGATTAAGGATCCGTGTAGGCTGGAGCTGCTTCGAAGTCCCTATACTTTCTAG
 AGAATAGGAACTTCGGAATAGGAACTAAGGAGGATATTCATATGGATAAAGCCGTAAGCATATAAGCATGGATAA
 GCTATTTATACTTTAATAAGTACTTTGTATACTTATTTGCGAACATTCCAGGCCGCGAGCATTAGCGCGGTGATC
 10 ACACCTGACAGGAGTATGTAATGTCCAAGCAACAGATCGGCGTAGTCGGTATGGCAGTGATGGGACGCAACCTT
 GCGCTCAACATCGAAAGCCGTGGTTATACCGTCTCTATTTTCAACCGTTCCCGTGAGAAAGACGGAAGAAGTATT
 GCCGAAAATCCAGGCAAGAACTGGTTCCTTACTATACGGTGAAGAGTTTGTGCAATCTCTGGAAACGCCTCGT
 CGCATCCTGTTAATGGTGAAAGCAGGTGACGGCAGGATGCTGCTATTGATTCCCTCAAACCATATCTCGATAAA
 GGAGACATCATATTGATGGTGAACACCTTCTTCCAGGACACTATTCGTCGTAATCGTGAGCTTTTCAGCAGAG
 15 GGCTTTAACTTCATCGGTACCGGTGTTTCTGGCGTGAAGAGGGGGCGCTGAAAGTTCCTTCTATTATGCCTGG
 TGGCCAGAAAGAAGCCTATGAATTGGTAGCACCGATCCTGACCAAAATCGCCGCGTAGCTGAAGACGGTGAAC
 CATGCGTTACCTATATTGGTGCCGATGGCGCAGGTCATATGTGAAGATGGTTCACAACGGTATTGAATACGGC
 GATATGCAGCTGATTGCTGAAGCCTATTCTCTGCTTAAAGGTGGCCTGAACCTCACCAACGAAGAACTGGCGCA
 GACCTTTACCGAGTGGAAATAACGGTGAAGTGAAGTACCTGATCGACATCACCAAGATATCTTACCAAAAA
 20 AGATGAAGACGGTAACACTCTGGTTGATGTGATCTGGATGAAGCGGCTAACAAAGGTACCGGTAATGGACCA
 GCCAGACGCGCTGATCTCGGCGAACCGTCTGCTGATTACCGAGTCTGTGTTGACGTTTATCTCTTCT
 CTGAAAGATCAGCGTGTGGCCGATCTAAAGTCTCTCTGGTCCGCAAGCACAGCCAGCAGGCGACAAGGCTGA
 GTTCATCGAAAAGTTCGTCGTGCGCTGTATCTGGGCAAAATCGTTTCTTACGCCAGGGCTTCTCTCAGCTGC
 GTGCTGCGTCTGAAGAGTACAACCTGGGATCTGAACTACGGCGAAATCGCGAAGATTTTCCGTGCTGGCTGCATC
 25 ATCCGTGCGCAGTTCCCTGCAGAAAATCACCGATGCTTATGCCGAAAATCCACAGATCGTAACCTGTTGCTGGC
 TCCGTACTTCAAGCAAATTGCCGATGACTACCAGCAGGCGCTGCGTGATGTCGTTGCTTATGCAGTACAGAACG
 GTATTCCGGTTCCGACCTTCTCCGCAGCGGTTGCCTATTACGACAGCTACCGTGCTGCTGTTCTGCCTGCGAAC
 CTGATCCAGGCACAGCGTGACTATTTTGGTGCGCATACTTATAAGCGTATTGATAAAGAAGGTGTGTTCCATACC
 GAATGGCTGGATTA

30 **SEQ ID NO: 19** (Ejemplo de secuencia de nucleótidos de locus *O75 rfb* – *O75*–producción EPA cepa stLMTB11737)

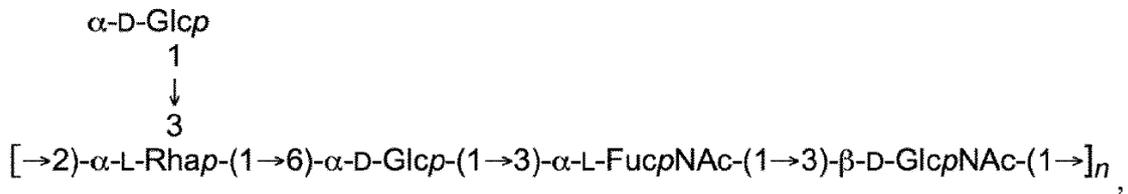
ATGACGAATTTAAAAGCAGTTATTCCTGTAGCGGGTCTCGGGATGCATATGTTGCCTGCCACTAAGGCGATACCC
 AAAGAGATGCTACCAATCGTCGACAAGCCAATGATTTCAGTACATTGTTGACGAGATTGTGGCTGCAGGGATCAAA
 GAAATCCTCCTGGTAACTCACGCGTCCAAGAAGCGGGTCAAAAACCACTTCGACACCTCTTATGAGTTAGAATCA
 35 CTCCTTGAGCAGCGCGTGAAGCGTCAACTGCTGGCGGAAGTACAGTCCATCTGTCCGCCGGGCGTGACCATTA
 TGAACGTGCGTCAGGGCGAACCTTTAGGTTTAGGCCACTCCATTTTGTGTGCGCGACCTGCCATTGGTGACAAC
 CCATTTGTGCTGGTACTGCCAGACGTTGTGATCGACGATGCCAGCGCCGACCCGCTACGTTACAACCTTGCTGC
 CATGATTGCAGTTTTCAACGAAACGGGCCGAGCCAGGTGCTGGCAAAACGATGCCGGGTGACCTCTCTGAAT
 ACTCCGTATCCAGACTAAAGAGCCGCTGGACCGTGAGGCTAAAGTCAGCCGCATTGTTGAATTTATGAAAAA
 40 CCGGATCAGCCGCAGACGCTGGACTCAGACATCATGGCCGTAGGTCGCTATGTGCTTTTCTGCCATTTGGCC
 GGAACGGAACGTAAGCCTGCTGATGGGGACGATTTACGCTGACTGATGCTATTGCCGAGCTGGCGAAA
 AAACAATCCGTTGATGCAATGCTGATGACCGGCGACAGTTACGACTGCGGCAAAAAAATGGGCTATATGCAGGC
 GTTTGTGAAGTATGGCCTACGCAACCTGAAAGAAGGGGCGAAGTCCGTAAGGTTATTGAGAAGCTGTTAAGCG
 AATAATGAAAATCTGACCGGATGTAACGGTTGATAAGAAAATTATAACGGCAGTGAAAATTCGCAGCAAAAGTAA
 45 TTTGTTGCGAATCTTCTGCGGTTGTTTTATATAAACCATCAGAATAACAACGAGTTAGCAGTAGGGTTTTATTCA
 AAGTTTTCCAGGATTTTCTTGTTCAGAGCGGATTGGTAAGACAATTAGCGTTTGAATTTTCCGGTTTTAGCGC
 GAGTGGGTAACGCTCGTCACATCATAGGCATGCATGCAGTCTGCTGGTAGCTGTAAGCCAGGGGCGGTAGCG
 TGCATTAATACCTCTATTAATCAAACGAGAGCCGCTATTTTACAGCATGCTCTGAAGTAATATGGAATAAATTA
 AGCTAGCAGTGAAGATACTTGTACTGGTGGCGCAGGATTTATTGGTCTGCTGTTGTTGCTCACATAATAAATAA
 50 TACGCAAGATAGTGTGTTAATGTCGATAAATTAACATACGCCGAAACCTGGAATCGCTCGCTGAAATTTCTGA
 TTCTGAACGTTATTCATTTGAGCATGCAGATATCTGCGATGCCGAAAGCGATGGCTCGTATTTTTCGCACAGCACA
 GCCAGACGCGGTGATGCACCTGGCAGCAGAGGCCAGCTGACCGCTCAATAACTGGCCCTGCGGCATTTATT
 GAAACCAATATTGGGTACTTATGTTCTTTTGAAGCGGCGCAATTTATTGGTCTGGTCTGATGATGAAAAA
 55 AAAAAAACTTCCGTTTTCATCATATTTCTACTGATGAGGTGTATGGTACTTACCCCATCCGGATGAAGTAAATA
 GCAATGAAACGTTGCCGCTATTTACGGAATGACAGCATACGCGCAAGTAGTCCATATTCTGCTTCTAAAGCTT
 CCAGCGATCATTGTTGCGCGCATGAAACGTAATGTTTACCGACCATTTGACTAATTGCTCGAACAACCT
 ATGGTCTTATCATTTCCCGGAAAAGCTTATTCCACTGGTTATTCTTAATGCACTGGAAGGTAAGGCATTACCTAT
 TTATGGCAAAGGAGATCAGATCCGCGACTGGTTGTATGTAGAGGATCATGCTCGAGCGTTATATACCGTCGTAA
 60 CCGAAGGTAAGCGGGCGAAAACCTTATAACATTGGTGACACAACGAAAAGAAAACATCGACGTAGTGTCACT
 ATTTGTGATTTGTTGGATGAGATGCTCCGAAAGAGAAAATCTTATCGTGAGCAAATTTACCTGTTGCTGATCGCC
 CAGGCGATGATCGCCGTTATGCAATGATGCCGATAAAAATAGCCGCAATTGGGCTGGAACACAGGAAACG
 TTTGAGAGCGGGATTCGTAAACTGTGGAATGGTATCTGTCCAATACAAAATGGGTTGATAATGTAAAAAGTGGT
 GCCTATCAATCGTGGATTGAACAGAACTATGGGGGCCCACTAATGAATATCCTCTTTTTGGCAAAACAGGG

