

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 731 913**

51 Int. Cl.:

C12Q 1/689 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **30.01.2015 PCT/EP2015/051926**

87 Fecha y número de publicación internacional: **06.08.2015 WO15114094**

96 Fecha de presentación y número de la solicitud europea: **30.01.2015 E 15702733 (5)**

97 Fecha y número de publicación de la concesión europea: **05.06.2019 EP 3099813**

54 Título: **Pruebas de resistencia genética**

30 Prioridad:

**30.01.2014 EP 14153260
01.08.2014 EP 14179456**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
19.11.2019

73 Titular/es:

**ARES GENETICS GMBH (100.0%)
Karl-Farkas-Gasse 18
1030 Vienna, AT**

72 Inventor/es:

**KELLER, ANDREAS;
KIRSTEN, JAN;
SCHMOLKE, SUSANNE;
STÄHLER, CORD, FRIEDRICH;
RENSEN, GABRIEL y
BACKES, CHRISTINA**

74 Agente/Representante:

SÁEZ MAESO, Ana

ES 2 731 913 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

Pruebas de resistencia genética

5 La invención se refiere a un método para determinar un perfil de resistencia a antibióticos para un microorganismo bacteriano y a un método para determinar la resistencia de un microorganismo bacteriano a un fármaco antibiótico.

10 La resistencia a los antibióticos es una forma de resistencia a fármacos mediante la cual una subpoblación de un microorganismo, por ejemplo, una cepa de una especie bacteriana, puede sobrevivir y multiplicarse a pesar de la exposición a un fármaco antibiótico. Es un problema grave y de salud para el paciente individual, así como un problema fundamental de salud pública. El tratamiento oportuno de una infección bacteriana requiere el análisis de aislados clínicos obtenidos de pacientes relacionados con la resistencia a antibióticos, para seleccionar una terapia efectiva.

15 La resistencia a los fármacos antibacterianos (ADR) representa un problema de salud muy importante. Según el informe global sobre vigilancia de la resistencia a los antimicrobianos de la Organización Mundial de la Salud, la ADR produce 25,000 muertes por año en Europa y 23,000 muertes por año en los Estados Unidos. En Europa, 2,5 millones de días extra de hospitalización suponen un coste social de 1.5 billones de euros. En los Estados Unidos, el costo directo de 2 millones de enfermedades lleva a un costo directo de 20 billones de dólares. Se estima que el costo general es sustancialmente más alto, lo que reduce el producto interno bruto (PIB) hasta en un 1,6%.

20 Actualmente, las pruebas de resistencia / susceptibilidad se llevan a cabo obteniendo un cultivo de bacterias sospechosas, sometiéndolas a diferentes protocolos de antibióticos y determinando en qué casos las bacterias no crecen en presencia de cierta sustancia. En este caso, las bacterias no son resistentes (es decir, susceptibles al fármaco antibiótico) y la terapia puede administrarse a los respectivos pacientes. El documento US7335485 describe un método para determinar la susceptibilidad de un microorganismo a los antibióticos, en donde el organismo se cultiva en presencia de un fármaco antibiótico a analizar. Más recientemente, se aplican tecnologías sensibles como la espectrometría de masas para determinar la resistencia, pero esto continúa requiriendo el cultivo del microorganismo a analizar, en presencia de un fármaco antibiótico que se ensayará. Además, en todas estas técnicas, cada microorganismo que se probará tiene que ensayarse contra los antibióticos individuales o combinaciones de fármacos, lo que requiere pruebas extensas, engorrosas y que consumen mucho tiempo.

25 Se sabe que la resistencia a los fármacos puede estar asociada con polimorfismos genéticos. Esto es válido para los virus, donde la prueba de resistencia está establecida en la práctica clínica (por ejemplo, la genotipificación del VIH). Más recientemente, se demostró que la resistencia también tiene causas genéticas en bacterias e incluso en organismos superiores, como los humanos, donde la resistencia de los tumores contra ciertos agentes citostáticos puede vincularse a mutaciones genómicas.

30 Wozniak y otros (BMC Genomics 2012, 13 (Suppl 7) :S23) describe los determinantes genéticos de la resistencia farmacológica en *Staphylococcus aureus* en base a los datos de genotipo y fenotipo. Stoesser y otros describe la predicción de susceptibilidades antimicrobianas para *Escherichia coli* y *Klebsiella pneumoniae* utilizando datos de secuenciación genómica completa (J Antimicrob Chemother 2013; 68: 2234-2244).

