

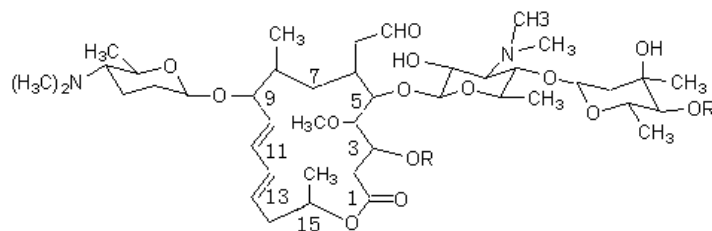
AGRUPACIÓN DE GENES BIOSINTÉTICOS DE CARRIMICINA

CAMPO TÉCNICO:

La presente descripción pertenece a los campos de recursos génicos microbianos y de la ingeniería génica, específicamente se relaciona con análisis de clones, análisis e investigación funcional de agrupaciones de genes biosintéticos de antibióticos en ingeniería génica y una aplicación de las agrupaciones de genes biosintéticos de antibióticos.

ANTECEDENTES:

La carrimicina ha usado los nombres, por ejemplo, shengjimicina y bioteccespiramicina, es un antibiótico macrólido de 16 elementos desarrollado mediante el uso de una tecnología de biología sintética [descrita en las patentes núms. ZL971044406 y ZL021487715]. La carrimicina es una espiramicina con 4"-hidroxilo multiacetilado, que acepta 4"-isovaleril espiramicina III, II y I como principales ingredientes, en donde el ingrediente III representa aproximadamente 30 % o más, el ingrediente II representa aproximadamente 25 %, y el contenido del ingrediente I no excede de 10 %.



Isovaleril espiramicina III: $R=COCH_2CH_3$ $R'=COCH_2CH(CH_3)_2$

Isovaleril espiramicina II: $R=COCH_3$ $R'=COCH_2CH(CH_3)_2$

Isovaleril espiramicina I: $R=H$ $R'=COCH_2CH(CH_3)_2$

La carrimicina tiene actividad relativamente alta contra bacterias grampositivas y tiene actividad antibacterial contra bacterias resistentes a eritromicina, resistentes a antibióticos betalactámicos, contra bacillus influenzae, gonococcus, legionella, bacteroidesfragilis y clostridium perfringens. Particularmente, la carrimicina tiene actividad relativamente alta contra mycoplasma

pneumoniae, chlamydia trachomatis y chlamydia pneumoniae [Yu, Lanxiang, y otros, Sichuan Journal of Physiological Sciences; 1998, 20 (3), patente número: 2003101224209] y tiene mejores efectos posantibiotico y efectos antibióticos por debajo de la MIC. La carrimicina está libre de completa resistencia cruzada a medicamentos con medicamentos similares. Se ha demostrado por

5 investigaciones farmacocinéticas, que la carrimicina tiene más alta lipofilidad y tiene gran actividad antibacterial intracelular, gran velocidad de absorción en la administración oral, gran biodisponibilidad absoluta, gran penetrabilidad tisular, amplia distribución en los tejidos, prolongada vida media de eliminación y prolongado tiempo de retención in vivo, y la concentración tisular de las carrimicinas es más alta que la concentración plasmática [Sun, Liwen, y otros, Chinese

10 Pharmacological Bulletin 2000, 16 (6): 694-8; Zhong, Dafang, y otros, J chromatography B. 2003, 791: 45; Shi, Xiangguo, y otros, Asian Journal of Drug Metabolism and Pharmacokinetics. 2003, 3 (2): 134; Shi, Xiangguo, y otros, Chinese Chemical letter 2004, 15: 431; Shi, Xiangguo, y otros, Acta Pharmacologica Sinica, 2004, 25: 1396]. Se ha demostrado por resultados de investigaciones clínicas completas de tres etapas, farmacológicas y toxicológicas, que la carrimicina se usa para el tratamiento

15 de infecciones del tracto respiratorio, es definitiva en el efecto del tratamiento y de baja tasa de reacciones adversas, particularmente tiene poco daño al hígado y buena seguridad [Lin, Futian, y otros, Eighth nationwide antibiotic academic conference paper compilation 1997, p.167; Zhao, Chunyan, y otros, Chinese Journal of Antibiotics, 1998, 23 (4): 306; Sun, Tao, y otros, Chinese Journal of Antibiotics, 2001, 26 (1): 49-51]. La carrimicina es un producto fermentado directo de

20 bacteria diseñada por ingeniería genética, obtenido mediante el uso de la tecnología de recombinación génica. El proceso de preparación es simple y conveniente, y puede efectivamente evitar la contaminación química y ahorrar la fuente de energía. Las preparaciones orales de carrimicina son convenientes de tomar y solamente se requiere tomarla una vez al día, lo cual es útil para mejorar la observancia de los pacientes con los medicamentos, y es conveniente además para

25 ingresar a la serie básica de medicamentos del seguro médico.

La carrimicina es un producto fermentado de bacteria diseñado por ingeniería genética (*Streptomyces spiramyceticus* WSJ-1), que se obtiene al someter el gen de 4"-isovaleril transferasa de bacteria productora de carbomicina a expresión mediante clonaje en bacteria productora de

espiramicina (*Streptomyces spiramyceticus* F21). La bacteria productora de espiramicina (*Streptomyces spiramyceticus* F21) se aisló del suelo en el Condado Yongchang, Provincia de Gansu, China en 1982 por nuestro laboratorio. Las características morfológicas, características bioquímica fisiológicas, la composición química de la pared celular, la secuencia del gen de ARNr 16S y los niveles de la proteína del gen 5 de mantenimiento de la bacteria en el análisis del árbol filogenético, no tiene nada en común con la bacteria productora de espiramicina *Streptomyces ambofaciens* ATCC23877 informada con anterioridad y la de estreptomiceto informada. Así, la bacteria productora de espiramicina (*Streptomyces spiramyceticus* F21) es muy posiblemente una nueva especie de estreptomiceto [Dai, Jianlu, y otros, *Journal of Microbiology*, 2012, 39 (4): 503-514].

La secuenciación de agrupaciones de genes relacionados con la biosíntesis de espiramicina en la bacteria productora de espiramicina *Streptomyces ambofaciens* ATCC23877 se ha completado [Karray F., *Microbiology*, 2007, 153: 4111-4122], y las secuencias de agrupaciones de genes biosintéticos de otros antibióticos macrólidos tales como avermectina, canavaliamicina, eritromicina, chalcemicina, tilosina y medemicina también se han informado [Ikeda H. y otros, *Nat. Biotechnol.* 2003, 21(5): 526-531, Haydock y otros, *Microbiology*, 2005, 151, 3161-3169; Oliynyk M. y otros, *Nat. Biotechnol.* 2007, 25(4): 447-453; Wards L. y otros, *Antimicrob. Agents & Chemotherapy*, 2004, 48(12): 4703-4712; Cundiffe E. y otros, *Antonie Van Leeuwenhoek*, 2001, 79(3-4): 229-234; Midoh Naoki y otros, patente de Estados Unidos núm. 7070980]. Las agrupaciones de genes biosintéticos de antibióticos macrólidos tienen una longitud completa de aproximadamente 50-80 kb y tienen las características comunes de estar compuestos de policétido sintasa (PKS) para codificar una estructura modular biosintética de un anillo de macrólido de 16 elementos, enzimas relacionadas con la unidad de extensión por síntesis de policetonas, enzimas responsables por la modificación de diferentes grupos radicales de un anillo de lactona, genes de enzimas relacionadas con la síntesis y transferencia de glicosilo y genes relacionados con la función de control, resistencia y regulación, etc. Los macrólidos se forman al llevar a cabo una reacción de condensación continua para catalizar algunas moléculas de ácidos carboxílicos simples mediante la PKS compuesta de estructuras modulares de manera similar a la biosíntesis de ácidos grasos. Cada módulo es responsable solamente por la reacción de condensación de una sola etapa en un proceso de formación de una cadena de policetona,

y el módulo contiene al menos un dominio estructural de beta-cetoacil sintetasa (KS), un dominio estructural de aciltransferasa (AT) y un dominio estructural de proteína portadora de acilos (ACP). Adicionalmente, el módulo posiblemente contiene además un dominio estructural de beta-cetoacil reductasa (KR), un dominio estructural de deshidrogenasa (DH) y un dominio estructural de acilester reductasa (ER), y los dominios estructurales deciden una etapa de reducción de unidades de extensión añadidas. Entretanto, se requiere, además, la acción de un dominio estructural de tioesterasa (TE) para catalizar la ciclización y liberación de cadenas de policetonas. Finalmente, se requiere, además, que se lleven a cabo etapas de modificación tales como hidroxilación, metilación, metoxilación y acilación para formar varias estructuras de antibióticos macrólidos. Generalmente, todos los macrólidos se conectan con grupos glicosilo (o glicosilamino) de cantidades diferentes, por ejemplo, la carrimicina contiene tres grupos glicosilo, es decir, forosamina, micaminosa y micarosa. Los grupos glicosilo son aceptados por enzimas relacionadas con la síntesis y transferencia de glicosilo. Los genes de resistencia dotan a las bacterias productoras con la capacidad para resistir a los antibióticos producidos por ellas mismas y, generalmente, se relacionan con la proteína de transporte ABC. Los genes relacionados con la función de regulación y control participan en la regulación y control de antibióticos biosintéticos en sí mismos.