CAGGTTGGTTGGGAACACTACAGCGTGCTCTGGCACCTCTGGGTAATTTGATTGCTCTTGATGTTCACTCCACTGAT
TACTGTGGTGATTTTAGTAACCCTGAAGGTGTGGCTGAAACCGTTAGAAGCATTGCGCCTGATATTATTGTCAAC
GCAGCCGCTCACACCGCAGTAGACAAAGCAGAATCAGAACCAGGAGTTTGCACAATTACTGAACGCGACGAGTGT
CGAAGCGATCGCGAAAGCAGCCAATGAAGTCGGCGCTTGGGTTATTTACTACTCTACTGACTACGTAATTTCCGG
5 GGACCGGTGAAATACCATGGCAGGAGGAGGATGCAACCGCACCCGCTAAATGTTTACGGTGAACCAAGTTAGC
AGGAGAAAAAGCATTACAAGAGCATTGTGCGAAGCACCTTATTTCCGGACCAGCTGGGTCTATGCAGGTAAGG
GAAATAACTTCGCCAAAACGATGTTGCGTCTGGCAAAAAGAGCGTGAAGAATTAGCCGTTATTAATGATCAGTTTG
GTGCGCCAACCTGGCGCAGAGTTGCTGGCTGATTGTACGGCACATGCCATTTCGTGTGGCACTGAATAAACCGGA
AGTTCGCAGGTTTGTACCATCTGGTAGCCAGTGGTACCACAACCTGGCACGATTATGCTGCGCTGGTTTTTGAAG
10 AGGCGCGCAAAGCAGGCATTCCCCTTGCCTCAACAAGCTCAACGCAGTACCAACAACAGTCTATCCTACACCA
GCTCGTCTCCACATAACTCTCGCCTTAATACAGAAAAATTTTACGACAGAACTTTGCGCTTGTCTTGCCTGACTGG
CAGGTTGGTGTGAAACGCATGCTCAACGAATTTTACGACTACAGCAATTTAATAGTTTTTGCATCTTGTTCGTG
ATGGTGGAAACAAGATGAATTAAGGAATGATGGAATGAATACGCGTAAAGGTATTTTACGCGGTGGTTCTG
GTACCTGCTTTTACTCTGTACTATGGCTGTCAGTAAACGCTTATACCATTATGATAAACCGATTATTA
15 CCCGCTCTCTACTGATGTTGGCGGGTATTTCGCGATATTTTATTATCAGCACGCCACAGGATACTCCTCGTTT
TCAACAACCTGCTGGGTGATGGGAGCCAGTGGGGGCTAAATCTTCACTACAAAGTGAACCGAGTCCGGATGGT
CTTTCGCAGGCATTTATCATCGGTGAAGAGTTTATCGGTGGTGTGATTGTGCTTTGGTACTTGGTGATAATATC
TTCTACGGTCACGACCTGCCTAAGTTAATGGATGCCGCTGTTAACAAAGAAAGTGGTGAACCGGTATTTGCCTAT
CACGTTAATGATCCTGAACGCTATGGTGTGCTTGGATTTGATAAAAAACGGTACTGCAATCAGCCTGGAAGAAAA
20 CCGTTACAACCAAAAAAGTAATTAATGCGGTAACCGGGCTTTATTTCTATGATAACTACGTTTGGAAATGGCGAAA
AATCTTAAGCCTTCTGCCCGGTGAACCTGAAATTTACCGATTAACCGTATCTATATGAAACGAGGGGCTTTA
TCTGTTGCCATGATGGGACGTGGATATGCCTGGCTGGACACGCGGGACACATCAAAGTCTTATTGAAGCAAGCAA
CTTCATTGCCACCATTGAAGAGCGCCAGGGCTTGAAGTTTCTGCCCAGAAAGAAATTGCTTACCGTAAAGGGT
TTATTGATGCTGAGCAGGTGAAAGTATTAGCTAAACCGCTGAAAAAAATGCTTATGGTCAGTATCTGCTAAAAAT
25 GATTAAGGTTTATAATAAATGAATGTTATTAACAGAAATTCAGATGACTGATTTTTGAACCGAAAGTTTTT
GGTGATGAGCGTGGTTTTCTTATGGAAAGCTTTAATCAGAAAGTTTTCGAAGAGGCTGTAGGGCGGAAGTTGA
ATTTGTTCAAGGATAATCATTCTAAATCGTGAAGGTGACTTAGAGGTTTACACTTTTACGCTTCTCCTTTTGGAG
CAGGCAAAATAGTAAGGTGTATAGTTGGCGAGGTATTTGATGTTGTCAGTAGACATTAGACCTAATTTCTGAAACA
TTTGGTTTATGGTTGGAGTAACTCTTTCGTCAGAAAAATAAAGGCAGCTATGGATTCCAGAAGGATTTCGCCCAT
30 GGTTTTTTAACTTTAAGTGATATTGCAGAGTTTGTATAAAACTAACAACCTATTATTCTTTAATCATGAAAGGGG
AGTCATTTGGAACGATGAGGAAATTAACATTGCCTGGCCCTCTCAATCAGAGAAGATTCTGTACAGAAAGATAT
TAATTTACCATCATTAGATTTGTTCAAATGTTTAGCAAGTAGTGTATCTTTACACTGCACATAGTCATCATTTTTT
ATGCTTTAAGTAAATTATATTGCACATCTATAACACAAAGCGCAATAATTTTCGACCTGATGAAGGTTTGTGGTTA
TTTATCTTTCTAGGCGTTTTTATGACTAAAATAGTTGTGGTTTTCTACAGCTCCAATATTTCCGACAAATAATGGGT
35 ACAAAGTTCTGTATTAGGAAGAATTGATGAGTTATTAATGAGGATAATGAGGTCTGTTTTGATTGAAATAAACCT
TGAAAATGTTACGGAAAAGAAAGATGAATTAATACCAACAAGATTTAATAATATTCAAAGATATGAAGTAAAAAAA
TATCTAGATCATTATTGCCGAGTTACAAATATTATTTGATATCAGAAGTTCGGTATGAACAATTATTTTCTTCTGCT
GACATTAGAGATAACATAAAAAAGATAATTGATTTAGAAAAACCTTCTATTATTATTGCTGAGTCTATATGGGCGTT
GCAAGCATTGCCTATTGAAATTAGTGCAGAAATACACTGTGTTATTCATGATGTGGCAACTGATTTCTTTAAAGAA
40 ATGTTTGTATCTCATAATGAGGTTGTACGAAAAATTTGTTTTTAAATGATTACCTAAAGTTGAAAATTAAGGAA
AAATATTATCAAACGTTTGTAGAGTTGAGCAATTTATCTTTCTGACAGAAGAAGATAAATGTTGGTATAAAAAAAGAT
ACAATATTGATGAGGTTGTTGTTCTTAGCGAGCAATCATCTTTATGTAGAAAAGATTAAGAGAAGTAACTCAATTT
CCAAAACCTTTTCTGCTTATTCCCGGTAGCATTGAATTTTACAAAAATTTTACGGCTTAAATGGTTTTATAAAAA
ATATATCCTGGATTAATAGGAAAAATAAGAATAGTTGTAACAGGAAAGGCATCAGATAAAAAAATAAAGATGTT
45 AAAGTGTGGAGAGGAAATTACCTTTACGGGAGAGCTTACTTTTCCACATATAATAAAGTGTGCTCAACATGCTTG
TGTGTTATTGCACCGATTACAACGGGCACTGGAATTAATAAAAAATATTAGAAGCTGTACAAAAAGGTATTCTG
TACTTACAACAAAATTTGCTTCAAAGGAATATGTTCCGATTTATGTTTTTATTGCGAGGAGGATACTGACACAAA
CTTTGTCAATTTAATTAACAGTTTTCTTGAACGACATTAAGAGTCCAAGAATGAATTTATTGCTTTTTTTCAGTCTT
GCGTTTTGGTTAATATTGGCTTTGGCCATAATAAAAAAGTGGAGATTAACGCATACTTAATGTTTTTCTCGT
50 GGTCCATAATGTTAATATACAGGGCTGCGTATGAATGATAGTATTATATCGAATACAGGAAAATGATAATGAA
GTGCCTATTTTATGTGACTTTAGTCTCGCATCTATAAGAGATATACATGGGGAGGTAGGCTATCTATTCTTATCAT
CAATCTTTAAACTTTTATGCTTGGCATTTCATTTCTTTTTTTTTATTGCTTTTTTATCACTCCTGCTTACATATTT
TTCATTACAGAAAAATAAGTTAATACCGATACTATCGTTAGTTTTTTATTTAAGCCATGCTTTTATAGTTAGAGATTT
GATTCAAATTAGGGCAGGATTAGCTGTTAGCATATCATTATATTCAATAATTAATTTAAAGGAAATAAAGTATAA
55 TTACAGGAGTTTTATTTGCTTCTTTGATTCACTTGGGGCGCTTATTATTGCTCTTTGTTATCCTTTTTTCAAAAAA
AATACATAACATTAATAATGATGTTGTTTTTATTTAGTGTCAATTTATTTTTCTTATTGATGGGCTTAATTTTATC
GATACAACTCTTATCAATATAGTTTGTCTTCAACTGCAATTTGCAATTAATGTTGGTTGGGAAAGATATGATTATC
GGGTGAGTATATTTACTAATCCGGTTTTTATTAAGGTGTTTTTTTTAATTGCTTAAATGCACAAATATGACTTTCA
GATATTAATAAATGAGAAAATTAGTGTCTTTATAACTTATATGTTTTAGGTGATTAGCTATGGTTGCATTGAGTGG
60 GATGGCTATTTCTTTCAGGCCGCTTTTCTGACACTAGGTGAAAGCATTTTAATTGATATGCTCTGTTT
TACAAAAGAAATACACCTCTGGCGTTTCTAATTTTTTCTTTTTTAAACAATTGTGCAATTAGGATATGATCTATTTATT
TCTAATGTGCATCCTGAGCTTACTCTGATTATTTTTGGGTGAATCTAAGTGAATAAATAAATAAAGGATACTTAT
CTCTAAAATACAAAATCTTGGACCTGTGAATGTAGTACGAGGATTGATAAAAAGAAAATAAATAAATGCTTTTACT
GTTTTTTGTTTAAACAATAGCGTAGATAAAAAATATATATGATGAGTTATGCTGTTTAGGAGCCAAGGTTATATTAAT