35 Barl y otros (International Journal of Antimicrobial Agents 32 (2008) 272-277) informa un estudio que presenta la integración y la aplicación de dos módulos de chips de ADN para la detección simultánea de polimorfismos de un solo nucleótido en *gyrA* (resistencia a quinolona) y *fimH* (aumento de la adhesión al epitelio del tracto urinario).

40 US 2005/0069897 A1 se relaciona con un método para detectar cepas de *E. coli* resistentes a la quinolona en un material de muestra biológica. Como se discute en la presente descripción, se ha investigado varios aislamientos clínicos de *E. coli* resistentes a la quinolona.

45 Jin y Gross (J. Mol. Biol. (1988) 202, 45-58) describen la resistencia a la rifampicina y la mutación genética relacionada *rpoB* en *Escherichia coli*.

50 Adams-Sapper y otros (Antimicrob Agents Chemother. 2013 Ene; 57 (1): 490-7) compara la composición clonal de aislados clínicos de *E. coli* resistentes y susceptibles a fármacos.

55 *Escherichia coli* (*E. coli*) es una bacteria gramnegativa, anaeróbica facultativa, con forma de bastón que generalmente se encuentra, por ejemplo, en el tracto gastrointestinal inferior de los mamíferos. Si bien muchas especies del género *Escherichia* son inofensivas, algunas cepas de algunas especies son patógenas en los seres humanos y causan infecciones del tracto urinario, enfermedades gastrointestinales, así como también una amplia gama de otras condiciones patológicas. *E. coli* es responsable de la mayoría de estas condiciones patológicas.

60 Continúa existiendo una necesidad de pruebas rápidas y eficientes de resistencia a los antibióticos.

65 La presente invención aborda esta necesidad.

Resumen de la invención

5 Los inventores realizaron estudios exhaustivos sobre el genoma de la bacteria E. coli resistente a los antibióticos y encontraron diferencias notables con la E. coli de tipo silvestre. Sobre la base de esta información, ahora es posible proporcionar un análisis detallado del patrón de resistencia de las cepas de E. coli basados en genes individuales o mutaciones a nivel de nucleótidos. Este análisis implica la identificación de una resistencia contra los antibióticos individuales, así como grupos de ellos. Esto permite no solo la determinación de una resistencia a un solo antibiótico, sino también a grupos de antibióticos como los antibióticos de lactama o quinolona, o incluso a todos los fármacos antibióticos relevantes.

10 Por lo tanto, la presente invención facilitará considerablemente la selección de un fármaco antibiótico apropiado para el tratamiento de una infección por E. coli en un paciente y, por lo tanto, mejorará en gran medida la calidad del diagnóstico y el tratamiento.

15 La presente invención se define en las reivindicaciones anexas.

En una modalidad de la invención, se determina la presencia de una mutación en un gen que se selecciona del grupo que consiste en potB, ycgK, ycgB, valS y yjjJ.

20 En una modalidad de la invención, el método comprende determinar la presencia de una mutación en más de un gen que se selecciona a partir de dicho grupo.

En una modalidad de la invención, la presencia de una mutación en un gen seleccionado de dicho grupo es indicativo de una resistencia a los fármacos antibióticos que pertenecen a diferentes grupos de fármacos antibióticos.

25 En una modalidad de la invención, la mutación en el gen potB es L267V. Preferiblemente, la mutación es indicativa de una resistencia a la ciprofloxacina (CP), levofloxacina (LVX) y tetraciclina (TE).

30 En una modalidad de la invención, la mutación en el gen ycgK es H63Q. Preferiblemente, la mutación es indicativa de una resistencia a ciprofloxacina (CP), levofloxacina (LVX), cefazolina (CFZ), tetraciclina (TE) y cefuroxima (CRM).

En una modalidad de la invención, la mutación en el gen ycgB es I497L. Preferiblemente, la mutación es indicativa de una resistencia a ciprofloxacina (CP), levofloxacina (LVX), cefazolina (CFZ) y tetraciclina (TE).