Por medio de la información de secuencias de la agrupación de genes y el análisis estructural, pueden realizarse, además manipulaciones genéticas en bacterias productoras para obtener nuevos antibióticos y más efectivos. Por ejemplo, se crean nuevos antibióticos macrólidos mediante el cambio de las estructuras modulares de síntesis de PKS de antibióticos macrólidos por manipulación genética, mediante cambio del anillo de lactona después de la modificación, y sustitución o modificación de grupos glicosilo. Y la producción de antibióticos puede aumentarse al llevar a cabo operación genética en genes de resistencia o genes reguladores. [Wilkinson B. y otros, Chem Biol. 2000, 7 (2): 111-117; Kalz L. y otros, Med Res Rev. 1999, 19 (6): 543-58; Goodman CD y otros, Antimicrobial Agents and Chemotherapy, 2013, 57(2): 907-913; Wang W y otros, ProcNatlAcadSci Estados Unidos, 2014, 111(15): 5688-93; Stratigopoulos G y otros, Mol Microbiol. 2004, 54(5): 1326-34; Novakova R y otros, Folia Microbiol. 2011, 56(3): 276-82].

SUMARIO

La presente descripción proporciona una agrupación de genes de ligamiento biosintéticos de carrimicina. La agrupación de genes tiene 44 marcos de lectura abiertos (orf) de genes en general, la longitud completa de la secuencia de nucleótidos es de 89315 pb (sec. con núm. de ident. 1). La agrupación de genes contiene 5 orf que codifican la policétido sintasa (orf10-14), que comprenden 8 módulos y 37 dominios estructurales, 9 orf relacionados con la unidad de extensión por síntesis de policetonas y modificación (1, 4-6, 15 y 36-39), 16 orf relacionados con la síntesis de glicosilo (9, 16-22, 24, 26, 28, 29, 33-35 y 41) y 6 orf relacionados con la transferencia de glicosilo (7, 8, 30-32 y 40). Adicionalmente, la agrupación de genes comprende además 2 orf relacionados con la resistencia (3 y 25) y 4 orf posiblemente relacionados con la regulación y control (2, 23, 27 y 42). Las secuencias de nucleótidos se seleccionan por separado de un grupo que consiste en orf1(1-645), orf2(1810-1208), orf3(3133-2285), orf4(3614-4840), orf5(4846-5511), orf6(7150-5801), orf7(8444-7179), orf8(9729-8482), orf9(10543-9830), orf10(16215-10543), orf11(21076-16328), orf12(32511-21124), orf13(38599-32585), orf14(52259-38643), orf15(53099-54310), orf16(54495-54845), orf17(54842-56041), orf18(56038-56946), orf19(56930-57967), orf20(57937-60174), orf21(60836-61984), orf22(62796-62077), orf23(63633-65645), orf24(67379-66318), orf25(69004-67352), orf26(69349-70650), orf27(72156-70708), orf28(72422-73462), orf29(74601-73561), orf30(74913-76160), orf31(76218-77486), orf32(77606-78781), orf33(78783-79775), orf34(79772-80779), orf35(82055-80823), orf36(83164-82052), orf37(84400-83279), orf38(84713-84393), orf39(85576-84710), orf40(85825-87042), orf41(87094-87702) y orf42(89315-88143) en las sec. con núm. de ident.: 1. Adicionalmente, la agrupación de genes comprende además el orf43 (866-60) y el orf44 (2337-1174) en un gen exógeno de sec. con núm. de ident.: 2 no unido a la sec. con núm. de ident.: 1 con la longitud completa de 2337 pb.

La presente descripción proporciona además una secuencia de aminoácidos de 4'-fosfopanteteinil transferasa (PPT), la secuencia de aminoácidos consiste en 214 aminoácidos en la sec. con núm. de ident.: 3 y se denomina como IA-W1, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 1-645 en la sec. con núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos de un factor de

regulación y control de la transcripción de la familia TetR, la secuencia de aminoácidos consiste en 200 aminoácidos en la sec. con núm. de ident.: 4 y se denomina como IA-W2, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 1810-1208 en la sec. con núm. de ident.: 1.

5 La presente descripción proporciona además una secuencia de aminoácidos de ARNt 23S metil transferasa, la secuencia de aminoácidos consiste en 282 aminoácidos en la sec. con núm. de ident.: 5 y se denomina como IA-W3, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 3133-2285 en la sec. con núm. de ident.: 1.

10 La presente descripción proporciona además una secuencia de aminoácidos de 3-O-aciltransferasa, la secuencia de aminoácidos consiste en 408 aminoácidos en la sec. con núm. de ident.: 6 y se denomina como IA-W4, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 3614-4840 en la sec. con núm. de ident.: 1.

15 La presente descripción proporciona además una secuencia de aminoácidos de O-metil transferasa, la secuencia de aminoácidos consiste en 221 aminoácidos en la sec. con núm. de ident.: 7 y se denomina como IA-W5, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 4846-5511 en la sec. con núm. de ident.: 1.

20 La presente descripción proporciona además una secuencia de aminoácidos de crotonil coenzima A reductasa, la secuencia de aminoácidos consiste en 449 aminoácidos en la sec. con núm. de ident.: 8 y se denomina como IA-W6, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 7150-5801 en la sec. con núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos de glicosil transferasa, la secuencia de aminoácidos consiste en 421 aminoácidos en la sec. con núm. de ident.: 9 y se denomina como IA-W7, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 8444-7179 en la sec. con núm. de ident.: 1.

25 La presente descripción proporciona además una secuencia de aminoácidos de proteína accesoria de glicosiltransferasa, la secuencia de aminoácidos consiste en 415 aminoácidos en la sec. con núm. de ident.: 10 y se denomina como IA-W8, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 9729-8482 en la sec. con núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos de NDP-aminoxosa N-dimetil transferasa, la secuencia de aminoácidos consiste en 237 aminoácidos en la sec. con núm. de ident.: 11 y se denomina como IA-W9, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 10543-9830 en la sec. con núm. de ident.: 1.