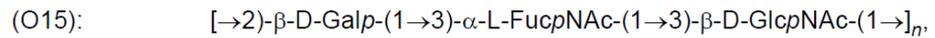
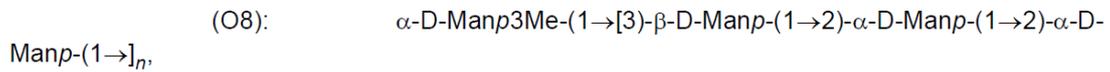
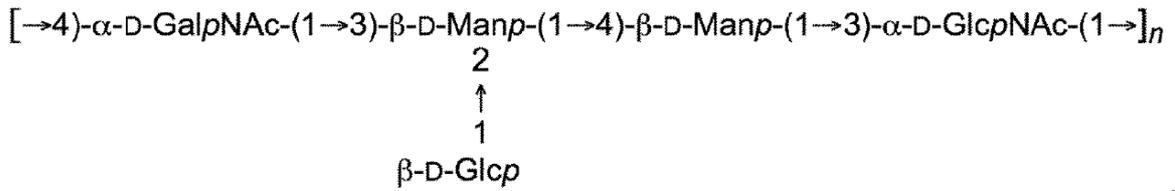
ACCAGATGGTACTTGGTTTCAGCAAAATTTTATTTGTGAGAAGTTTTTTAAAGGAACATCCACATAATATCTTACATT
CACATGGGATCACGGCCGATATGTTTTCTTACTTTCTGAATGGCGTGAAAATATCTACTATTCACAATAGACTAGA
TGAGGATTATATCCCATTTTGGCGCGGTTAAAGGGAATGCTATATATTATCTTCATCGTTTTATATTACGAAGAT
TTAATCATATCGTTGCTTGCAGCAGCGGTCCAATCAAAACTGAAACAATCGAAAGTAAAAACTAAAATAACCAC
5 CATCCAGAATGGGATTGATATAACTAGGTTTAAAGACACTTGAGTCTGATAAAAAAATTTATTGAGGAAAAACAC
GGATTTGATAGTGAAAAAAGAATATTTATATATTGTGGCTCGTTATCATTAAAGGAAAAATATTGCTTACCTCTTGG
ACACTTAGCCATCGAAGAAAAATGATATTTTTAATTTAGGTGATGGTGAACTTTTTAGATATTGTAAGGATAAAT
ATTCTAAAGATTTACGGTATATTTTTATGGGAAAAGTTGAATGCCCTCTTGAATATTATCAATTATCAGATTTTTT
10 GTTCCGCTTCTTTATCGGAAGGGCTCCCCTTGGCACTATTAGAAGCTGCCTCTACTGGGTGCTATTTATATGTT
AGCGATATAGAGCCCATAGAGAAATTGCATCTCTATTAGGAGAGGAAAATATTTCTATGTTTAAAATTAAGGATG
GATCATATAATTATTTGCAACCTAAAATAAAAAAAGCTGACTATAACGCTCTTCTGACGATAAACCCTTACAATATA
TCCGATAAAAAAATGTCAAACTTTTATGACAACTTTTTGTTCTTTATTAGAGCAGAGGCACTAATAATGATTT
ATGTTTCGGTAATTTCTCATGGTCATTTCAAACCTTAAAGGAATTAGGAGCAGTATCAAATTAATAATCACAGC
15 AGAATTAAGATTAAGATAAAGATAATTTAGGAGAGCGCAATTTGGATTTTTGCAAGAAAAAATAAATTAACCT
ATTTAAGGTCTAAAGAGAAAAAGGATTTGGAGAGAATAAATGAAGTTTTTCTCTATATCTCTTAACTTACT
AAGGAAGATTTTTTGTGGTTATGAATCCTGATATATATTGAGTGTCTGATCTATTAGATGTCGTAGATGAGT
GTGGTTCAGCGAATGTTAATCTAGCAACGATAAATTTATACAGGGATTTTGATAAAAAACATATGATAACTCAGT
AAGGAAATTTCCCTCGGCAATTGATTTTTTATGTCATTTTTATTTAAGAAAAATGACTGTGTAGTAAATAAGAACA
AAATAACGAAACCAACATATGTTGATTGGGCTGCAGGTTCTTTCTAATATTTAATGCCTTCTTTTATCAAACCTC
20 AACGGATTCACGAAAAAGTATTTTATGTATTGCGAAGATATTGATATATGTTGGCGAGCTAAAAACACTTCAATA
CTTCAGTTTTACTACTCCATGCTATGCAGCAATTCATTTGACAAATTTAAACAATCGTAGGATTTTTAGTAGACAT
TTCATTTGGCATATAAAAAAGTATTATCCTTTTTTATATAAAAAATGGTATGCTGCGTTCTAGTAAGTTGCTTTAA
TGCTAATATCTTTTAAAGAGGTGAGAATGATACCTGTTATTTTGGCTGGTGGTTCCGGAAGTCGCTTGTGGCCAC
TTTCACGAGAAAAGTCCCAAGCAGTTTTTAAAGTTGACTGGCAGTTTGACAATGTTGCAGTCAACATTGTCAC
25 GTCTTAATAATTTAAATGCTGATGATTCAATAGTTATATGCAACGAAGAGCATAGATTTATTGTTGCAGAACAATTA
AGAGAGTTAGGCAAACCTTCAAATAACATTATTCTTGAACCCAAAGGTCGTAATACAGCCCCTGCTATAACACTCG
CAGCATTAGCAGCAAAAAGAAAATTCGCTGATGAAGATCCATTGATTCTTATTTAGCTGCAGATCACAACATCCA
AGACGAACATGTTTTCTGTGAGGCAATTAATAAGGCGTCATCTTTAGCTAGTTATGGAAAACACTAGTACTTTTGGT
ATCGTTCCATTCAAACCTGAAACTGGGTATGGCTATATTCGTCGCGGTGATGAAGTGCCTGTAGATGAGCAGCAT
30 GCGGTGGCCTTTGAAGTGGCGCAGTTTGTGCAAAAACCGAATCTGGAACCGCGCAGGCCTATGTGGCAAGCG
GCGAATATTACTGGAACAGCGGTATGTTCTGTTCCGTGCCGGACGCTATCTCGAAGAAGTAAAAAGTATCGT
CCGGATATTCTCGATGCCTGTGAAAAAGCGATGAGCGCCGTCGATCCGGATCTCGATTTTATTCGTGTGGATGA
AGAGGCGTTTCTCGCTTGTCCGGAAGAGTCCGTGGATTACCGGTCATGGAATGCACGGCAGATGCCGTTGTG
GTGCCGATGGATGCGGGCTGGAGCGATGTCGTTCTGGTCTTATTATGGGAGATCAGCGCCCACACCGCCG
35 AGGCAACGTTTCCCACGGCGATGTGATTAATCACAAAACGAAAACAGCTATGTGTACGCCGAATCTGGCCTG
GTCACCACCGTCGGGGTGAAGATTTGGTGGTAGTGACAGCCAAAGATGCAGTGCCTGATTGCCGACCGTAATG
CGGTGCAGGATGTGAAGAAAAGTGGTGCAGCAGATCAAAGCTGATGGTCCGATGAGCATCGGGTGCATCGCGA
AGTGTATCGTCCGTGGGGCAAATATGACTCTATCGACGCGGGCGACCGCTACCAGGTGAAACGCATCACCGTG
AAACCGGGCGAAGTTTTGTGCGGTACAGATGCATTATCATCGCGCGGAACACTGGGTGGTTGTGCGGGGAACGG
40 CAAAAGTCACTATCAACGGTGATATCAAACCTGCTTGGTGAACCGAGTCCATTTATATTCCGCTGGGGGCGATGC
ACTGCTTGGAAAACCCGGGAAAATAGATTTAGAATTAATTGAAGTTCCGCTCTGGTGCATATCTTGAAGAAGATG
ATGTTATTAGATGTTATGATCGCTATGGACGAAAGTAAATATAATAATTAATTTAGAAATTAGAAAATGATAATATAA
GTTTTCGTGTGATAAACAATAGATAGTATGGGTTGAAAATATGAGTTCTTTAACTGTTTTAAAGCTTACGCAAT
TCGCGGGAATTTAGGTGAAGAACTGAATGAAGATATCGCCTGGCGCATTGGTTCGCGCCTATGGCGAATTTCTCA
45 AACCGAAAACCATTTGTTAGGCGGTGATGTCCGTCTCACCAGCGAAAACCTTAAAACCTGGCGCTGGCAAAGGT
TTACAGGATGCGGGCGTCGATGTGCTGGATATTGGCATGTCCGGCACCGAAGAGATTTATTTCCGCCAGTTCCA
TCTCGGCGTGGATGGCGGCATTGAAGTTACCGCCAGCCATAATCCGATGGATTACAACGGCATGAAGCTGGTG
CGCGAAGGGGCTCGCCGATCAGCGGTGATACCGGACTGCGCGACGTCCAGCGTCTGGCAGAAGCTAACGAC
TTTCTCCCGTCGATGAAACCAACCGGTCGCTATCAGCAAATCAATCTGCGTGACGCTTACGTTGATCACCTG
50 TTCGGTTATATCAATGTCAAAAACCTTACGCGCTCAAGCTGGTATCAACTCCGGGAATGGCGCAGCGGGTCC
GGTGGTGGACGCTATCGAAGCCCGCTTTAAAGCCCTCGGCGCACCGGTGGAGTTAATCAAAGTGCATAACACG
CCGGACGGCAATTTCCCAACGGTATTCTAACCCTGTTGCTGCCGGAATGTCGCGACGACACCCGCAATGCGG
TCATCAAACACGGCGCGGATATGGGCATTGCCTTTGATGGCGATTTTGACCGCTGTTTCTGTTTACGAAAAAG
GGCAGTTTATTGAGGGCTACTACATTGTGCGCCTGCTGGCAGAAGCGTTCCTCGAAAAAATCCCGCGCGGAAG
55 ATCATCCACGATCCACGTCTCTCCTGGAACACCATTGATGTGGTGGACGGCCGCGGGCGGCACGCCGGTGTATG
CGAAAACAGGACACGCCTTTATTAAGAAGATGTCGCAAGGAAGACGCCATCTACGTTGGCGAAATGAGCGCT
CACCATTACTCCGCTTTCCGCTTACTGTGACAGCGCATGATCCCGTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGG
CCTGAAAGGAAAAACGCTGGGCGAACTGGTGCAGCACCAGGATGGCGGCGTTTTCCGGCAAGCGGTGAGATCAA
CAGAAAACCTGGCGCACCTGTTGAGGCGATTAACCGCTGGAACAGCATTTTAGCCGTGAGGTGCTGGCGGTG
60 GATCGCACCGATGGCATCAGCATGACCTTGGCGACTGGCGCTTAACTGCGCTCTTCAAACACCGAACCAGGT
GGTGCAGCTGAATGTGGAATCTCGCGGTGATGTTGAGTTATGGTAATCCATACTCAAGAAATATTATCAATTTT
GACGTCATAAAGAATAAGCCCTGACAAGTTAGGGCTTAATTAATATATATTTTTTTGAATTGGGGATTTGTGGTA
AGATTTTTAATATGTTATTTAATGTGGTTGAATTAATGTTGACTGGAAAATAAATGAGAACGAAAAAAGCATTAC
ACAACTTAAAGTTGATTTATTAATTACTTTTTTATTGTTTTGCTAGGGTTTTATTCGAACTGTTTTTGTTCAA