35 En una modalidad de la invención, la mutación en el gen valS se selecciona del grupo que consiste en D469E, N444K, Y439*, Y413 *, K277N, D244E, Y242* y D176E, en donde el asterisco designa una mutación que conduce a un nuevo codón de parada. Preferentemente en esta modalidad, una mutación seleccionada del grupo que consiste en D469E, N444K, Y439*, Y413 *, K277N y D176E es indicativa de una resistencia a ciprofloxacina (CP), levofloxacina (LVX) y cefazolina (CFZ). También preferentemente en esta modalidad, la mutación D244E es indicativa de una resistencia a ciprofloxacina (CP), levofloxacina (LVX) y gentamicina (GM) y la mutación Y242* es indicativa de una resistencia a ciprofloxacina (CP), levofloxacina (LVX), cefazolina (CFZ) y tetraciclina (TE).

40 En una modalidad de la invención, la mutación en el gen yjjJ es H184R. Preferentemente, la mutación es indicativa de una resistencia a la ciprofloxacina (CP), levofloxacina (LVX) y cefazolina (CFZ).

45 En cualquier modalidad de la invención, la presencia de una mutación en un gen seleccionado de dicho grupo se prueba en relación con la cepa de referencia E. coli K12 subcepa DH10B.

50 La presente descripción describe adicionalmente un método para determinar un perfil de resistencia a antibióticos para un microorganismo bacteriano que pertenece a la especie E. coli que comprende las etapas de
 a) proporcionar una muestra que contenga o se sospeche que contiene el microorganismo bacteriano;
 b) determinar la presencia de una mutación en al menos un gen del microorganismo bacteriano seleccionado del grupo de genes enumerados en la Tabla 4;
 en donde la presencia de una mutación es indicativa de una resistencia a un fármaco antibiótico.

55 La tabla 4 se muestra a continuación:

abgB	mhpA	ybfD	yfcO	yncG	torZ	fadA	thiE
allA	mnmC	ybfQ	yfdF	yneK	uidB	fdx	thiM
argI	mukB	ycbF	yfdR	yphG	ycjX	fhuB	trpC
caiC	norW	ycbS	yfdX	zraS	ydiU	fhuC	udp
csiE	ompA	yccE	yfhM	agaD	yejA	fhuD	uxaA
cynX	ompC	yceH	ygbN	astE	yfbL	fmt	ybiB

ES 2 731 913 T3

	dadX	parC	ycgB	ygcQ	chbG	yfiK	gudP	ybiU
	elaD	pgaA	ycgK	ygeK	eutE	ygcR	helD	ydfI
5	fcl	potB	yciQ	ygeO	eutQ	ygcU	hrpB	ydgA
	fhuA	puuC	ydbD	ygiD	figL	ygfZ	ilvA	yecA
	flhA	puuE	yddV	yhaC	gcvP	ygiF	kdpD	yehT
10	flu	rem	ydeK	yhaI	gspO	ygjM	ldcA	yfcN
	frwC	rhcS	ydjO	yhdP	gudD	yjjU	lplA	yheN
	gyrA	rhcD	yeaU	yhgE	hemF	yjjW	menB	yhgF
15	hofB	Rz	yeaX	yhiJ	kdpE	yohG	metH	yhhQ
	htrL	stfR	yeeJ	yjbl	ldrA	ypdB	pbpC	yhjE
	hybA	tilS	yefM	yjcF	livG	yqjA	purH	yjjG
20	hyfB	valS	yegE	yjfF	murB	yrfB	purK	ynfA
	hyfG	xseA	yegl	yjfZ	murP	ytfG	purL	
	hyfl	yach	yehB	yjgL	nepl	aspS	queF	
25	hypF	yafE	yehI	yjgN	pphB	birA	rhaA	
	ilvY	yafT	yehM	yjhS	ptrB	cysD	rhaB	
	lsrC	yagR	yeil	yjjJ	rhaD	dapB	rplO	
	lsrF	yaiO	yfaL	ymdC	speC	dxs	srlD	
30	menE	ybbB	yfaW	ynbB	tiaE	eutA	thiC	

La presencia o ausencia de una mutación en estos genes se analiza en relación con la cepa de referencia E. coli K12 subcepa DH10B (consulte también información más detallada a continuación y en el Ejemplo 1). En una modalidad preferida de la descripción, la etapa b) comprende determinar la presencia de una mutación en al menos dos o más genes seleccionados del grupo de la Tabla 4, y en donde la presencia de una mutación en al menos dos genes es indicativa de una resistencia a un antibiótico.

En lugar de probar solo genes individuales o mutantes, una combinación de varias posiciones variantes puede mejorar la precisión de la predicción y reducir aún más los resultados falsos positivos que están influenciados por otros factores. Por lo tanto, en particular se prefiere determinar la presencia de una mutación en 2, 3, 4, 5, 6, 7, 8 o 9 (o más) genes seleccionados de la Tabla 4.