5 La presente descripción proporciona además una secuencia de aminoácidos que comprende un dominio estructural de policétido sintasa de cetosintasa (KS)₈-aciltransferasa (AT)₈-cetoreductasa (KR)₈-proteína portadora de acilos (ACP)₈-tioesterasa liberadora de cadena (TE), la secuencia de aminoácidos consiste en 1890 aminoácidos en la sec. con núm. de ident.: 12 y se denomina como IA-W10, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 16215-10543
10 en la sec. con núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos que comprende un dominio estructural de policétido sintasa de KS₇-AT₇-KR₇-ACP₇, la secuencia de aminoácidos consiste en 1582 aminoácidos en la sec. con núm. de ident.: 13 y se denomina como IA-W11, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 21076-16328 en la sec. con
15 núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos que comprende un dominio estructural de policétido sintasa KS₅-AT₅-KR₅-ACP₅-KS₆-AT₆-DH₆ (deshidrogenasa)-ER₆ (enoilreductasa)-KR₆-ACP₆, la secuencia de aminoácidos consiste en 3795 aminoácidos en la sec. con núm. de ident.: 14 y se denomina como IA-W12, y la secuencia de nucleótidos de un gen
20 codificante se selecciona de las bases 32511-21124 en la sec. con núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos que comprende un dominio estructural de policétido sintasa KS₄-AT₄-DH₄-KR₄-ACP₄, la secuencia de aminoácidos consiste en 2004 aminoácidos en la sec. con núm. de ident.: 15 y se denomina como IA-W13, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 38599-32585 en la sec. con
25 núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos que comprende un dominio estructural de policétido sintasa KS₁-AT₁-ACP₁-KS₂-AT₂-KR₂-ACP₂-KS₃-AT₃-DH₃-KR₃-ACP₃, la secuencia de aminoácidos consiste en 4538 aminoácidos en la sec. con núm. de ident.: 16

y se denomina como IA-W14, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 52259-38643 en la sec. con núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos de citocromo P-450 oxidasa, la secuencia de aminoácidos consiste en 403 aminoácidos en la sec. con núm. de ident.: 17 y se denomina como IA-W15, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 53099-54310 en la sec. con núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos de NDP-hexosa isomerasa, la secuencia de aminoácidos consiste en 116 aminoácidos en la sec. con núm. de ident.: 18 y se denomina como IA-W16, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 54495-54845 en la sec. con núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos de NDP-hexosa aminotransferasa, la secuencia de aminoácidos consiste en 399 aminoácidos en la sec. con núm. de ident.: 19 y se denomina como IA-W17, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 54842-56041 en la sec. con núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos de NDP-glucosa sintasa, la secuencia de aminoácidos consiste en 302 aminoácidos en la sec. con núm. de ident.: 20 y se denomina como IA-W18, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 56038-56946 en la sec. con núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos de NDP-glucosa-4,6-deshidrogenasa, la secuencia de aminoácidos consiste en 345 aminoácidos en la sec. con núm. de ident.: 21 y se denomina como IA-W19, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 56930-57967 en la sec. con núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos de NDP-hexosa-2,3-deshidrogenasa/tioesterasa, la secuencia de aminoácidos consiste en 745 aminoácidos en la sec. con núm. de ident.: 22 y se denomina como IA-W20, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 57937-60174 en la sec. con núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos de NDP-hexosa aminotransferasa, la secuencia de aminoácidos consiste en 382 aminoácidos en la sec. con núm. de

ident.: 23 y se denomina como IA-W21, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 60836-61984 en la sec. con núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos de NDP-aminohexosa N-dimetiltransferasa, la secuencia de aminoácidos consiste en 239 aminoácidos en la
5 sec. con núm. de ident.: 24 y se denomina como IA-W22, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 62796-62077 en la sec. con núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos de un factor de regulación y control de la transcripción, la secuencia de aminoácidos consiste en 670 aminoácidos en la sec. con núm. de ident.: 25 y se denomina como IA-W23, y la secuencia de nucleótidos de un
10 gen codificante se selecciona de las bases 63633-65645 en la sec. con núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos de NDP-aminohexosa isomerasa, la secuencia de aminoácidos consiste en 354 aminoácidos en la sec. con núm. de ident.: 26 y se denomina como IA-W24, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 67379-66318 en la sec. con núm. de ident.: 1.

15 La presente descripción proporciona además una secuencia de aminoácidos de la proteína de transporte ABC, la secuencia de aminoácidos consiste en 550 aminoácidos en la sec. con núm. de ident.: 27 y se denomina como IA-W25, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 69004-67352 en la sec. con núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos de NDP-hexosa
20 deshidrogenasa, la secuencia de aminoácidos consiste en 433 aminoácidos en la sec. con núm. de ident.: 28 y se denomina como IA-W26, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 69349-70650 en la sec. con núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos de enzima similar a GTP, la secuencia de aminoácidos consiste en 482 aminoácidos en la sec. con núm. de ident.: 29 y
25 se denomina como IA-W27, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 72156-70708 en la sec. con núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos de NDP-hexosa isomerasa, la secuencia de aminoácidos consiste en 346 aminoácidos en la sec. con núm. de ident.:

30 y se denomina como IA-W28, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 72422-73462 en la sec. con núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos de NDP-hexosa cetorreductasa, la secuencia de aminoácidos consiste en 346 aminoácidos en la sec. con núm. de
5 ident.: 31 y se denomina como IA-W29, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 74601-73561 en la sec. con núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos de proteína accesorio de glicosiltransferasa, la secuencia de aminoácidos consiste en 415 aminoácidos en la sec. con núm. de ident.: 32 y se denomina como IA-W30, y la secuencia de nucleótidos de un gen
10 codificante se selecciona de las bases 74913-76160 en la sec. con núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos de glicosiltransferasa, la secuencia de aminoácidos consiste en 422 aminoácidos en la sec. con núm. de ident.: 33 y se denomina como IA-W31, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 76218-77486 en la sec. con núm. de ident.: 1.

15 La presente descripción proporciona además una secuencia de aminoácidos de glicosiltransferasa, la secuencia de aminoácidos consiste en 391 aminoácidos en la sec. con núm. de ident.: 34 y se denomina como IA-W32, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 77606-78781 en la sec. con núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos de NDP-hexosa
20 cetorreductasa, la secuencia de aminoácidos consiste en 330 aminoácidos en la sec. con núm. de ident.: 35 y se denomina como IA-W33, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 78783-79775 en la sec. con núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos de NDP-hexosa reductasa, la secuencia de aminoácidos consiste en 335 aminoácidos en la sec. con núm. de ident.:
25 36 y se denomina como IA-W34, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 79772-80779 en la sec. con núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos de NDP-hexosa metiltransferasa, la secuencia de aminoácidos consiste en 410 aminoácidos en la sec. con núm. de

ident.: 37 y se denomina como IA-W35, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 82055-80823 en la sec. con núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos de metoximalonil sintetasa, la secuencia de aminoácidos consiste en 370 aminoácidos en la sec. con núm. de ident.: 38
5 y se denomina como IA-W36, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 83164-82052 en la sec. con núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos de deshidrogenasa, la secuencia de aminoácidos consiste en 373 aminoácidos en la sec. con núm. de ident.: 39 y se denomina como IA-W37, y la secuencia de nucleótidos de un gen codificante se selecciona de las
10 bases 84400-83279 en la sec. con núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos de proteína portadora de acilos, la secuencia de aminoácidos consiste en 106 aminoácidos en la sec. con núm. de ident.: 40 y se denomina como IA-W38, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 84713-84393 en la sec. con núm. de ident.: 1.

15 La presente descripción proporciona además una secuencia de aminoácidos de metoximalonil deshidrogenasa, la secuencia de aminoácidos consiste en 288 aminoácidos en la sec. con núm. de ident.: 41 y se denomina como IA-W39, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 85576-84710 en la sec. con núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos de glicosil
20 transferasa, la secuencia de aminoácidos consiste en 405 aminoácidos en la sec. con núm. de ident.: 42 y se denomina como IA-W40, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 85825-87042 en la sec. con núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos de NDP-hexosa isomerasa, la secuencia de aminoácidos consiste en 202 aminoácidos en la sec. con núm. de ident.:
25 43 y se denomina como IA-W41, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 87094-87702 en la sec. con núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos de una proteína factor de regulación y control de la transcripción, la secuencia de aminoácidos consiste en 390

aminoácidos en la sec. con núm. de ident.: 44 y se denomina como IA-W42, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 89315-88143 en la sec. con núm. de ident.: 1.

La presente descripción proporciona además una secuencia de aminoácidos de una ARNr 23S metilasa exógena (Tioestreptona y marcador de resistencia *tsr* relacionado), la secuencia de aminoácidos consiste en 269 aminoácidos en la sec. con núm. de ident.: 45 y se denomina como IA-W43, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 866-57 en la sec. con núm. de ident.: 2.

La presente descripción proporciona además una secuencia de aminoácidos de 4"-micarosa glucósido isovaleril transferasa exógena, la secuencia de aminoácidos consiste en 388 aminoácidos en la sec. con núm. de ident.: 46 y se denomina como IA-W44, y la secuencia de nucleótidos de un gen codificante se selecciona de las bases 2337-1171 en la sec. con núm. de ident.: 2.