AAATGGGAAGTGATATTACTGGAGTGATGTTACTATTCACACAGTTGACAGCATATCTCAATTTGGCAGAATTAGG
 TATTGGAATTGCAGCTGCCAGCGTATTATATAAACCGCTCAGCGAGAATGAATACAATAAAATAACTTACATAATA
 TCTTTGCTCTCAGTCATATACAAATATATATTTGTGTTTGTGTTTATTCTTGGCGTTGTTATAGGTATCTGTATTTAT
 5 TACTTTATTGATTCTGTAAAGGTTGTAATGGCGTTTTTTTATATTGGGCTTTTGTTCGTTTTTAATACATCGTTGAC
 ATATAGTTATGCTAAATACTCCACATTATTAAGTCTAATCAGCGGTAAGAAAAATTCAAGGTGGC
 GGAAAAGTTATAATAATTGTATTTTCAGATATTAATTTGTGCTTTACGCAAAGTTTCATACTTTATTTGTTAGTTGAG
 ACTTTAGGTATTTTTCTCAATATTTGATTTTTAAAAAATAATTGGGAACGGAAATCAATATCTCAGTAATGAGGT
 TTTACTTATTGAAAGCGATAAACTTTTGATAAAAAAAGAATTAATAAATAAGAATAAAAAATATGTTCTTCCATAAAAT
 10 AGGTGCTGTGCTTGTCTTAATACAGACTACCTGCTTGTATCAAAGTTTCTGACATTAAGTTATGTGACAATTTTT
 GGCAGCTATATGATGGTATTTTCAGATAGTAAGTGTGTTGATGTCAAGTTTTGTTAATGCTATTACTGCAGGAATGG
 GTAATTACTTAATTAATAAAAGTAATTTAGAAATTAAGGAAATTACACGTCAATTTTATGTGATATTTATCGCCTTTG
 CAACATTCATATCACTAAATATGTTTTTCTTGTAAATGATTTTATCGCAAAATGGATAGGTGTTAATTATACATTA
 GTAACACCCTAGTTGCATTAATGATTGTTAACGTATTCATTAGTGTGTCAGGGTACCTTCTGATATTTAAAAAAC
 15 GCAAGTGGACATTTTTGGTGATATTTATTATCCATTATTAGAAGGTGTGCTGAATATTACGATATCCATTTTTGG
 CTATCATTATTGGATTACCTGGCATTATTATAGGGACAATAGTATCTAACTTAATAGTAATAATGCTTGGCAAACCA
 TTATATCTTTACTCTAAGTTATTTAATCTTAGAAATCCGACGAGGGTTTTATTTTGAATTTATTTCTCGGCCTATGTTA
 TATTCATTATGTGTGATTGGGGTGAGCTATTTATTGCGCGATGAAATATATTCATTTAAAGTAAGTACATGGTTGG
 ATTTTATTAACAAGCTACTCTTAGTCTCTACTCCTAGCATATTGGTAATATGTGCTATTTTCTCTACGGATAGTGAC
 TTTAGATTATTTTTCAGAAAAATATATATGTGATTATGAAGAAATAAAAAATTTGAAAATGTATTAATCGAAATTAT
 20 GCAACGAGCTTTATTTTTATAAATGATATGTGATCTTTTCGCGAATAGGAGTAAGGATCCGTGTAGGCTGGAGCT
 GCTTCGAAGTTCCTATACTTTCTAGAGAATAGGAACCTTCGGAATAGGAACCTAAGGAGGATTCATATGGATAAA
 GCCGTAAGCATATAAGCATGGATAAGCTATTTATACTTTAATAAGTACTTTGTATACTTATTTGCGAACATTCCAGG
 CCGCGAGCATTACGCGCGGTGATCACACCTGACAGGAGTATGTAATGTCCAAGCAACAGATCGGCGTAGTCGG
 TATGGCAGTGATGGGACGCAACCTTGCCTCAACATCGAAAGCCGTGTTTATACCGTCTCTATTTTCAACCGTTC
 25 CCGTGAGAAGACGGAAGAAGTGATTGCCGAAAATCCAGGCAAGAACTGGTTTCTTACTATACGGTGAAAGAGT
 TTGTCGAATCTCTGGAAACGCCTCGTCGCATCCTGTTAATGGTGAAAGCAGGTGCAGGCACGGATGCTGCTATT
 GATTCCTCAAACCATATCTCGATAAAGGAGACATCATCATTGATGGTGGTAACACCTTCTTCCAGGACACTATTC
 GTCGTAATCGTGAGCTTTTCAGCAGAGGGCTTTAACTTCATCGGTACCGGTGTTTCTGGCGGTGAAGAGGGGGC
 GCTGAAAAGTCTTCTATTATGCCTGGTGCCAGAAAGAAGCCTATGAATTGGTAGCACCGATCCTGACCAAAT
 30 CGCCGCCGTAGCTGAAGACGGTGAACCATGCGTTACCTATATTGGTGCCGATGGCGCAGGTCACTATGTGAAG
 ATGGTTCACAACGGTATTGAATACGGCGATATGCAGCTGATTGCTGAAGCCTATTCTCTGCTTAAAGGTGGCCTG
 AACCTCACCAACGAAGAACTGGCGCAGACCTTTACCGAGTGGAATAACGGTGAAGTACCTGATCGA
 CATCACCAAAGATATCTTACCACAAAAGATGAAGACGGTAACTACCTGGTTGATGTGATCCTGGATGAAGCGGC
 TAACAAAGGTACCGTAAATGGACCAGCCAGAGCGCGCTGGATCTCGGCGAACCGCTGTGCTGATTACCGAG
 35 TCTGTGTTTGCACGTTATATCTTCTCTGAAAGATCAGCGTGTGGCCGATCTAAAGTTCTCTCTGGTCCGCAA
 GCACAGCCAGCAGGCGACAAGGCTGAGTTCATCGAAAAAGTTTCGTCTGCGCTGTATCTGGGCAAAATCGTTTT
 TTACGCCAGGGCTTCTCTCAGCTGCGTGTGCTGCGTCTGAAGAGTACAACCTGGGATCTGAACTACGGCGAAATCG
 CGAAGATTTTCCGTGCTGGCTGCATCATCCGTGCGCAGTTCCTGCAGAAAATCACCGATGCTTATGCCGAAAAAT
 40 CCACAGATCGCTAACCTGTTGCTGGCTCCGTACTTCAAGCAAATTGCCGATGACTACCAGCAGGCGCTGCGTGA
 TGTCGTTGCTTATGCAGTACAGAACGGTATTCCGGTTCCGACCTTCTCCGACGCGGTTGCCATTACGACAGCTA
 CCGTGCTGCTGTTCTGCCTGCGAACCTGATCCAGGCACAGCGTGACTATTTTGGTGCGCATACTTATAAGCGTA
 TTGATAAAGAAGGTGTGTTCCATACCGAATGGCTGGATTA

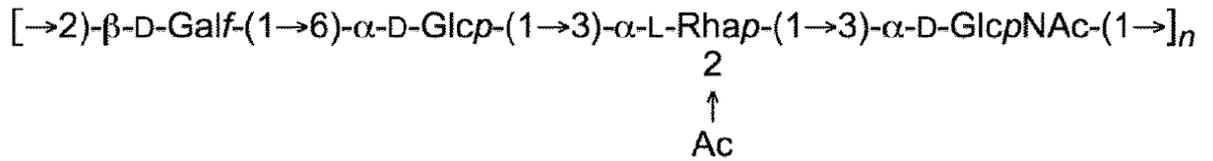
(O4-Glc+):



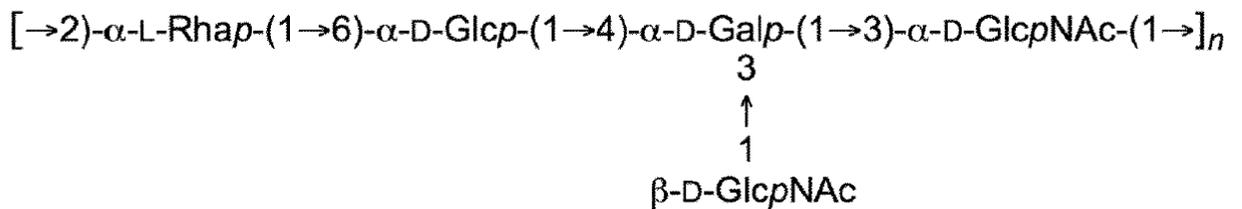
(O6A):



(O16):

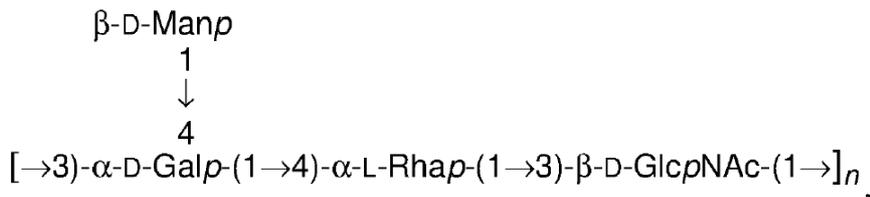


(O18A):



y

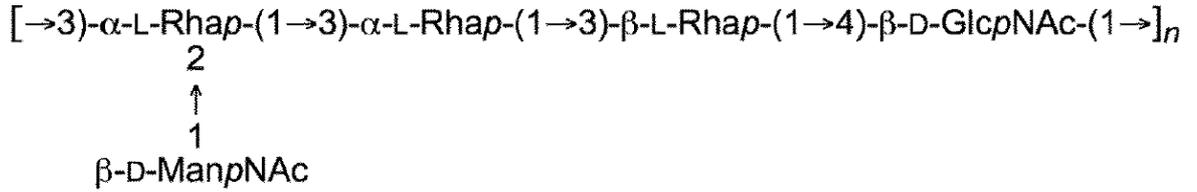
(O75):



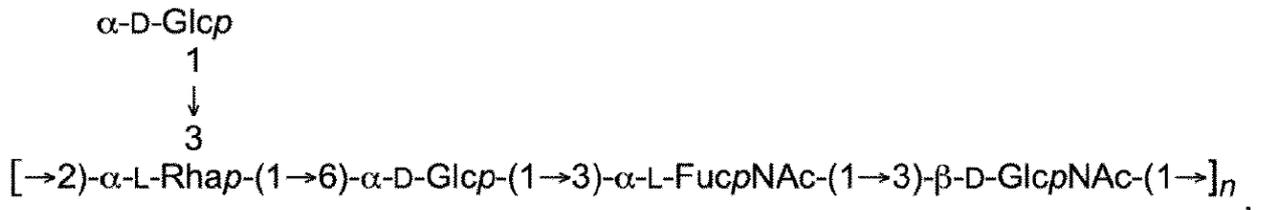
respectivamente y cada

n es independientemente un número entero de 1 a 100, preferiblemente de 3 a 50, por ejemplo 5 a 40, por ejemplo 7 a 25, por ejemplo 10 a 20.

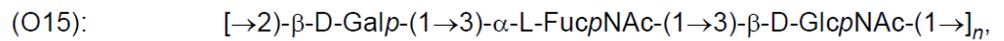
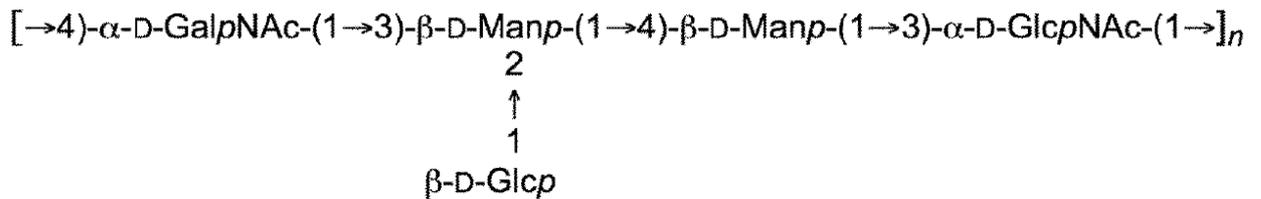
- 5 2. El método de la reivindicación 1, en donde el O_x-antígeno es O4 antígeno polisacárido glucosilado, y el PglB_y comprende la mutación de aminoácido N311V o las mutaciones de aminoácidos Y77H y N311V en relación con el PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6.
- 10 3. El método de la reivindicación 2, en donde la célula huésped procariota recombinante comprende además una secuencia que codifica un GtrS que tiene la secuencia de aminoácidos de SEQ ID NO: 4, y secuencias de nucleótidos que codifican un GtrA y un GtrB que tienen las secuencias de aminoácidos de SEQ ID NOs: 7 y 8, respectivamente.
4. Un método de preparación de un bioconjugado de un O_x antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora, comprendiendo el método:
- (i) proporcionar una célula huésped procariota recombinante que comprende:
- una secuencia de nucleótidos de un grupo de genes *rfb* para el O_x-antígeno polisacárido;
- 15 una secuencia de nucleótidos que codifica la proteína portadora que comprende al menos un sitio de glicosilación que comprende una secuencia de consenso de glicosilación Asn-X-Ser(Thr), en donde X es cualquier aminoácido excepto Pro (SEQ ID NO: 1), preferiblemente una secuencia consenso de glicosilación Asp(Glu)-X-Asn-Z-Ser(Thr), en donde X y Z se seleccionan independientemente de cualquier aminoácido excepto Pro (SEQ ID NO: 2); y
- una secuencia de nucleótidos que codifica una oligosacaril transferasa PglB_y; y
- 20 (ii) cultivar la célula huésped procariota recombinante en condiciones para la producción del bioconjugado,
- en donde el PglB_y comprende la mutación de aminoácido N311V con relación al PglB de tipo salvaje que tiene la secuencia de aminoácidos de SEQ ID NO: 6,
- en donde el O_x-antígeno es O1A antígeno polisacárido, O4 antígeno polisacárido glucosilado, O6A antígeno polisacárido, O15 antígeno polisacárido, O16 antígeno polisacárido u O75 antígeno polisacárido,
- 25 y cuando el O_x-antígeno es O4 antígeno polisacárido glucosilado, la célula huésped procariota recombinante comprende además una secuencia que codifica una glucosiltransferasa GtrS que tiene al menos 80% de identidad con SEQ ID NO: 4 y que es capaz de modificar un O4 antígeno polisacárido de *E. coli* mediante la adición de glucosa para producir el O4 antígeno polisacárido glucosilado de *E. coli*, y secuencias de nucleótidos que codifican una translocasa GtrA y una glicosiltransferasa GtrB que tienen al menos 80% de identidad de secuencia con SEQ ID NOs:
- 30 7 y 8, respectivamente, en donde la translocasa es capaz de translocar glucosa unida a bactoprenol y la glicosiltransferasa es capaz de glucosilar bactoprenol, y
- en donde los O1A, O4 glucosilado, O6A, O15, O16 y O75 antígenos polisacáridos tienen las estructuras de las Fórmulas (O1A),