En una modalidad preferida adicional de la divulgación, el presente método comprende en la etapa b) determinar la presencia de una mutación en al menos un gen seleccionado del grupo de genes enumerados en la Tabla 5, y en donde la presencia de una mutación en dicho al menos un gen es indicativa de una resistencia a un fármaco antibiótico.

Los genes de acuerdo con la Tabla 5 no se han descrito anteriormente en el contexto de la resistencia a los antibióticos de la bacteria E. coli. Pueden usarse para la determinación de una resistencia a fármacos antibióticos de E. coli solo o en combinación con otros genes que se describen en la presente descripción.

	abgB	yegl	ymdC	ycjX	ldcA	yhjE
	frwC	yehM	ynbB	ydiU	lplA	yjjG
55	hofB	yeil	yncG	yfbL	menB	
	htrL	yfaW	yneK	yfiK	metH	
	hybA	yfcO	yphG	ygcR	pbpC	
60	hyfB	yfdF	zraS	ygcU	purH	
	hyfl	yfdR	agaD	ygfZ	purL	
	lsrF	yfdX	chbG	ygiF	queF	
65	potB	ygbN	eutE	ygjM	rhaB	

ES 2 731 913 T3

5	puuC	ygcQ	eutQ	yjjU	rplO	
	yafT	ygeK	flgL	yjjW	srlD	
	yagR	ygeO	gcvP	yohG	thiC	
	yaiO	ygiD	gspO	yqjA	thiE	
	ybbB	yhaC	gudD	yrfB	thiM	
10	ybfD	yhgE	hemF	ytfG	uxaA	
	ybfQ	yhiJ	ldrA	aspS	ybiB	
	ycbF	yjbl	livG	cysD	ybiU	
15	ycbS	yjcF	murP	eutA	ydfI	
	ycgB	yjfF	nepl	fadA	ydgA	
	ycgK	yjfZ	pphB	fdx	yecA	
20	yciQ	yjgL	ptrB	fhuC	yehT	
	yddV	yjgN	rhaD	gudP	yfcN	
	ydjO	yjhS	tiaE	helD	yheN	
25	yeaX	yjjJ	torZ	hrpB	yhgF	
	yeeJ	ynfA	uidB	kdpD	yhhQ	

Para E. coli, se identificaron 86 pares de posiciones genéticas y resistencia a los fármacos altamente significativas (tabla 2). Las 86 combinaciones corresponden a 35 posiciones genéticas, ya que los sitios suelen ser significativos para más de un solo fármaco. Lo más importante es que los sitios respectivos están ubicados en 9 genes: hofB, allA, mukB, ymdC, potB, ycgK, ycgB, valS, yjjJ. Por lo tanto, estos genes parecen ser críticos para la resistencia/susceptibilidad a los antibióticos. Todas las mutaciones identificadas conducen a alteraciones de aminoácidos, ya sea a un cambio de aminoácidos en la posición respectiva o la aparición de un nuevo codón de parada. Para obtener información más detallada, consulte el Ejemplo 1, a continuación.

Más general, la presente descripción en un aspecto adicional se refiere a un método para determinar la resistencia o susceptibilidad de un microorganismo bacteriano que pertenece a la especie E. coli a un fármaco antibiótico, que comprende:

- proporcionar una muestra que contenga o se sospeche que contiene el microorganismo bacteriano que pertenece a la especie E. coli;
- determinar a partir de dicha muestra una información de secuencia de ácido nucleico de al menos un gen seleccionado del grupo de hofB, allA, mukB, ymdC, potB, ycgK, ycgB, valS, y yjjJ; y
- a partir del análisis de dicha información genética determinar la resistencia o susceptibilidad al fármaco antibiótico.

En una modalidad adicional de la descripción, se determina la presencia de una mutación en al menos un gen seleccionado del grupo de hofB, allA, mukB, ymdC, potB, ycgK, ycgB, valS, yjjJ. Por lo tanto, puede analizarse la presencia de una mutación en al menos uno o 2, 3, 4, 5, 6, 7, 8 o 9 de estos genes.

En una modalidad adicional de la descripción, se determina la presencia de una mutación en al menos un gen seleccionado del grupo de la siguiente tabla 6. Con mayor preferencia, se determina el cambio exacto de aminoácidos que se indica en la Tabla 6.