Se describen además los 44 genes completos de la agrupación de genes biosintéticos de carrimicina de la presente descripción, sobre la base de obtener información de la agrupación de genes biosintéticos de carrimicina, y analizar las posibles funciones de proteínas codificadas de cada gen por medio de bloqueo de genes y comparación de homólogos, y la agrupación de genes tiene una estructura que se muestra en la Figura 1, específicamente:

- (1) cinco genes de policétido sintasa, que incluye orf10-14;
- (2) nueve genes relacionados con la unidad de extensión de síntesis de policetonas y genes relacionados con la modificación, que incluyen orf1, orf4-6, 15 y 36-39;
- (3) dieciséis genes relacionados con la síntesis de glicosilo, que incluyen orf9, 16-22, 24, 26, 28, 29, 33-35 y 41;
- (4) seis genes relacionados con transferencia de glicosilo, que incluyen orf7, 8, 30-32 y 40;
- (5) dos genes relacionados con la resistencia, que incluyen orf3 y 25;
- (6) 4 genes relacionados en general con la regulación y control de la biosíntesis, que incluyen orf2, 23, 27 y 42; y
- (7) dos genes, que incluyen el orf43 de gen marcador diseñado por ingeniería genética de un gen exógeno (tioestreptona y un gen de resistencia *tsr*) y el orf44 de un gen de micarosa 4"-O-

isovaleril transferasa unido al orf43.

Cinco genes de policétido sintasa (orf10-14) en la sec. con núm. de ident.: 1, las secuencias de nucleótidos complementarias y la secuencia de aminoácidos de estas son esenciales para la síntesis de un anillo de lactona de carrimicina. Los 5 genes de la policétido sintasa comprenden 8 módulos y 37 dominios estructurales que se muestran en la Figura 2. El orf14 comprende 3 módulos: un dominio estructural 1 de carga, un módulo 2 y un módulo 3; en un dominio del módulo de carga, KS₁, AT₁ y ACP₁ son responsables por la síntesis inicial del anillo de lactona, y el ácido acético se cataliza como una unidad de iniciación. El módulo 2 comprende los dominios estructurales KS₂, AT₂, KR₂ y ACP₂; y el módulo 3 comprende los dominios estructurales KS₃, AT₃, DH₃, KR₃ y ACP₃ y es responsable por la introducción de 2 unidades de extensión adicionales de ácido acético para formar, finalmente, un esqueleto de cadena carbonada de C11-15 de carrimicina. El orf13 comprende un módulo 4 que comprende KS₄, AT₄, DH₄, KR₄ y ACP₄ y es responsable por la extensión de una tercera unidad de ácido acético para formar, finalmente, un esqueleto de cadena carbonada de C9-10 de carrimicina. El orf12 comprende un módulo 5 y un módulo 6, y el módulo 5 comprende un dominio estructural KS₅-AT₅-KR₅-ACP₅ y es responsable por la introducción de una unidad de extensión de ácido propiónico; y el módulo 6 comprende un dominio estructural KS₆-AT₆-DH₆-KR₆-ER₆-KR₆-ACP₆ y es responsable por la introducción de una unidad de extensión de ácido butírico para formar finalmente un esqueleto de cadena carbonada de C5-C8 de carrimicina. El orf11 comprende un módulo 7 que contiene un dominio estructural KS₇-AT₇-KR₇-ACP₇ y es responsable por la introducción de una unidad de extensión de ácido glicólico para formar finalmente un esqueleto de cadena carbonada de C3-C4 de carrimicina. El orf10 comprende a módulo 8 que contiene un dominio estructural KS₈-AT₈-KR₈-ACP₈-TE y es responsable por la introducción de una unidad de extensión de ácido acético, y la ciclización y la liberación de la cadena carbonada se completan con la participación de tioesterasa (TE). Un diagrama esquemático estructural de los genes de policétido sintasa de carrimicina se muestra en la Figura 2. Todos los dominios estructurales y las posiciones de aminoácidos de estos de los genes de policétido sintasa se muestran en la Tabla 1.

Las secuencias de nucleótidos o secuencias complementarias y las secuencias de aminoácidos

correspondientes de estas de la unidad de extensión de síntesis de policetonas y los orf1, orf4-6, 15 y 36-39 de los genes relacionados con la modificación son: IA-W1, que codifica una proteína portadora de acilos (ACP) sintetizada por policétido modificado por PPT que habilita la proteína para convertirla en una proteína activa; IA-W4, que codifica 3-O-aciltransferasa y es responsable por la acilación del hidroxilo en la posición 3 de carrimicina; IA-W5 y IA-W6, que codifican por separado O-metilasa y crotonoil coenzima A reductasa y son responsables por el suministro de unidades de extensión de policétidos; IA-W15, que codifica citocromo P450 mono-oxidasa y es responsable por la oxidación de una cadena carbonada de policétido; y IA-W36-39, que codifican por separado metoximalonil sintetasa, deshidrogenasa, proteína portadora de acilos y metoximalonil deshidrogenasa. Y todos los genes participan en la síntesis y modificación de las unidades de extensión de policétidos.

Tabla 1 Todos los dominios estructurales y posiciones de aminoácidos de estos de los genes de policétido sintasa:

Tabla 1.1 Todos los dominios estructurales y posiciones de aminoácidos de estos del gen IA-W14 de policétido sintasa

Módulo	Dominio estructural	Posición del aminoácido
Módulo de carga-módulo 1	KS ₁	1-400
	AT ₁	511-814
	ACP ₁	918-989
Módulo 2	KS ₂	1018-1444
	AT ₂	1551-1854
	KR ₂	2160-2338
	ACP ₂	2473-2546
Módulo 3 KS ₃	KS ₃	2570-2995
	AT ₃	3109-3412
	DH ₃	3483-3653
	KR ₃	4068-4248

	ACP ₃	4369-4441
--	------------------	-----------

Tabla 1.2 Todos los dominios estructurales y posiciones de aminoácidos de estos del gen IA-W13 de policétido sintasa

Módulo	Dominio estructural	Posición del aminoácido
Módulo 4	KS ₄	36-461
	AT ₄	575-878
	DH ₄	945-1158
	KR ₄	1532-1711
	ACP ₄	1831-1904

5

Tabla 1.3 Todos los dominios estructurales y posiciones de aminoácidos de estos del gen I IA-W12 de policétido sintasa

Módulo	Dominio estructural	Posición del aminoácido
Módulo 5	KS ₅	37-463
	AT ₅	662-957
	KR ₅	1245-1413
	ACP ₅	1519-1592
Módulo 6 3632-3702	KS ₆	1613-2039
	AT ₆	2157-2458
	DH ₆	2524-2686
	ER ₆	3025-3329
	KR ₆	3338-3518
	ACP ₆	3632-3702

Tabla 1.4 Todos los dominios estructurales y posiciones de aminoácidos de estos del gen IA-W11 de policétido sintasa

Módulo	Dominio estructural	Posición del aminoácido
Módulo 7	KS ₇	35-460
	AT ₇	566-864
	KR ₇	1149-1328
	ACP ₇	1433-1505

5 Tabla 1.5 Todos los dominios estructurales y posiciones de aminoácidos de estos del gen IA-W10 de policétido sintasa

Módulo	Dominio estructural	Posición del aminoácido
Módulo 8	KS ₈	36-461
	AT ₈	579-878
	DH ₈	831-993
	KR ₈	1232-1411
	ACP ₈	1513-1584
	TE ₈	1659-1872

La cantidad de genes relacionados con la síntesis glicosídica de carrimicina es 12 en general,
10 que incluyen orf9, 16-22, 24, 26, 28, 29, 33-35 y 41, en donde, orf18, 19 y 28 codifican las enzimas de síntesis, deshidratación e isomerización de unidades glicosídicas básicas de carrimicina; orf9, 20, 21, 24, 26 y 29 codifican las enzimas de N-dimetilación, 2,3 deshidratación, aminación, isomerización, deshidratación y ceto reducción de NDP-hexosamina en la síntesis de forosamina; orf16, 17 y 22 codifican las enzimas de isomerización, aminación y N-dimetilación de NDP-

hexosamina de micaminosa; y orf33, 34, 35 y 41 codifican enzimas de ceto reducción, metilación y isomerización de NDP-hexosamina de micarosa.