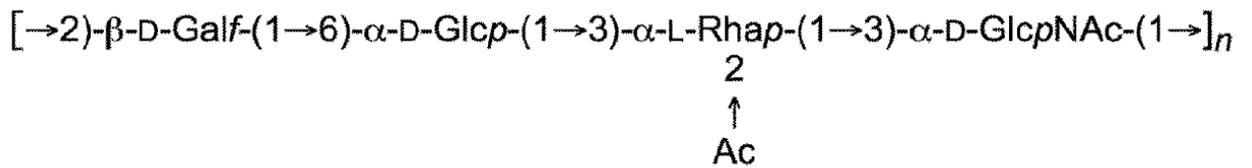
(O4-Glc+):



(O6A):

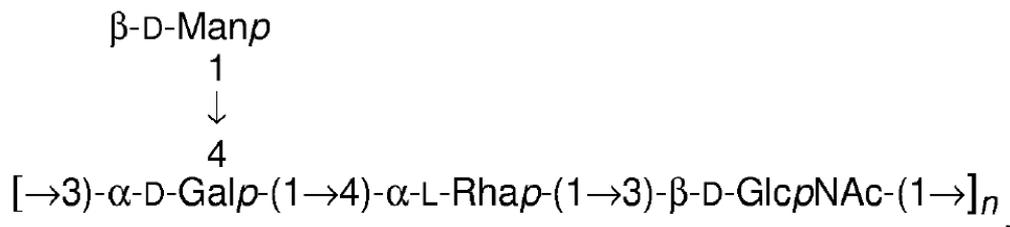


(O16):



y

(O75):



respectivamente, y cada n es independientemente un número entero de 1 a 100, preferiblemente de 3 a 50, por ejemplo 5 a 40, por ejemplo 7 a 25, por ejemplo 10 a 20.

5. El método de una cualquiera de las reivindicaciones 1 a 4, que comprende además aislar el bioconjugado de la célula huésped procarionta recombinante.
6. El método de una cualquiera de las reivindicaciones 1 a 5, en donde la proteína portadora se selecciona del grupo que consiste en Exotoxina A desintoxicada de *P. aeruginosa* (EPA), flagelina de *E. coli* (FliC), CRM197, proteína de unión a maltosa (MBP), toxoide diftérico, toxoide tetánico, hemolisina A desintoxicada de *S. aureus*, factor de aglutinación A, factor de aglutinación B, enterotoxina termolábil de *E. coli*, variantes desintoxicadas de enterotoxina termolábil de *E. coli*, toxina del cólera subunidad B (CTB), toxina del cólera, variantes desintoxicadas de la toxina del cólera, proteína Sat de *E. coli*, el dominio pasajero de la proteína Sat de *E. coli*, neumolisina de *Streptococcus pneumoniae*, hemocianina de lapa californiana (KLH), PcrV de *P. aeruginosa*, proteína de membrana externa de *Neisseria meningitidis* (OMPC) y proteína D de *Haemophilus influenzae* no tipificable,
7. El método de una cualquiera de las reivindicaciones 1 a 6, en donde la proteína portadora es Exotoxina A desintoxicada de *Pseudomonas aeruginosa* (EPA), preferiblemente en donde la proteína portadora EPA comprende 1–10, preferiblemente 2–4, más preferiblemente 4, de los sitios de glicosilación, preferiblemente en donde cada sitio de glicosilación comprende una secuencia de consenso de glicosilación que tiene la SEQ ID NO: 2, preferiblemente en donde la proteína portadora EPA comprende la SEQ ID NO: 3.
8. El método de una cualquiera de las reivindicaciones 1 a 7, en donde la proteína portadora es exotoxina A desintoxicada de *Pseudomonas aeruginosa* (EPA), en donde la proteína portadora EPA comprende 2-4 sitios de glicosilación, en donde cada sitio de glicosilación comprende una secuencia consenso de glicosilación que tiene SEQ ID NO: 2.
9. El método de una cualquiera de las reivindicaciones 1-8 en donde la célula huésped procarionta recombinante es una célula de *E. coli*, por ejemplo una cepa K-12 de *E. coli*, tal como la cepa W3110.
10. Una célula huésped procarionta recombinante para la preparación de un O_x antígeno polisacárido de *E. coli* unido covalentemente a una proteína portadora, comprendiendo la célula huésped procarionta recombinante:
- una secuencia de nucleótidos de un grupo de genes *rfb* para el O_x-antígeno polisacárido;
- una secuencia de nucleótidos que codifica la proteína portadora que comprende al menos un sitio de glicosilación que comprende una secuencia consenso de glicosilación Asn-X-Ser(Thr), en donde X es cualquier aminoácido excepto Pro (SEQ ID NO: 1), preferiblemente una secuencia consenso de glicosilación Asp(Glu)-X-Asn-Z-Ser(Thr), en donde X y Z se seleccionan independientemente de cualquier aminoácido excepto Pro (SEQ ID NO: 2); y
- una secuencia de nucleótidos que codifica una oligosacaril transferasa PglB_y
- en donde:
- cuando el O_x- antígeno es O1A antígeno polisacárido, el PglB_y comprende las mutaciones de aminoácidos N311V, K482R, D483H y A669V;
- cuando el O_x-antígeno es O4 antígeno polisacárido glucosilado, el PglB_y comprende la mutación de aminoácido N311V o las mutaciones de aminoácidos Y77H y N311V, y la célula huésped recombinante comprende además una secuencia que codifica una glucosiltransferasa GtrS que tiene al menos 80% de identidad con SEQ ID NO: 4 y que es capaz de modificar un O4 antígeno polisacárido de *E. coli* mediante la adición de glucosa para producir el O4 antígeno polisacárido glucosilado de *E. coli*, y secuencias de nucleótidos que codifican una translocasa GtrA y una glicosiltransferasa GtrB que tienen al menos 80% de identidad de secuencia con SEQ ID NOs: 7 y 8, respectivamente, en donde la translocasa es capaz de translocar glucosa unida a bactoprenol y la glicosiltransferasa es capaz de glucosilar bactoprenol;
- cuando el O_x-antígeno es O6A antígeno polisacárido, el PglB_y comprende las mutaciones de aminoácidos N311V, K482R, D483H y A669V;
- cuando el O_x-antígeno es O8 antígeno polisacárido, el PglB_y no comprende mutaciones de aminoácidos en las posiciones 77, 80, 287, 289, 311, 482, 483 y 669;
- cuando el O_x-antígeno es O15 antígeno polisacárido, el PglB_y comprende las mutaciones de aminoácidos N311V, K482R, D483H y A669V;
- cuando el O_x-antígeno es O16 antígeno polisacárido, el PglB_y comprende las mutaciones de aminoácidos Y77H, S80R, Q287P, K289R y N311V;
- cuando el O_x-antígeno es O18A antígeno polisacárido, el PglB_y no comprende mutaciones de aminoácidos en las posiciones 77, 80, 287, 289, 311, 482, 483 y 669; y

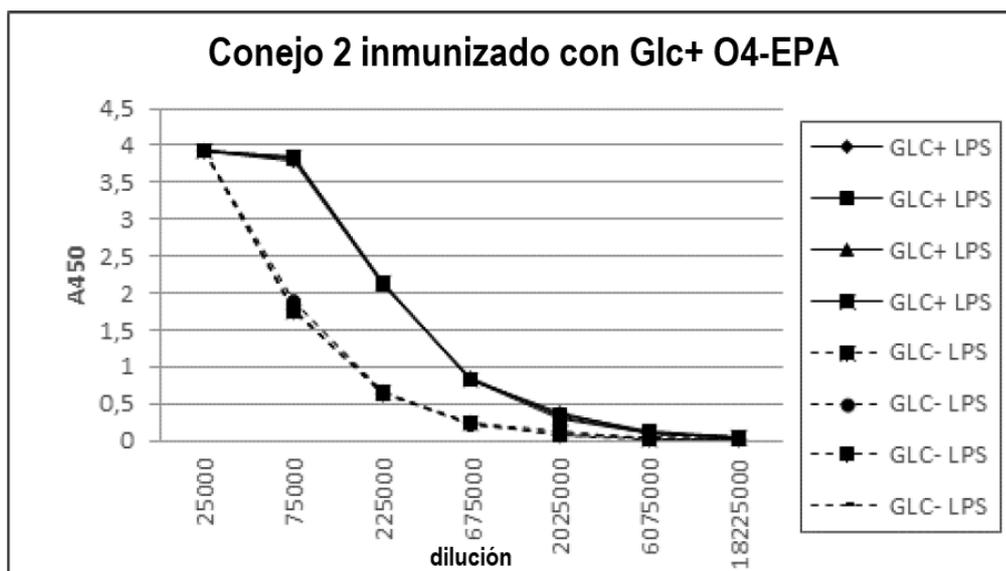
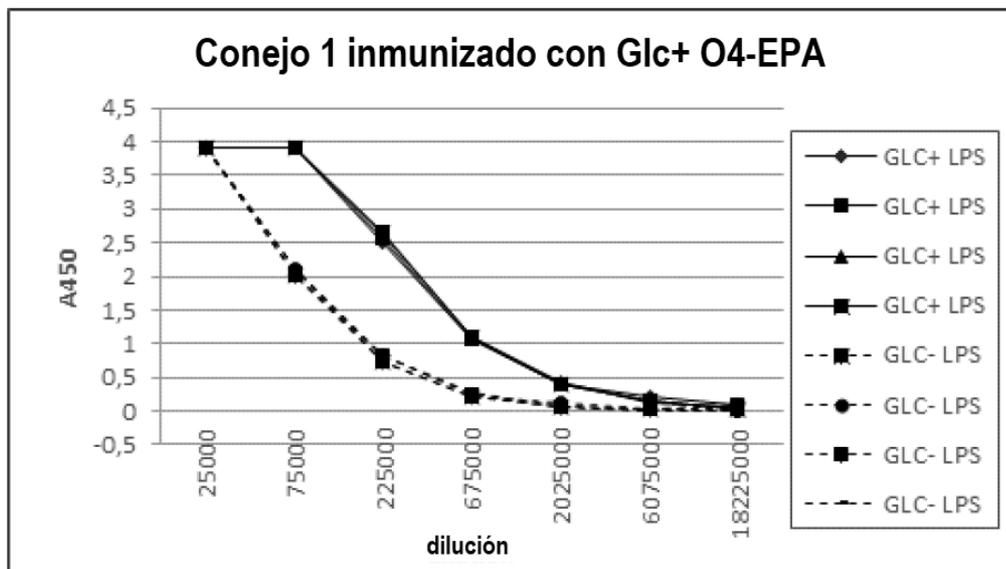