Nombre del gen	Cambio de aminoácido	Nombre del gen	Cambio de aminoácido	
55	aspS	D382E	purK	N137D
	birA	Q113H	purL	D615E
	cysD	D232N	queF	K126E
60	dapB	N87K	rhaA	S406N
	dxs	A541T	rhaB	T407A
	eutA	A210V	rplO	K39N
	fadA	V387I	srlD	M54T
65	fhuB	G448V	thiE	A121E

ES 2 731 913 T3

	fhuC	A122V	thiE	R43Q
	fhuD	D76E	thiM	A122T
5	fmt	V30I	trpC	L378F
	gudP	A448V	udp	I147M
	gyrA	D87N;D87Y	uxaA	E236A
10	heiD	E671D	ybiB	G35S
	hrpB	A413T	ybiU	M419I
	hrpB	V240A	ydfI	A146V
15	ilvA	D401E	ydgA	F416L
	kdpD	E376D	yecA	I195V
	ldcA	R167Q	yehT	A106V
20	lplA	A279T	yfcN	I39V
	menB	T31A	yheN	Q49H
	metH	E1124;E1124D	yhgF	E737D
25	mukB	S1015N	yhhQ	R138H
	parC	S80I	yhjE	I323V
	parC	S80R	yjjG	A57V
	pbpC	H37Q	ynfA	T84S
30	purH	T366I		

Sorprendentemente, los inventores encontraron que existe una superposición de mutaciones en proteínas funcionalmente similares de *E. coli* y *K. pneumoniae*. Interesantemente, al considerar las proteínas que se asociaron significativamente con al menos un fármaco, los inventores encontraron una superposición de 1.746 proteínas (el mismo nombre oficial y más del 80 por ciento de positivos en BLAST en comparación por pares) que se ven afectadas en *E. coli*, así como en *K. pneumoniae*. Al extender el análisis a los cambios de AA exactos en estas proteínas, los inventores aún detectaron una superposición de 55 posiciones mutadas que son iguales en ambos organismos. Por lo tanto, los genes anteriores podrían constituir una base valiosa para la determinación del patrón de resistencia a los antibióticos en ambos microorganismos, *E. coli* y *K. neumonía*.

De acuerdo con un aspecto opcional de la descripción, la información de la secuencia de ácido nucleico puede ser la determinación de la presencia de un solo nucleótido en una sola posición en al menos un gen.

Por lo tanto, la descripción comprende un método en donde se detecta la presencia de un polimorfismo o mutación de un solo nucleótido en una única posición de nucleotídica.

Por ejemplo, esto puede hacerse en al menos un gen seleccionado del grupo de hofB, allA, mukB, ymdC, potB, ycgK, ycgB, valS y yjjJ. Por lo tanto, de acuerdo con un aspecto opcional de la descripción, la mutación es una mutación que se selecciona del grupo de mutaciones enumeradas en la tabla 2 (ver más adelante en el Ejemplo 1). Por lo tanto, la presente descripción incluye además un método para determinar un perfil de resistencia a antibióticos para un microorganismo bacteriano que pertenece a la especie *E. coli* que comprende los pasos de

- a) proporcionar una muestra que contenga o se sospeche que contiene el microorganismo bacteriano;
 - b) determinar la presencia de una mutación en al menos una posición como se identifica en la Tabla 2;
- en donde la presencia de una mutación es indicativa de una resistencia a un fármaco antibiótico.

La determinación puede hacerse en base a 1, 2, 3, 4, 5, 6, 7 y hasta las 35 posiciones genéticas identificadas en la Tabla 2.

Generalmente, el método de acuerdo con la presente descripción implica determinar la resistencia de *E. coli* a uno o más fármacos antibióticos. Estos fármacos incluyen, pero no se limitan a, antibióticos seleccionados del grupo que consiste en ampicilina sulbactam (AS), ampicilina (AM), clavulánico de amoxicilina (AUG), aztreonam (AZT), ceftriaxona (CAX), ceftazidima (CAZ), cefotaxime (CFT), cefepima (CPM), ciprofloxacina (CP), ertapenem (ETP), levofloxacina (LVX), cefuroxima (CRM), piperazilina tazobactam (P/T), trimetoprim sulfametil-oxazol (T/S), tobramycin (A/C), gentamicina (GM), cefazolina (CFZ), cefalotina (CF), imipenem (IMP), meropenem (MER) y tetraciclina (TE). Ver además la Tabla 1.