La cantidad de genes relacionados con la transferencia glicosídica de carrimicina es 6 en general, que incluyen orf7, 8, 30-32 y 40, en donde orf7 codifican la glicosilasa de micaminosa; orf8 codifica una proteína accesoria de glicosilación de micaminosa; orf31 y 32 codifican la glicosilasa de forosamina; orf30 codifica una proteína accesoria de glicosilación de forosamina; y orf40 codifica una enzima de glicosilación de micarosa.

La cantidad de genes relacionados con resistencia de carrimicina 2 en general, que incluye orf3 y 25, en donde orf3 codifica una ARN r 23S metilasa; y orf25 codifica una proteína transportadora ABC, orf3 y orf25 dotan a la bacteria productora de carrimicina con resistencia que se produce por antibióticos propios a través de una metilación y mecanismo de bombeo para el ARN del ribosoma.

La cantidad de genes de carrimicina relacionados con la regulación y control de la biosíntesis es 4 en general, que incluyen orf2, 23, 27 y 42, en donde orf2 codifica un factor de la familia TetR de inhibición de la regulación y control de la transcripción y posiblemente participa en la regulación y control negativos en la biosíntesis de carrimicina; orf23 y orf42 codifican por separado dos factores de transcripción de regulación y control positivos, y el último factor de regulación y control positivos de la transcripción sirve como un factor de regulación y control positivos de la transcripción especial de la vía y se usa para regular y controlar directamente la biosíntesis de carrimicina; y orf27 codifica una enzima GTP y posiblemente regula y controla la biosíntesis of carrimicina a través de la regulación y control de las funciones de las células.

Los orf43 y orf44 exógenos se relacionan con la biosíntesis de carrimicina, en donde el orf43 codifica el gen de ARNr 23S metilasa relacionado con resistencia a tioestreptona, el gen se une a orf44 de un gen de micarosa 4"-O-hidroxil isovaleril transferasa, y la expresión de la resistencia de orf43 puede proporcionar un marcador de identificación para la bacteria diseñada por ingeniería genética de carrimicina.

Las secuencias complementarias de la sec. con núm. de ident.: 1 y la sec. con núm. de ident.: 2 de la presente descripción pueden obtenerse en cualquier momento de acuerdo con el principio de apareamiento de bases complementarias. La secuencia de nucleótidos o parte de la secuencia de

nucleótidos de la sec. con núm. de ident.: 1 y la sec. con núm. de ident.: 2 pueden obtenerse a través de una reacción en cadena de la polimerasa (PCR), o digestión enzimática del ADN correspondiente mediante el uso de una endonucleasa de restricción adecuada o mediante el uso de otras tecnologías adecuadas. Pueden obtenerse genes similares a los genes de la biosíntesis de carrimicina de otros organismos, a través de la secuencia de nucleótidos o parte de la secuencia de nucleótidos proporcionadas por la presente descripción, mediante el uso de un método de reacción en cadena de la polimerasa (PCR), o un método para llevar a cabo hibridación Southern mediante el uso de ADN que contiene las secuencias de la presente descripción como una sonda.

La presente descripción proporciona además una vía de obtener al menos parte de la secuencia de ADN en la sec. con núm. de ident.: 1 y la sec. con núm. de ident.: 2 para construir un vector recombinado.

La presente descripción proporciona además una vía de bloquear los genes de la biosíntesis de carrimicina, en donde al menos uno de los genes comprende la secuencia de nucleótidos en la sec. con núm. de ident.:1.

Pueden obtenerse nuevos derivados de carrimicina a través del bloqueo de una o más etapas de la biosíntesis de carrimicina mediante el uso de clon de genes o fragmentos de ADN de la secuencia de nucleótidos o al menos parte de la secuencia de nucleótidos proporcionada por la presente descripción. La secuencia de nucleótidos comprende los fragmentos de ADN o genes y puede usarse para aumentar la producción de carrimicina o derivados de esta.

El clon de ADN de la secuencia de nucleótidos o al menos parte de la secuencia de nucleótidos proporcionados por la presente descripción puede usarse para localizar más genotecas de plásmidos de una genoteca genómica. Estas genotecas de plásmidos comprenden al menos parte de las secuencias de la presente descripción y contienen además ADN de regiones, adyacentes a las genotecas de plásmidos, en un genoma de bacteria productora de carrimicina.

La secuencia de nucleótidos proporcionadas por la presente descripción pueden modificarse o mutarse. Las formas de modificación o mutación comprenden la inserción o sustitución, una reacción en cadena de la polimerasa, una reacción en cadena de la polimerasa mediada por errores, la mutación específica de sitio, la reconexión de secuencias y la mutación causada por luz ultravioleta o reactivos

químicos.

La secuencia de nucleótidos proporcionada por la presente descripción puede evolucionar directamente (intercambio de cadenas de ADN) a través de diferentes partes de las secuencias o secuencias homólogas de otras fuentes.

5 Los fragmentos o dominios estructurales o módulos o genes de las secuencias de nucleótidos o al menos parte de las secuencias de nucleótidos proporcionadas por la presente descripción, pueden usarse para construir una genoteca de la policétido sintasa o una genoteca de derivados de la policétido sintasa o una genoteca de colección. Nuevos compuestos de policetona se producen a través de la delección o desactivación de uno o más dominios estructurales, módulos o genes de
10 policétido sintasa de los mismos sistemas de policétido sintasa o diferentes, o mediante aumento de uno o más dominios estructurales, módulos o genes de policétido sintasa.

Las secuencias de nucleótidos de los genes modificadores de la biosíntesis y genes de la síntesis de glicosilo y de la glicosiltransferasa de la presente descripción proporcionan una vía de obtener derivados de carrimicina a través de delección, sustitución o reformación de estos genes modificadores
15 y de la síntesis y transferencia de glicosilo.

Los fragmentos o dominios estructurales o módulos o genes de las secuencias de nucleótidos o al menos parte de las secuencias de nucleótidos proporcionadas por la presente descripción pueden usarse para aumentar la producción de carrimicina o derivados de esta mediante duplicación de la cantidad.

20 Los clones de genes de las secuencias de nucleótidos o al menos parte de las secuencias de nucleótidos proporcionadas por la presente descripción pueden expresarse en huéspedes exógenos a través de sistemas de expresión adecuados para obtener carrimicina modificada o carrimicina con mayor bioactividad o mayor rendimiento. Los huéspedes exógenos comprenden *Streptomyces*, *Escherichia coli*, *Bacillus*, levadura, plantas, animales, etc.

25 Los genes o agrupaciones de genes de las secuencias de nucleótidos o al menos parte de las secuencias de nucleótidos proporcionadas por la presente descripción pueden expresarse en huéspedes heterólogos, y las funciones de los genes o agrupaciones de genes en cadenas del metabolismo de los huéspedes se comprenden a través de una tecnología de chip de ADN.

Los polipéptidos de las secuencias de aminoácidos o al menos parte de las secuencias de aminoácidos proporcionadas por la presente descripción pueden aún tener bioactividad, incluso nueva actividad biológica después de que uno o algunos aminoácidos se eliminan o sustituyen, o se aumenta el rendimiento, o se optimizan las características dinámicas u otras propiedades perseguidas de las proteínas. Pueden obtenerse nuevas proteínas o enzimas a través de la conexión de las secuencias de aminoácidos de la presente descripción mediante de la tecnología de delección apropiada, y después, se producen productos nuevos o asociados.

Las secuencias de aminoácidos proporcionadas por la presente descripción pueden usarse para separar las proteínas requeridas y pueden aplicarse en la preparación de anticuerpos.

Las secuencias de aminoácidos proporcionadas por la presente descripción proporcionan la posibilidad de predecir una estructura tridimensional de la policétido sintasa.

Los genes proporcionados por la presente descripción y las proteínas y anticuerpos de estas pueden usarse además para la selección y desarrollo de compuestos o proteínas para medicamentos, para la industria y la agricultura.

BREVE DESCRIPCIÓN DE LOS DIBUJOS:

Figura 1: Una estructura de una agrupación de genes biosintéticos de carrimicina.

Figura 2: Una estructura de un gen de la policétido sintasa de carrimicina.

Figura 3: Un diagrama esquemático de la construcción de los plásmidos recombinantes de bloqueo de IA-W4, el gen de la 3-O-aciltransferasa y el doble intercambio.

Figura 4: Un diagrama esquemático de la construcción de los plásmidos recombinantes de bloqueo del gen IA-W42 de regulación y control de la transcripción, etc.