Fig. 1

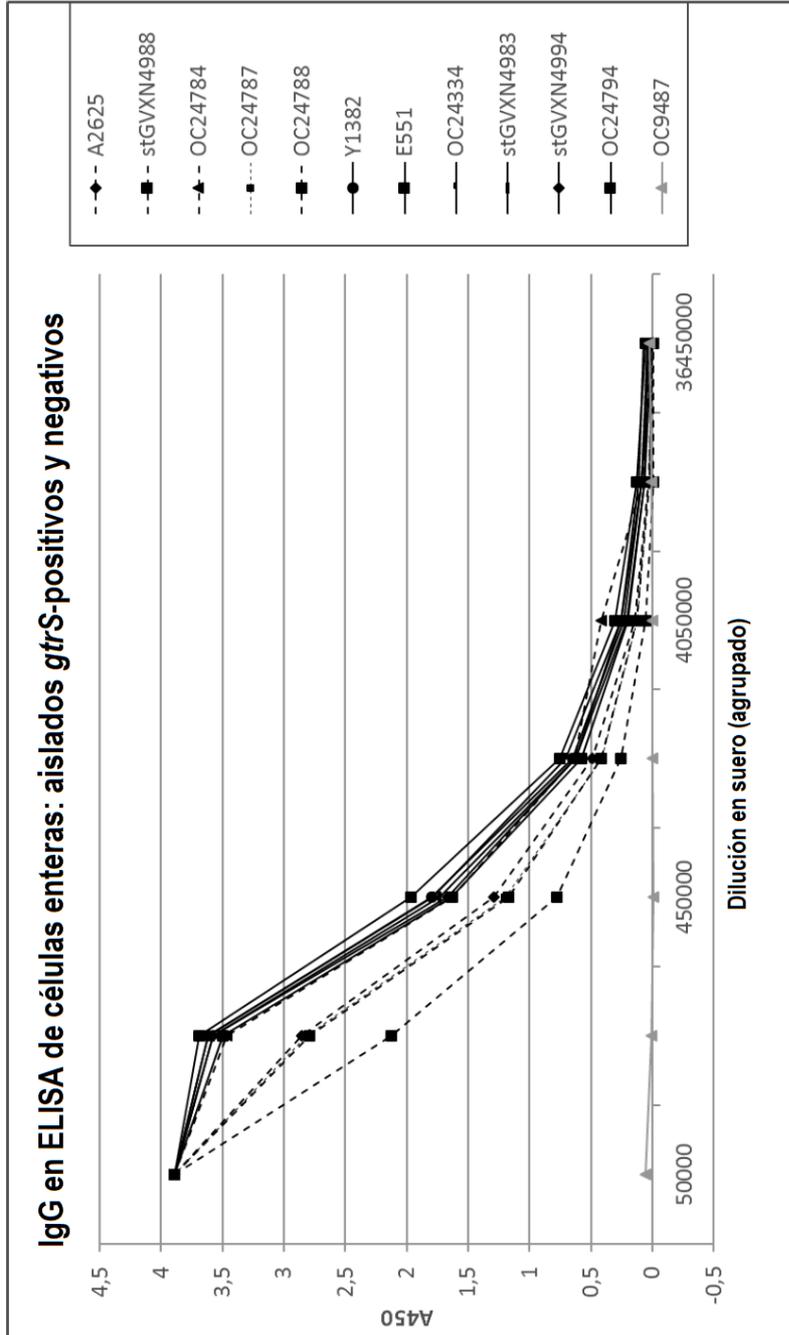


Fig. 2

Fig. 3

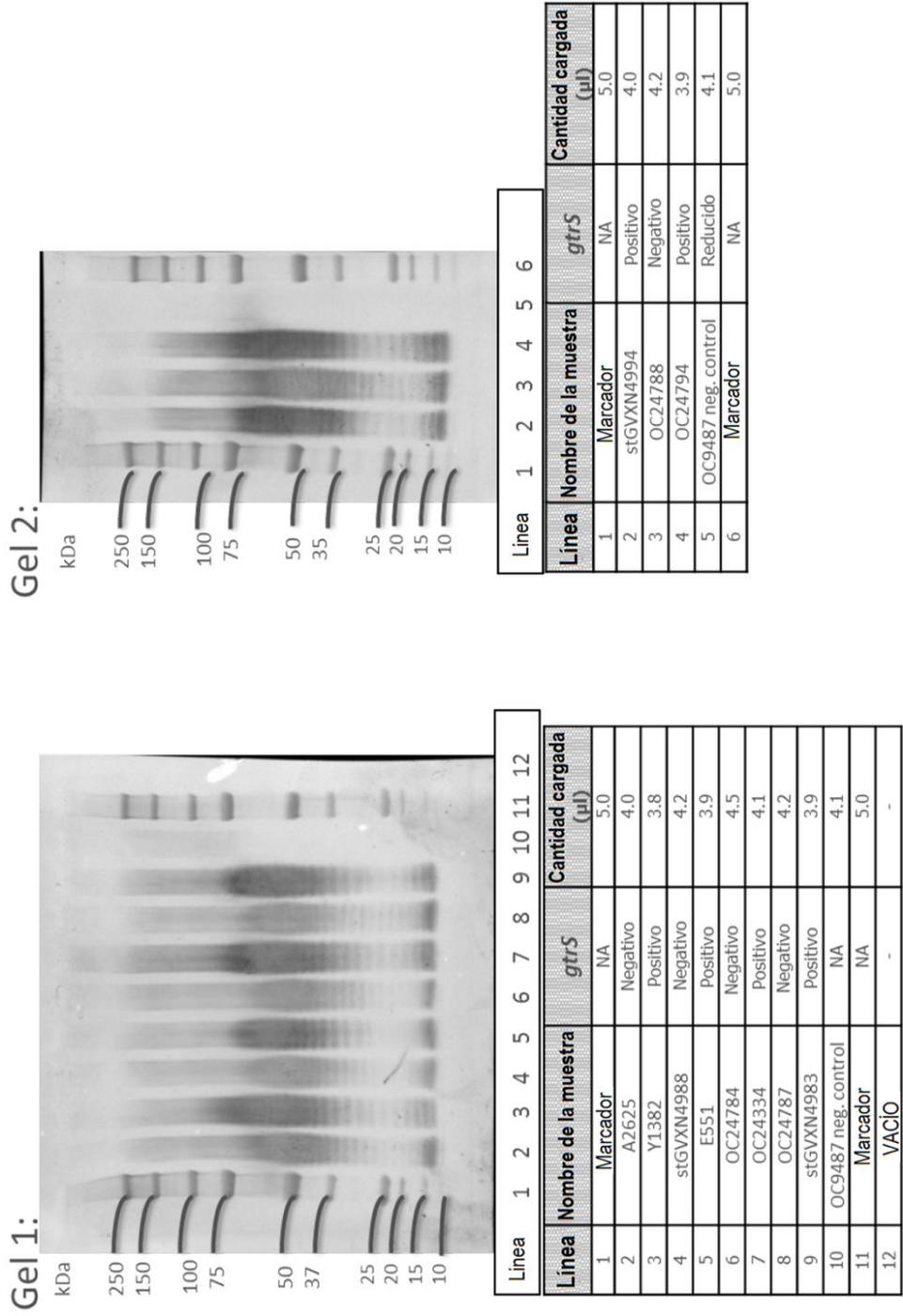


Fig. 4A

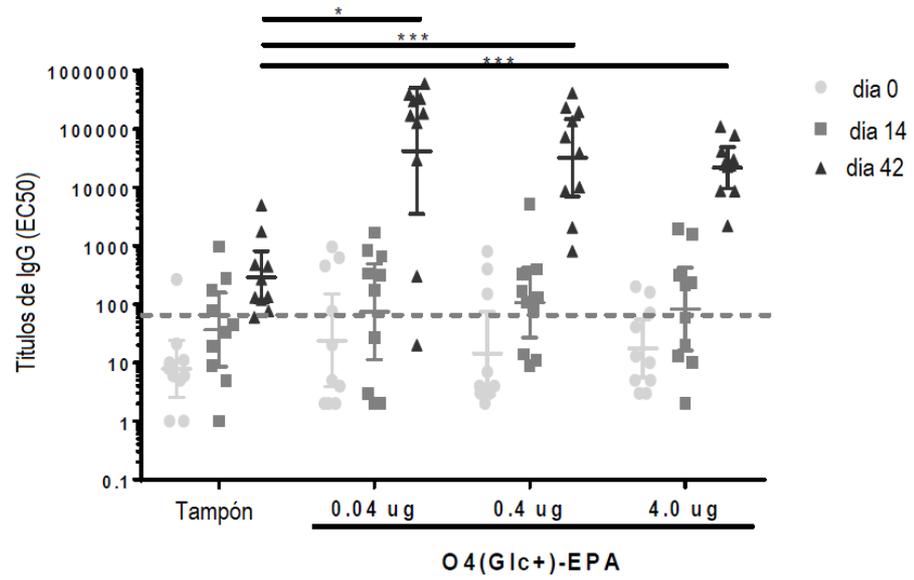


Fig. 4B

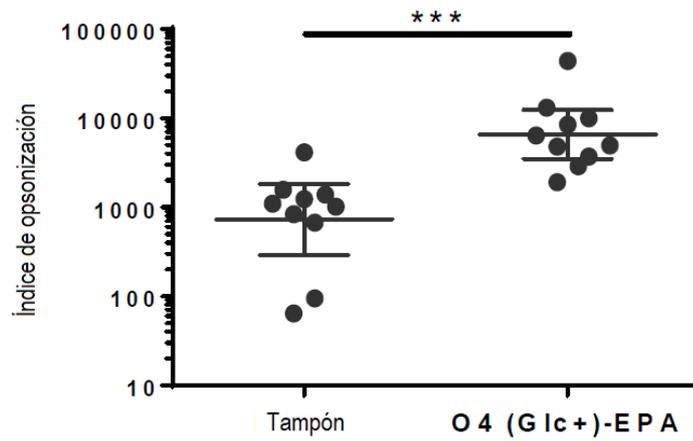


Fig. 5

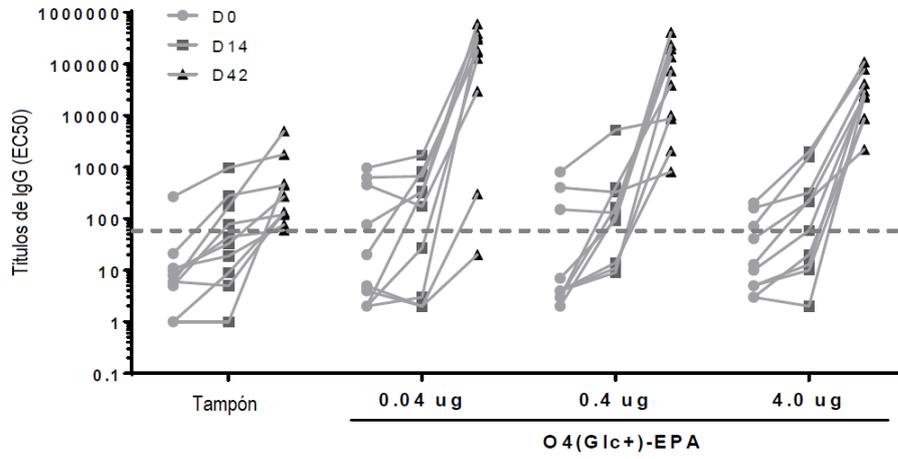


Fig. 6

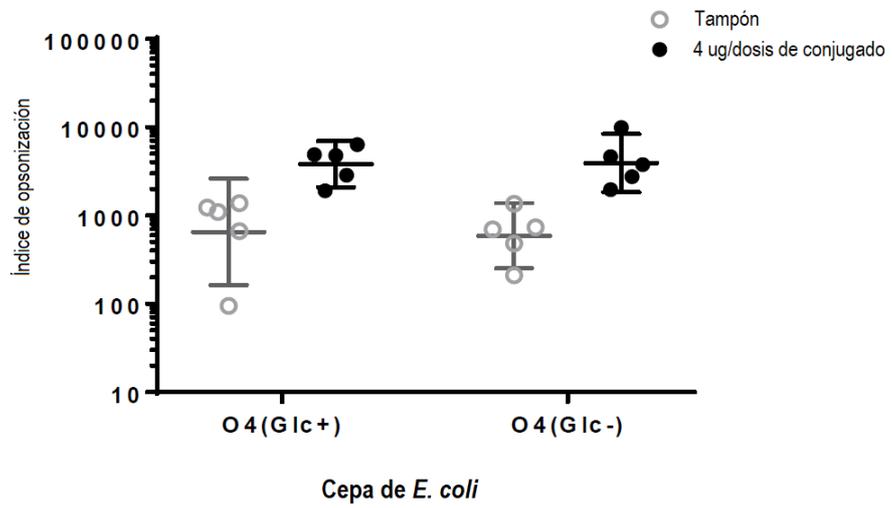
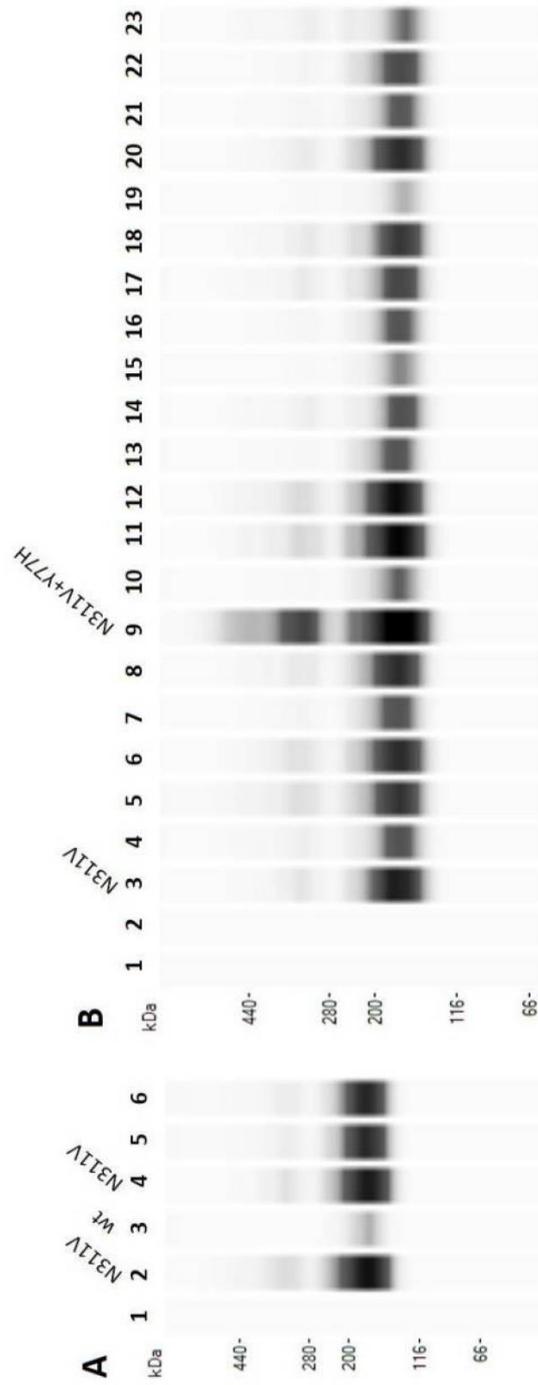