ES 2 731 913 T3

Los inventores sorprendentemente encontraron que las mutaciones en ciertos genes son indicativas no solo de una resistencia a un solo antibiótico, sino a grupos que contienen varios fármacos.

Por ejemplo, resultó que en el caso del grupo de antibióticos lactámicos, la presencia de una mutación en los siguientes genes: *chbG*, *eutQ*, *flgL*, *gudD*, *gyrA*, *ldrA*, *menE*, *murB*, *murP*, *nepl*, *parC*, *pphB*, *ptrB*, *rhaD*, *ydiU*, *yegE*, *yegl*, *yfbL*, *yfiK*, *ygcR*, *ygiF*, *ygjM*, *yohG* y/o *yrfB* es indicativo de la presencia de una resistencia contra los antibióticos de este grupo.

El grupo de antibióticos lactámicos comprende preferentemente A.S., AM, AUG, AZT, CFZ, CPE, CFT, CAZ, CAX, CRM, CF, CP, IMP, MER, ETP y/o P/T. El umbral del valor p para estos genes identificados es $\leq 10^{-45}$.

Está dentro del alcance de la presente descripción que la determinación anterior se realice en base a un solo gen o 2, 3, 4, etcétera genes de este grupo, sin embargo, se prefiere determinar una mutación en todos estos genes en relación con la cepa de referencia K12 subcepa DH10B (vea también más abajo para más información).

En una modalidad adicional de la presente descripción, el antibiótico se selecciona de los antibióticos quinolona o aminoglucósidos y se determina la presencia de una mutación en los siguientes genes: *agaD*, *chbG*, *eutE*, *eutQ*, *gcvP*, *gspO*, *gyrA*, *livG*, *menE*, *nepl*, *parC*, *speC*, *tiaE*, *torZ*, *uidB*, *yegE*, *yegl*, *yejA*, *ygcU*, *ygfz*, *ygiF*, *ygjM*, *yjjU*, *yjjW*, *ymdC*, *ypdB*, *ycjA*, y/o *ytfG*.

Los antibióticos de quinolona y aminoglucósidos se seleccionan preferentemente de CP, LVX, GM y TO.

Sorprendentemente, los genes relevantes se superponen completamente con respecto a la resistencia a los antibióticos quinolona y aminoglucósidos; el umbral del valor p para estos genes es $\leq 10^{-53}$. También aquí, está dentro del alcance de la presente descripción que la determinación se realice en base a un solo gen o en 2, 3, 4 o más genes de este grupo solamente, sin embargo, se prefiere determinar una mutación en todos estos genes en relación a la cepa de referencia K12 subcepa DH10B.

En una modalidad adicional de la presente descripción, el antibiótico se selecciona de tetraciclina y se determina la presencia de una mutación en al menos uno o más de los siguientes genes: *astE*, *chbG*, *eutQ*, *flgL*, *gudD*, *gyrA*, *hemF*, *hypF*, *kdpE*, *ldrA*, *menE*, *murB*, *murP*, *nepl*, *ompC*, *parC*, *pphB*, *ptrB*, y/o *rhaD*. El umbral del valor p es $\leq 10^{-47}$.

En otra modalidad adicional de la presente descripción, el antibiótico se selecciona de trimetoprim sulfametoxazol y se determina la presencia de una mutación en al menos uno o más de los siguientes genes: *astE*, *chbG*, *eutQ*, *flgL*, *gudD*, *gyrA*, *ldrA*, *menE*, *murB*, *nepl*, *parC*, *ycjx*, *ydiU*, *yegE*, *yfiK*, *ygcR*, *ygiF*, y/o *yrfB*. El umbral del valor p es $\leq 10^{-48}$.

En una modalidad preferida de la presente descripción, el método de la presente descripción comprende determinar una mutación, en donde la mutación se selecciona del grupo de mutaciones enumeradas en la Tabla 7. La tabla 7 se muestra a continuación:

Posición genómica	Terapia	Referencia	Alt	AA	Alt AA	Gen	Cambio
37032	P/T	C	T	V	I	<i>caiC</i>	V270I
206427	AM	C	A, G	T	R, K	<i>yafE</i>	T133R;T133K
319290	P/T	A	T	D	E	<i>yaiO</i>	D36E
1181357	AM	C	T	T	M	<i>yceH</i>	T178M
1368519	P/T	G	T	AA	S	-	A137S
1516808	P/T	G	A, T	AA	T, S	<i>stfR</i>	A114T;A114S
1517573	P/T	G	C	E	Q	<i>stfR</i>	E369Q
1567286	P/T	G	A	AA	T	<i>ynbB</i>	A148T
1615473	AZT, CAX	A	C, T	E	V, A	<i>yncG</i>	E203V;E203A
1684413	AM	C	T	M	I	<i>ydeK</i>	M441I
1974644	A/S	T	A, C	C	R, S	<i>yeaX</i>	C69R;C69S
2052365	ETP	A	T, C	I	Codón de parada, M	<i>flhA</i>	1427;I427M
2178525	P/T	C	T	G	D	<i>yefM</i>	G74D

ES 2 731 913 T3

	2216164	P/T	C	T, G	R	K, T	fcl	R20K;R20T
5	2233638	ETP, P/T	G	A, T	L	F, Codón de parada	yegE	L447F;L447
	2428172	CP, LVX	C	A, T	D	N, Y	gyrA	D87N;D87Y
	2428183	A/S, AM,	G	A	S	L	gyrA	S83L
		AZT,						
10		CAX,						
		CAZ,						
		CFT,						
15		CPE,						
		CRM, GM,						
		T/S, TO						
20	2463877	AUG	A	G	V	A	menE	V46A
	2565236	ETP	G	T, A	A	S, T	yfdR	A156S;A156T
	2725302	P/T	G	A	M	I	xseA	M428I
25	2755319	P/T	T	C	M	T	csiE	M33T
	2924554	P/T	A	T	T	S	norW	T27S
	3240296	P/T	G	A, C	F	Codón de parada, L	hybA	F204;F204L
30	4054212	A/S	C	A, T	E	Codón de parada, D	ilvY	E184;E184D
	4525576	AM	T	C	I	V	yjfZ	I78V
	4553471	ETP	C	A, T	L	I, F	yjfF	L20I;L20F
	4575887	P/T	T	C, G	L	R, P	yjgL	L207R;L207P
35	4636902	AM	G	A	A	V	-	A175V

Los inventores encontraron que, aparte de los genes anteriores indicativos de una resistencia a los antibióticos, también los polimorfismos de un solo nucleótido (=SNP) pueden tener un alto significado para la presencia de una resistencia contra los antibióticos definidos. El análisis de estos polimorfismos a nivel de nucleótidos puede mejorar y acelerar aún más la determinación de una resistencia a los fármacos antibióticos en E. coli.

Por ejemplo, una resistencia de E. coli contra el fármaco antibiótico AM puede determinarse por la presencia de un polimorfismo de un solo nucleótido en al menos uno, por ejemplo 1, 2, 3, 4, 5 o 6 de las siguientes posiciones de nucleótidos: 2428183, 4525576, 1684413, 4636902, 1181357, 206427.

En una modalidad de la presente descripción, el fármaco antibiótico es A/S y se detecta un SNP en al menos uno, por ejemplo 1, 2 o 3 de las siguientes posiciones de nucleótidos: 2428183, 4054212, 1974644.

En una modalidad adicional de la presente descripción, el fármaco antibiótico es AUG y se detecta una mutación en la siguiente posición de nucleótido: 2463877.

Para una resistencia al fármaco antibiótico AZT, se detecta una mutación en al menos una de las siguientes posiciones de nucleótidos: 2428183, 1615473.

En una modalidad aún más preferida de la presente descripción, el fármaco antibiótico es CAX y se detecta una mutación en al menos una de las siguientes posiciones de nucleótidos: 2428183, 1615473.

Una resistencia a los fármacos antibióticos CFT, CP, CPE, CRM, GM, LVX, TO, T/S o CAZ puede detectarse por una mutación en la posición del nucleótido 2428183.

Cuando el fármaco antibiótico es ETP, se detecta una mutación en al menos una, por ejemplo 1, 2, 3 o 4 de las siguientes posiciones de nucleótidos: 2052365, 2233638, 4553471, 2565236.

En una modalidad adicional de la presente descripción, el fármaco antibiótico es P/T y se detecta una mutación en al menos una, por ejemplo 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13 o 14 de las siguientes posiciones de nucleótidos: 2233638,

2216164, 2725302, 1567286, 2755319, 319290, 3240296, 1517573, 2178525, 2924554, 1516808, 37032, 1368519, 4575887.