Figura 5A: Verificación del bloqueo de un gen IA-W4 de 3-O-aciltransferasa mediante PCR, en la cual, 1: cepa original; 2, 3, 4: mutante con gen bloqueado; y 5: marcador de ADN III.

Figura 5B: Verificación del bloqueo de otros genes mediante PCR, en la cual, III: marcador de ADN III; C: cepa original; y M: mutante con gen bloqueado.

Figura 6: Análisis por HPLC de productos fermentados de mutantes con gen IA-W4 de 3-O-aciltransferasa bloqueado, en la cual: a: un control de carrimicina; y b: extracto fermentado del

mutante con gen bloqueado, I, II y III son picos de absorción de tres ingredientes principales, es decir, isovaleril espiramicina I, isovaleril espiramicina II e isovaleril espiramicina III de la carrimicina, respectivamente.

De acuerdo con la presente descripción, las cepas mutantes se obtienen a través de experimentos de bloqueo de genes; y se comprueba por los experimentos que el cambio de ingrediente de carrimicina de las cepas mutantes es causado por el bloqueo de genes, o que la carrimicina ya no se produce. Y así es indicador que la información obtenida de la agrupación de genes se relaciona con la biosíntesis de carrimicina. De acuerdo con la presente descripción, un gen marcador exógeno de resistencia a tioestreptona (orf43) y un gen de micarosa 4"-O-hidroxil isovaleril transferasa (orf44) unido al orf43 se integran a los cromosomas de la bacteria productora de carrimicina a través de recombinación génica homóloga. Nuestro laboratorio ha demostrado que el orf43 y el orf44 son esenciales para la biosíntesis de carrimicina a través de investigaciones (Chinese Journal of Biotechnology, volumen 15, número 2, 1999, 171-176).

DESCRIPCIÓN DETALLADA

Las modalidades proporcionadas más abajo se usan sólo para ayudar a los expertos en la técnica a comprender mejor la presente descripción, en lugar de limitar, de cualquier manera, la presente descripción.

Modalidad 1: Extracción del ADN total de las bacterias productoras de carrimicina (*S. spiramyceticus*)

Formula de medio de cultivo de R₂YE (g/100ml):

Sacarosa 10.3

Glucosa 1.0

Extracto de levadura 0.4

Tryptona 0.2

Peptona 0.4

Hidrolizado de caseína 0.1

K₂SO₄ 0.025

CaCl₂ 0.216
 KH₂PO₄ 0.005
 MgCl₂.6H₂O 1.012
 NaOH(1M) 0.5ml

5 Tris-HCl(0.25mol/L pH7.2) 10ml

Se añaden 0.2 mL de solución de oligoelementos y se añade agua destilada hasta que el volumen sea de 100 mL y el pH sea de 6.5.

Solución de oligoelementos (g/100mL):

10 ZnCl₂ 0.004

FeCl₃.6H₂O 0.02

CuCl₂.2H₂O 0.001

MnCl₂.4H₂O 0.001

Na₂B₄O₇.10H₂O 0.001

15 (NH₄)₆Mo₇O₂.4H₂O 0.001

La esterilización a presión de 15 libras se realiza durante 20 min a una temperatura de 121 grados centígrados.

S. spiramyceticus se inoculó en 25 mL de medio de cultivo R₂YE, se realizó un cultivo en mesa vibratoria durante 48 horas a una temperatura de 28 grados centígrados. Después se realizó el subcultivo en 100 mL de medio de cultivo R₂YE, se realizó un cultivo en mesa vibratoria durante 24 horas a una temperatura de 28 grados centígrados. Después los talos se recogieron (aproximadamente 10 g) después de la centrifugación durante 10-15 min a una velocidad de 5,000 rpm. La operación se realizó principalmente de acuerdo con las especificaciones de producto de la Empresa de Ciencias de la Vida UPTECHTM. Se añadieron 50 mL de solución de 25 mM de EDTA a los talos para lavado bajo vibración, y después se centrifugó la solución y se descartó el sobrenadante. Los micelios se suspendieron con 25 mL de solución de lisozima (10 mg/mL, preparados a partir de 10 mg de Tris-HCl con un pH de 8.0, 2 mg de EDTA y 1.2 % de TritonX-100 mediante la adición de 0.5 mL de 100 mg/mL de ARNasa), y se cultivaron durante 1-2 h a una temperatura de 37 grados centígrados hasta

que las células se volvieron translúcidas. Después se añadió 2.5 mL de solución de proteasa K y se realizó el cultivo durante 30 minutos a una temperatura de 55 grados centígrados. Después se añadió 20 mL de solución de SDS al 10 % y se realizó el cultivo durante 10 min a una temperatura de 70 grados centígrados. Después se transfirió la solución a una columna de purificación de ADN, se centrifugó durante 1 min a una velocidad de 12,000 rpm, después se añadió 50 mL de solución que contenía proteasa para lavar la columna y se centrifugó durante 1 min a una velocidad de 12,000 rpm a temperatura ambiente. Después, la columna se lavó dos veces con 50 mL de solución de enjuague y se centrifugó durante 1 min a una velocidad de 12,000 rpm a la vez. Después se añadió 5-10 mL de eluyente TE, y se colocó durante 2-5 min a temperatura ambiente, y se centrifugó durante 1 min a una velocidad de 12,000 rpm. Se recogió la solución y se guardó el ADN total a una temperatura de -20 grados centígrados.

Modalidad 2: Verificación de funciones de información de los genes en la sec. con núm. de ident.: 1 a través de bloqueo génico

Se realiza mediante bloqueo de genes tales como IA-W1, IA-W42 en dos extremos de la agrupación de genes, e IA-W4, 17, 21, 23 y 27 seleccionados para obtener cepas mutantes. Se demuestra por experimentos, que la capacidad de estas cepas bloqueadas para producir carrimicina ha cambiado la carrimicina o se desaparece. Así, la información obtenida de la agrupación de genes es esencial para la producción de carrimicina. Los cebadores se diseñan de acuerdo con los genes codificantes antes mencionados y las secuencias aguas arriba y aguas abajo de estas y se insertan en los sitios apropiados de digestión enzimática, y las secuencias de cebadores se muestran en la tabla 2.

Tabla 2 Secuencias de cebadores diseñados para el experimento de bloqueo génico

	Gen	Secuencia del cebador	Gen	Secuencia del cebador
5	IA-W1	CCGGAATTCGCCCTTGAACGCTTGT CCG <i>EcoRI</i>	IA-W42	CTAGTCTAGACCGCCGACCGCA AACTCTC <i>XbaI</i>
		CGCGGATCCGCTCACTCGGCAGGAT GGG <i>BamHI</i>		AACTGCAGCGACGTTCTCCTCC TCACCG <i>PstI</i>
		AACTGCAG TCCGTCTACAAGGCGT GGTT <i>PstI</i>		CGGGGTACCCGACCTGTGGCTG ACCGAC <i>KpnI</i>
		CTAGTCTAGACGTATCGGTTTCGTCG AGGTCT <i>XbaI</i>		CCGGAATTCGTGGACGACACCT GTATGAAC <i>EcoRI</i>
10	IA-W4	AGTGTCTAGACGGCGCGCGGCACG GGGTTGAACTC <i>XbaI</i>	IA-W17	CCGGAATTCATCCCCTTCCTCG ACGCAG <i>EcoRI</i>
		GACAAGCTTTGGATTCTCGCTCCTC TTTCGGGATGG <i>Hind III</i>		AACTGCAGGGCGGTACGGGGT AGTGGAT <i>PstI</i>
		GACAAGCTTTGAGCGTGGCAGACC AGACCGCTCT <i>Hind III</i>		CGCGGATCCGAGAGCCTCAGC CTTCCC <i>BamHI</i>
		AGTGGAAATTCACACAGGGCAAGGT CGGCGTGCTCTG <i>EcoRI</i>		CTAGTCTAGACCGTACTCCCTG GCGTTGTT <i>XbaI</i>
15	IA-W21	CCGGAATTCGACCGCATCCGCTACG ACG <i>EcoRI</i>	IA-W23	CCGGAATTCTTACCTGGATTATG GTGAAG <i>EcoRI</i>
		CGCGGATCCGAGCCATTGGTCGTCG AAGA <i>BamHI</i>		CGGCCGAGCGGGCTGCAGA <i>PstI</i>
		AACTGCAGACCGACGGCATCTACAC CAC <i>PstI</i>		CGGGGTACCGGAGTACAACGC CGGCTTC <i>KpnI</i>
		CTAGTCTAGACCGAGGACCGCAAGG ACTACG <i>XbaI</i>		CTAGTCTAGACCGAGCACGGTC CGGGAGG <i>XbaI</i>
20	IA-W27	CCGGAATTCGCGTGCTCACCGACAA CCTG <i>EcoRI</i>	AACTGCAG TCGGGCCATCTTGTCGTTG <i>PstI</i>	
		CGCGGATCCGGGAAGTCCTCACTGC TCAAC <i>BamHI</i>	CTAGTCTAGACCTTCAGGGTGCCGTAG TC <i>XbaI</i>	