Fig. 7



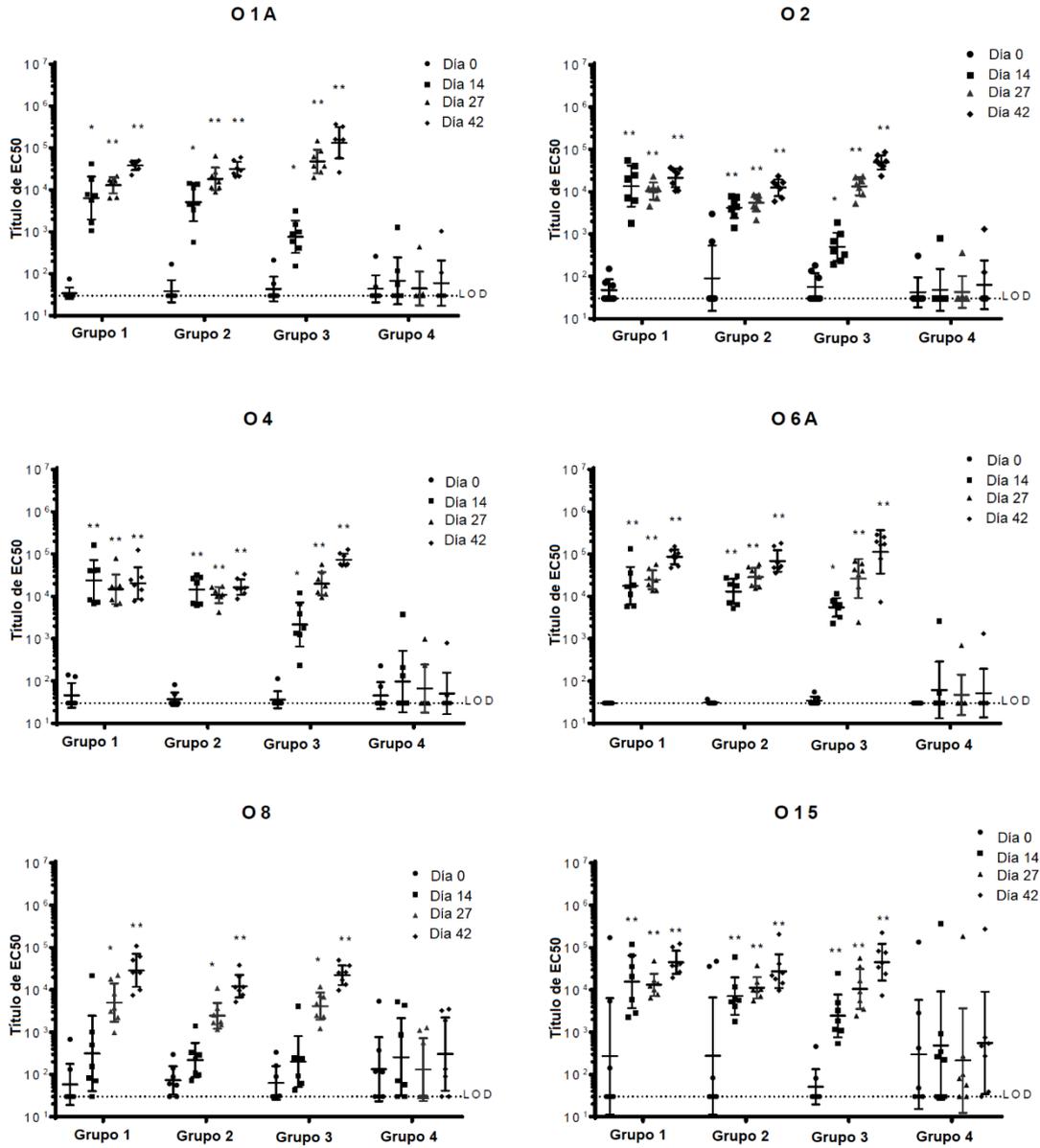


Fig. 8

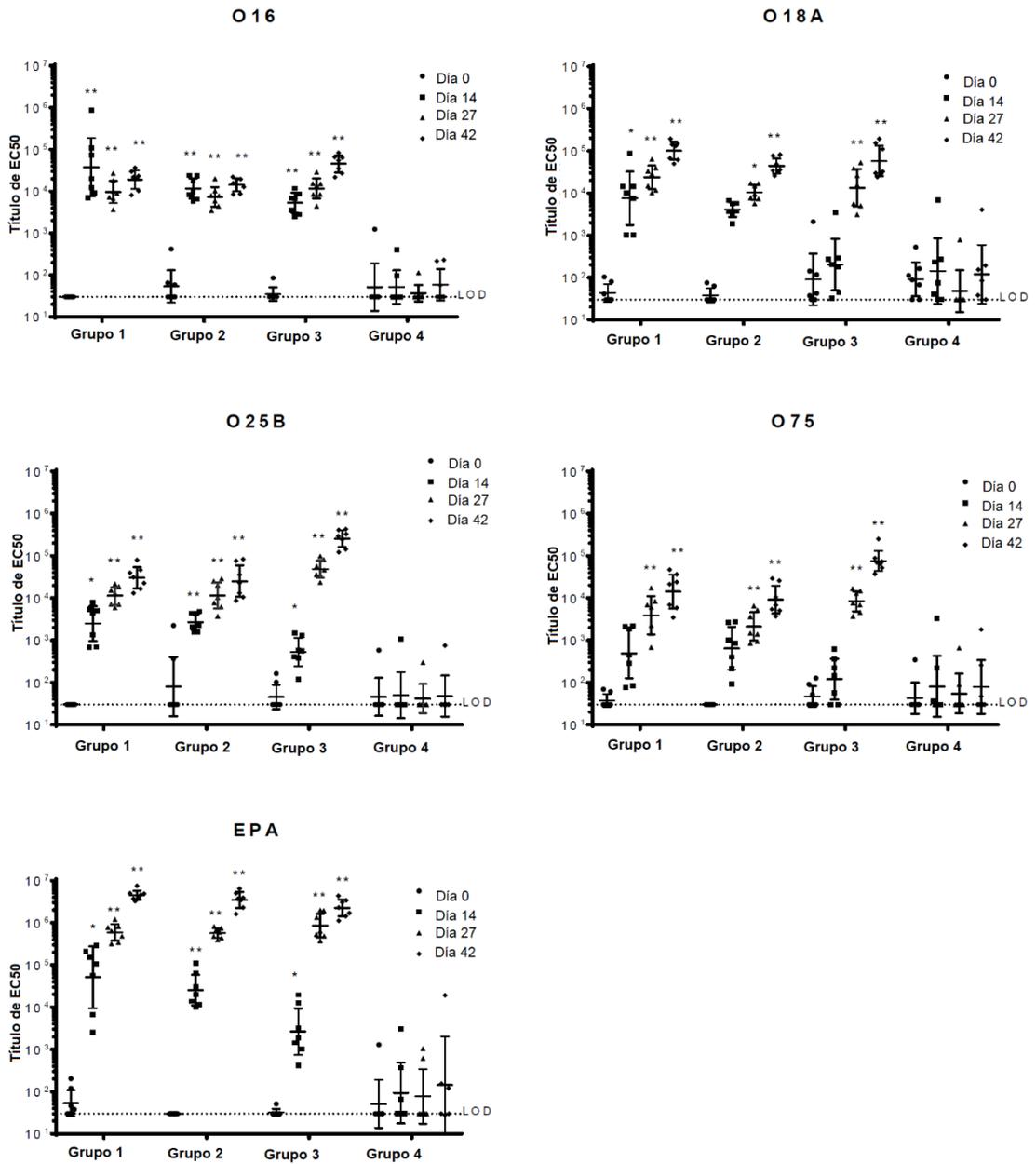


Fig. 8 - continuación

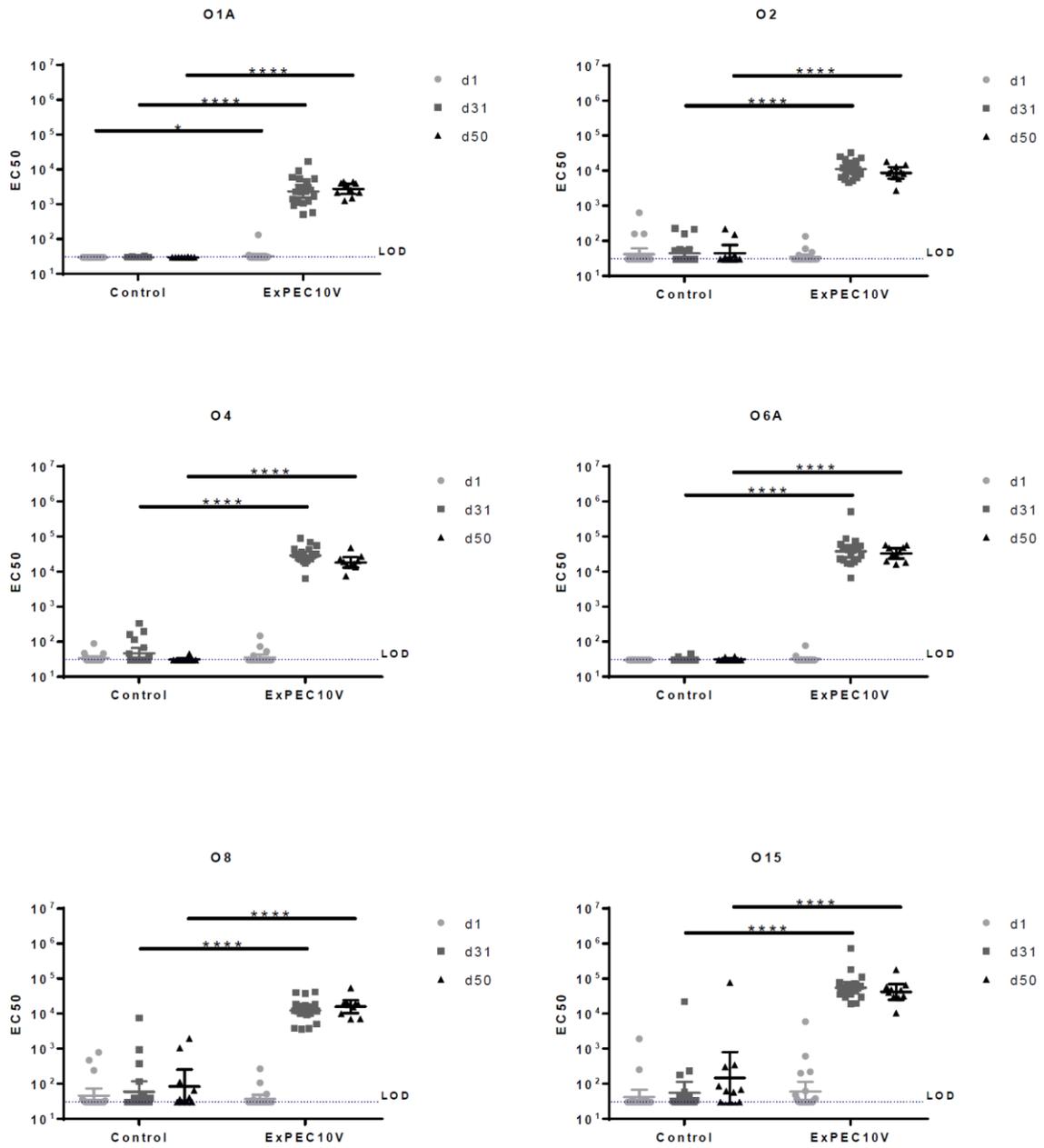


Fig. 9

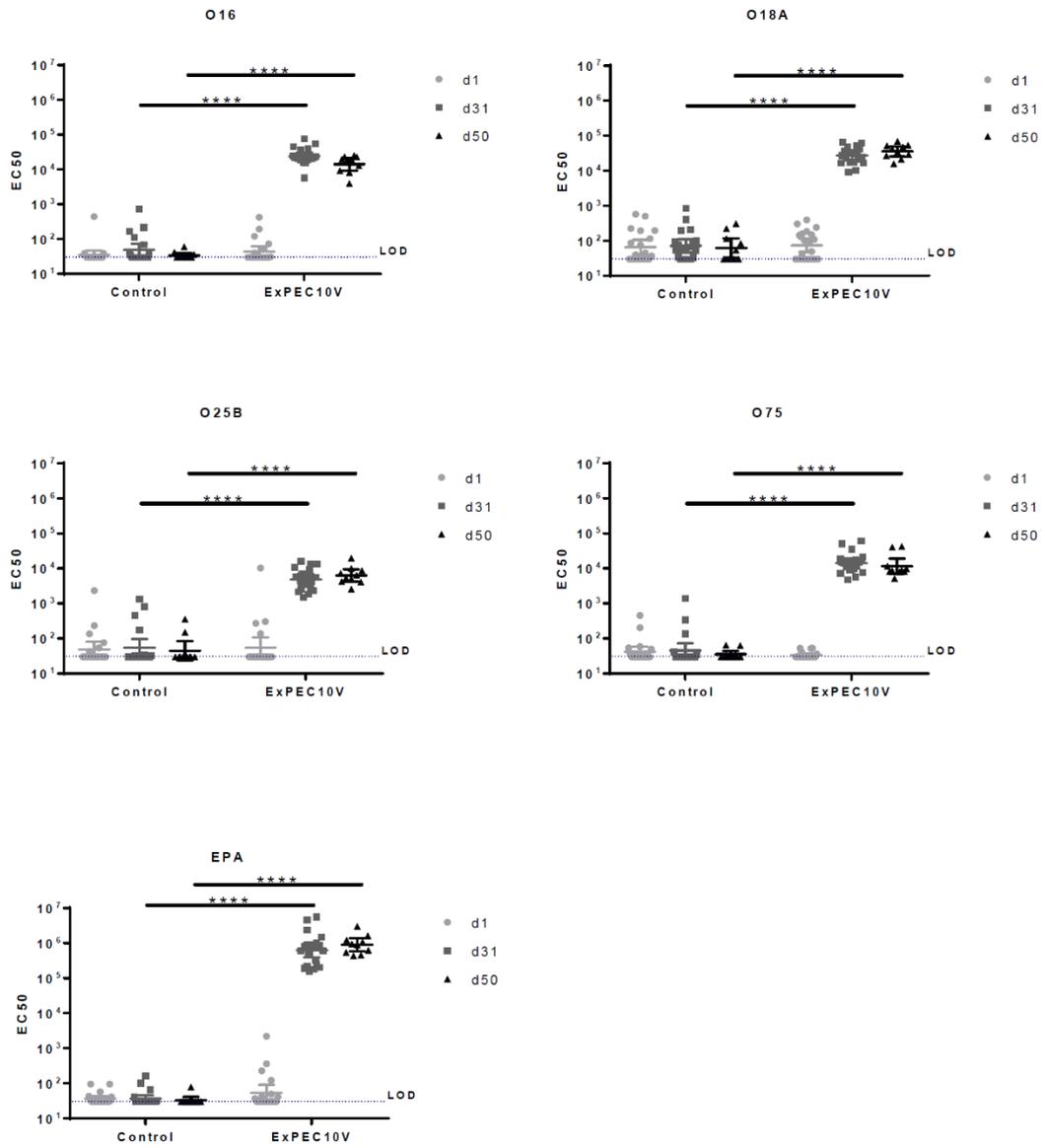


Fig. 9 - continuación

Fig. 10A

Fase 1 Dosis baja Fase 1 Dosis med. Fase 1 Dosis alta Fase 2a

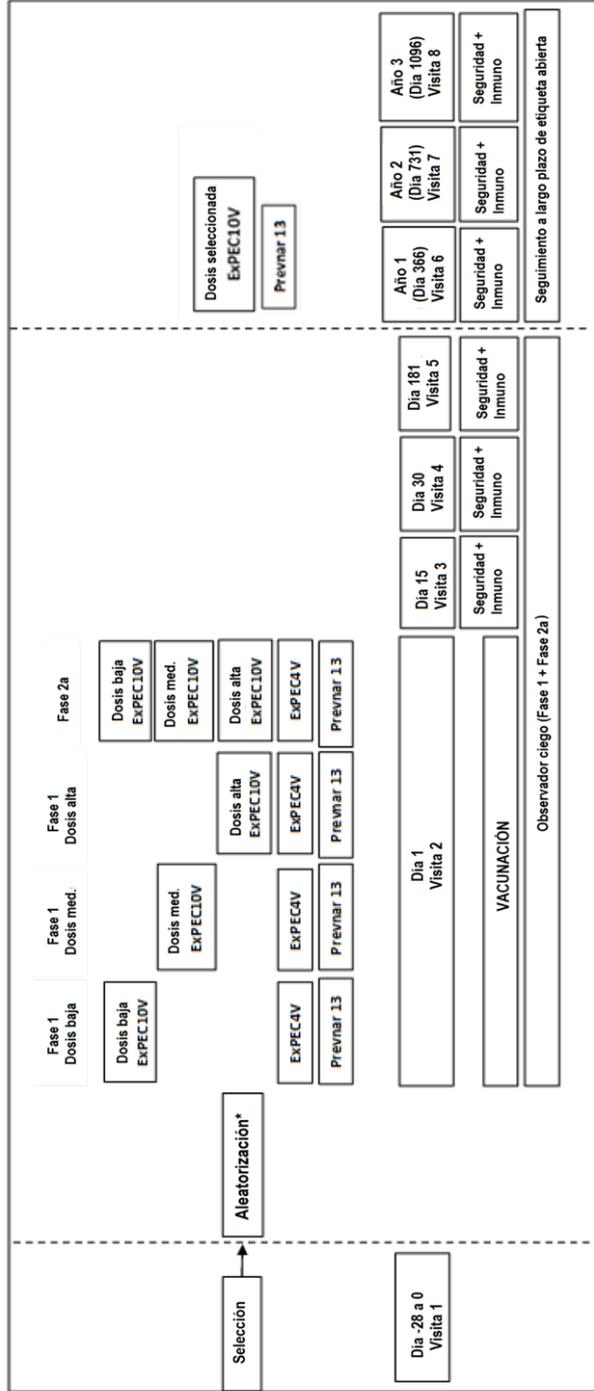


Fig. 10B