Los fragmentos de genes homólogos correspondientes se obtienen por separado mediante
amplificación por PCR, y el plásmido recombinante que contiene genes homólogos se obtuvo
mediante la adopción de los sitios de digestión enzimática correspondientes, por inserción de genes
resistentes marcados y seleccionados (Apramicina-Am) y conexión de los genes con un vector
pKC1139 sensible a la temperatura [Bierman M. y otros, Gene, 1992;116(1): 43-9] o vector de
Escherichia coli/Streptomyces pGH112 [Youbao Biology Company]. El plásmido recombinante se

transformó con protoplastos y se transfirió a la bacteria productora de carrimicina. Después del cultivo, se aislaron colonias individuales para obtener el fragmento homólogo de la cepa de bloqueo genético de doble intercambio. En la Figura 3 se muestra un diagrama esquemático de la construcción de plásmidos recombinantes bloqueadores de un gen IA-W4 de 3-O-aciltransferasa y de doble
5 intercambio. En la Figura 4 se muestra un diagrama esquemático de la construcción de los plásmidos recombinantes bloqueadores de genes tal como el gen IA-W42 de regulación y control de la transcripción.

El ADN total de las cepas bloqueadas y el ADN total de las cepas originales se sometieron a la verificación por PCR mediante adopción por separado de los cebadores correspondientes, como se
10 muestra en la Figura 5A y Figura 5B. Se muestra como resultado de la Figura 5A, que se delecionaron 613 pb en el gen que codifica orf4 de una cepa mutante con el gen IA-W4 bloqueado. En la Figura 5B se ilustra un resultado de verificación por PCR, comparado con las cepas originales, la longitud de los productos de PCR aumenta a medida que los genes resistentes marcados y seleccionados se insertan en la cepa mutante con genes codificantes relacionados bloqueados.

Se comprueba mediante experimentos de fermentación y detección de productos por HPLC, que las cepas mutantes de IA-W4 bloqueado ya no producen 4"-isovaleril espiramicina III y II y están dominadas por 4"-isovaleril espiramicina I como ingrediente principal (Figura 6). Se comprueba que un gen IA-W4 de 3-O-acetiltransferasa en la información genética en la sec. con núm.
15 de ident.:1 proporcionada por la presente descripción participa en la biosíntesis de carrimicina. Debido al bloqueo del gen, las cepas mutantes pierden la función de acilación del hidroxilo en posición 3 de un anillo de lactona de carrimicina.

Se comprueba por experimentos de fermentación en otras cepas con genes bloqueados y de actividad antibacteriana y de detección por HPLC en los productos, que las cepas bloqueadas ya no producen carrimicina activada.

Se demuestra que la agrupación de genes en la sec. con núm. de ident.: 1 proporcionada por la presente descripción participa en la biosíntesis de carrimicina.
25

Modalidad 3: Selección de cepas para la transferencia y el bloqueo génicos de bacterias productoras

de carrimicina

3.1 Preparación of protoplastos:

Se inocularon con inclinación esporas frescas de bacterias productoras de carrimicina en medio de cultivo líquido R₂YE y se agitaron durante 48 h a una temperatura de 28 grados centígrados a una velocidad de 220 rpm. El cultivo fluido se inoculó a una velocidad de inoculación de 10 % en medio de cultivo líquido fresco R₂YE que contenía 0.5 % de glicina y se agitó durante 20 h a una temperatura de 28 grados centígrados. Se introdujeron 10 mL de solución bacteriana en un tubo de centrífuga y se centrifugó a una velocidad de 3,000 rpm para recoger el micelio. El micelio se lavó con tampón P.

10	Tris-HCl (pH8.0)	1mol/L	3.1ml
	CaCl ₂ ·2H ₂ O	3.68	
	MgCl ₂ ·6H ₂ O	2.04	
	Sacarosa	103	
	Glucosa	1.0	
15	Solución de lemento traza	2.0 ml	
	PH 7.6		

La esterilización a presión de 15 libras se realiza durante 30 min a una temperatura de 121 grados centígrados.

Después de lavar dos veces, el micelio se suspendió con un volumen adecuado de tampón P, se añadió una solución de tampón P con lisozima (concentración final es de 2 mg/mL), se mezcló de manera uniforme y se incubó durante 30-45 min en un baño de agua a una temperatura de 37 grados centígrados, y se agitó una vez cada 10-15 min. Las condiciones de formación de protoplastos se observaron con un microscopio de diferencia de fase de 10×40. La enzimólisis se interrumpió cuando el examen microscópico muestra que la mayoría de los micelios han formado protoplastos. Después de filtrar a través de algodón absorbente, el licor del filtro se sometió a un lavado centrífugo dos veces con tampón P. Finalmente, el protoplasto se suspendió con 1 mL de tampón P, y la suspensión se cargó separadamente en tubos EP por 100 µL/tubo, y se conservó a una temperatura de -70 grados

centígrados para su uso posterior.

3.2 Transformación de protoplastos por ADN plasmídico:

Se tomaron 100 µL de protoplasto y se añadió a 10 µL de solución de ADN plasmídico, se volteó una pared de tubo para realizar una mezcla uniforme. 400 µL de tampón P que contenía 25 % de PEG-1000 (un producto de la empresa Britain Koch-light) se añadió rápidamente, se realizó succión-soplado y mezclado uniforme, y se colocó por 5 minutos a temperatura ambiente. Se cubrió una placa plana deshidratada de R₁R₂YE con 200 µl de mezcla, se cultivó durante 20 h a una temperatura de 28 grados centígrados, cubierta con 50 µg/mL de agua estéril con tsr, se cultivó durante 5-7 días a una temperatura de 28 grados centígrados, y se recogieron los transformantes.

3.3 Selección de cepas mutantes con genes bloqueados

Los transformantes se recogen en un medio de cultivo que contiene 50 µg/mL de Tsr.

Polvo de torta de frijol de soya 20

15 Glucosa 10

Almidón 30

CaCO₃ 5.0

NaCl 4.0

agar 18

20 Se usa agua desionizada para la preparación, y se realiza una esterilización a presión de 15 libras durante 30 min a una temperatura de 121 grados centígrados a un valor natural de pH.

El cultivo se realizó durante 5-7 días a una temperatura de 28 grados centígrados, el paso de la generación 4-5 se llevó a cabo en un medio de cultivo no dosificado. Las monoesporas se separaron.

25 Las monoesporas se examinaron por separado y de manera correspondiente en un medio de cultivo que contenía Am (Am 50 µg/mL) para seleccionar cepas con genes bloqueados que crecen en Am y no en Tsr. Se recogieron cepas bloqueadas con expresión estable de marcadores de resistencia, se extrajo el ADN de los genomas de las cepas bloqueadas, se realizó la amplificación por PCR

mediante la adopción de los cebadores correspondientes en la modalidad 2, y se valoró la corrección del bloqueo genético de acuerdo con el tamaño del producto y de la secuenciación del ADN.

Modalidad 4: Fermentación de con genes bloqueados y bacterias productoras de carrimicina y
5 detección e identificación de la actividad del producto

4.1 Fermentación

Medio de cultivo inclinado (g/L):

Polvo de torta de frijol de soya 20.0

Glucosa 10.0

10 Almidón 30.0

CaCO₃ 5.0

NaCl 4.0

Agar 18.0

Se usa agua desionizada para la preparación, y se realizó una esterilización a presión de 15 libras
15 durante 30 min a una temperatura de 121 grados centígrados a un valor natural de pH.

Las cepas se cultivaron en un medio de cultivo inclinado de 10-12 d a una temperatura de 28 grados centígrados. Después que las cepas crecieron, se inocularon en un matraz triangular de 100 mL que contenía 30 mL de medio de cultivo de fermentación y se agitaron durante 96-120 h a una temperatura de 28 grados centígrados.

20 Medio de cultivo de fermentación (g/L):

Glucosa 5.0

Cloruro de sodio 10

Almidón 60

Sulfato de magnesio 1.0

25 Carbonato de calcio 5.0

Nitrato de amonio 6.0

Fosfato de dihidrógeno potásico 0.5

Levadura en polvo 5.0

Harina de pescado 20.0

valor natural de pH

Se usa agua desionizada para la preparación, y se realizó una esterilización a presión de 15 libras durante 30 min a una temperatura de 121 grados centígrados.

5

4.2 Detección de la actividad del producto de fermentación:

Se centrifugó el licor de fermentación, se tomó el sobrenadante y se diluyó, y después se realizó la detección con *Bacillus subtilis* como bacteria de detección en referencia a un ensayo microbiológico de acetil espiramicina (II), 2005 <Farmacopea de la República Popular China>. La detección se realiza mediante adopción de un método de placa cilíndrica con un método de curva estándar.

10

4.3 Extracción e identificación del producto de la fermentación:

El licor de fermentación se centrifugó durante 15 min a una velocidad de 3,000 rpm a temperatura ambiente, el pH del sobrenadante se ajustó a 8.5 con NaOH 1M. Después, el sobrenadante se extrajo con 1/2 volumen de acetato de etilo. Se extrajo una fase de acetato de etilo y se sometió a un secado por soplado en un plato plano, después se disolvió la sustancia seca en metanol cromatográficamente puro y después se introdujo 10-20 µL de una muestra después del filtrado. Instrumentos de cromatografía: un cromatógrafo de líquidos Shimadzu LC-10ATvp y un detector de matriz de diodos; una columna cromatográfica: Kromasil C₁₈(4.5mm×150mm, 5µm); fase que fluye: CH₃OH/1%NaH₂PO₄ (55:45); longitud de onda de detección: 231 nm; velocidad de flujo: 1 mL/min; y temperatura de la columna: 25 grados centígrados. Los productos de fermentación de cepas mutantes se identifican con un producto estándar de carrimicina como control (comprado al Instituto Nacional para el Control de Productos Farmacéuticos y Biológicos).

20

Los genes y proteínas involucrados en la presente descripción se muestran en una tabla de secuencias.

25

REIVINDICACIONES

1. Una agrupación de genes biosintéticos de carrimicina, que comprende 44 genes que comprende:

- 1) cinco genes de policétido sintasa, que incluyen orf10-14;
- 2) nueve genes relacionados con la unidad de extensión de síntesis de policetonas y con la modificación, que incluyen orf1, orf4-6, 15 y 36-39;
- 3) dieciséis genes relacionados con la síntesis de glicosilo, que incluyen orf9, 16-22, 24, 26, 28, 29, 33-35 y 41;
- 4) seis genes relacionados con transferencia de glicosilo, que incluyen orf7, 8, 30-32 y 40;
- 5) dos genes relacionados con la resistencia, que incluyen orf3 y 25;
- 6) cuatro genes relacionados con la regulación de la biosíntesis, que incluyen orf2, 23, 27 y 42; y
- 7) dos genes, que incluyen el orf43 del gen marcador diseñado por ingeniería genética de un gen exógeno y el orf44 de un gen de micarosa 4"-O-hidroxisovaleril transferasa unido al orf43.

2. La agrupación de genes biosintéticos de acuerdo con la reivindicación 1, caracterizada porque los cinco genes de policétido sintasa codifican una enzima de biosíntesis de policétidos, la enzima de biosíntesis de policétidos cataliza la síntesis de un anillo de lactona de 16 elementos de carrimicina, y

secuencia de aminoácidos correspondientes a la secuencia de nucleótidos o secuencias complementarias de los 5 genes orf10-14 de policétido sintasa comprenden IA-W10, IA-W11, IA-W12, IA-W13 e IA-W14.

3. La agrupación de genes biosintéticos de acuerdo con la reivindicación 1 o la reivindicación 2, caracterizada porque cinco policétido sintasas comprenden las secuencias de aminoácidos correspondientes a las secuencias de nucleótidos o secuencias complementarias de dominios estructurales de la cetosintasa (KS), la aciltransferasa (AT), la cetorreductasa (KR), la deshidratasa (DH), la enoilreductasa (ER), una proteína portadora de acilos (ACP) y la tioesterasa.

4. La agrupación de genes biosintéticos de acuerdo con la reivindicación 3, caracterizada porque las cinco policétido sintetas comprenden módulos o dominios estructurales.
- 5 5. La agrupación de genes biosintéticos de acuerdo con la reivindicación 1, caracterizada porque esas secuencias de aminoácidos correspondientes a las secuencias de nucleótidos o secuencias complementarias de los genes relacionados con la unidad de extensión de síntesis de policetonas y con la modificación que incluyen los orf1, orf4-6, 15 y 36-39, comprenden IA-W1, IA-W4, IA-W5, IA-W6, IA-W15, IA-W36, IA-W37, IA-W38 e IA-W39.
- 10 6. La agrupación de genes biosintéticos de acuerdo con la reivindicación 1, caracterizada porque esas secuencias de aminoácidos correspondientes a las secuencias de nucleótidos o secuencias complementarias de los genes relacionados con la síntesis de glicosilo que incluyen los orf9, 16-22, 24, 26, 28, 29, 33-35 y 41, comprenden IA-W9, IA-W16, IA-W17, IA-W18, IA-W19, IA-W20, IA-
- 15 W21, IA-W22, IA-W24, IA-W26, IA-W28, IA-W29, IA-W33, IA-W34, IA-W35 e IA-W41.
7. La agrupación de genes biosintéticos de acuerdo con la reivindicación 1, caracterizada porque esas secuencias de aminoácidos correspondientes a las secuencias de nucleótidos o secuencias complementarias de los genes relacionados con la transferencia de glicosilo que incluyen los orf7, 8,
- 20 30-32 y 40, comprenden IA-W7, IA-W8, IA-W30, IA-W31, IA-W32 e IA-W40.
8. La agrupación de genes biosintéticos de acuerdo con la reivindicación 1, caracterizada porque esas secuencias de aminoácidos correspondientes a las secuencias de nucleótidos o secuencias complementarias de los genes relacionados con la resistencia que incluyen los orf3 y 25, comprenden
- 25 IA-W3 y IA-W25.
9. La agrupación de genes biosintéticos de acuerdo con la reivindicación 1, caracterizada porque esas secuencias de aminoácidos correspondientes a las secuencias de nucleótidos o secuencias complementarias de los genes relacionados con la regulación de la biosíntesis que incluye los orf2,
- 30 23, 27 y 42, comprenden IA-W2, IA-W23, IA-W27 e IA-W42.

10. La agrupación de genes biosintéticos de acuerdo con la reivindicación 1, caracterizada porque esas secuencias de aminoácidos correspondientes a las secuencias de nucleótidos o secuencias complementarias del orf43 y el orf44 del gen exógeno marcador diseñado por ingeniería genética
5 unido al orf43 comprenden IA-W43 y IA-W44.

RESUMEN

La presente descripción proporciona una agrupación de genes biosintéticos de carrimicina. La agrupación de genes biosintéticos comprende 44 marcos abiertos de lectura de genes (orf), es decir, 5 orfs (orf10-14) que codifican la policétido sintasa, 9 orf (orf1, 4-6, 15 y 36-39) relacionados con la unidad de extensión de la síntesis de policetonas y con la modificación, 16 orf (orf9, 16-22, 24, 26, 28, 29, 33-35 y 41) relacionados con la síntesis de glicósilos, 6 orf (orf7, 8, 30-32 y 40) relacionados con la transferencia de glicosilo, 2 orf (orf3 y 25) relacionados con la resistencia, 4 orf (orf2, 23, 27 y 42) posiblemente relacionados con la regulación, un orf de gen marcador de resistencia tsr (orf43) y un orf de gen de 4"-micarosa glucósido isovaleril transferasa (orf44).