5 Preferentemente, la resistencia al fármaco antibiótico respectivo se prueba de acuerdo con los diagramas de decisión de las Figuras 5-20.

10 Un diagrama de decisión o "árbol de decisión" es un gráfico similar a un árbol para tareas de predicción, por ejemplo, clasificación. Dado un conjunto de datos que consiste de una serie de muestras con valores de características (cualquier medida, aquí SNP) y etiquetas de clase (aquí resistente/no resistente a un determinado fármaco), un árbol de decisión modela el proceso de decisión de inferir la etiqueta de clase de la muestra a partir de sus valores característicos.

15 Para construir el modelo se usa un conjunto de datos dado (como se describió anteriormente): Entre todas las características (SNP) y sus posibles valores (bases de ADN A, C, T y G), se selecciona el valor de la característica que logra la separación óptima de la muestra con respecto a las etiquetas de muestra dadas. Idealmente, este sería un SNP cuyo valor para muestras no resistentes es diferente con respecto a las muestras resistentes. El valor de la característica seleccionada se convierte en la raíz del árbol (el primer nodo del árbol, a menudo dibujado en la parte superior) y las muestras se dividen de acuerdo con esa característica, es decir, las muestras que tienen ese valor de característica y las muestras con otro valor. Los subconjuntos resultantes de muestras forman nuevos nodos y el proceso de selección y división de características se repite para cada uno de ellos por separado. Este procedimiento se detiene si se cumple un criterio específico (por ejemplo, no se logra ninguna mejora adicional o se alcanza el tamaño máximo del árbol).

La representación gráfica utilizada se define de la siguiente manera:

La raíz del árbol siempre se dibuja en la parte superior. Cada nodo contiene la siguiente información:

- 25 • Su característica y su(s) valor(es) se dibujan debajo del nodo, por ejemplo, SNP 2428183 = G.
- Etiqueta de clase: 0 = no resistente, 1 = resistente.
- Distribución de clase: La proporción de muestras contenidas en ese nodo que pertenece a la clase 0 o 1.
- Proporción de muestras contenidas en ese nodo (w.r.t. al número de muestra utilizado para construir el árbol).
- Color: verde = 0, azul = 1, cuanto más intenso sea el color, mayor será la certeza de la etiqueta de la clase elegida.

30 Generalmente, el modelo se basa en el llamado conjunto de entrenamiento y su poder de predicción se prueba en el denominado conjunto de prueba (para evaluar el rendimiento del modelo en datos invisibles). Ambos conjuntos de datos deben ser independientes y no tienen intersección. Sin embargo, si el conjunto de datos disponible no es lo suficientemente grande como para formar un conjunto de datos de prueba y entrenamiento lo suficientemente grande, aplicamos un procedimiento llamado validación cruzada de k-veces (CV): Dividimos nuestro conjunto de datos en k subconjuntos de igual tamaño, luego cada uno de los k subconjuntos se usa una vez como datos de prueba y el resto como datos de entrenamiento. El árbol final se basa en todo el conjunto de datos, por lo que el CV solo se utiliza para estimar el rendimiento del modelo final.

La clasificación de una nueva muestra funciona de la siguiente manera:

- 40 • Uno comienza en la raíz del árbol: se comprueba el valor del atributo raíz en la muestra. Si el valor es igual al valor raíz, entonces uno va a la izquierda al siguiente nodo. De lo contrario, uno va a la derecha.
- Se comprueba el valor del atributo de nodo actual en la muestra y se decide nuevamente si se va a la izquierda o a la derecha. Y así.
- 45 • El proceso se detiene si uno está en un nodo hoja (nodo terminal, nodo sin bordes salientes). La muestra obtiene la misma etiqueta que el nodo hoja.

50 De acuerdo con un aspecto opcional de la descripción, una mutación detectada es una mutación que conduce a una secuencia de aminoácidos alterada en un polipéptido derivado de un gen respectivo en el que se encuentra la mutación detectada. De acuerdo con este aspecto, la mutación detectada conduce a una versión truncada o polipeptídica (en la que se crea un nuevo codón de parada por la mutación) o una versión mutada del polipéptido que tiene un cambio de aminoácidos en la posición respectiva.

55 De acuerdo con un aspecto opcional de la descripción, determinar la información de la secuencia de ácido nucleico o la presencia de una mutación comprende determinar una secuencia parcial o una secuencia completa de al menos un gen.

60 De acuerdo con un aspecto opcional de la descripción, determinar la información de la secuencia de ácido nucleico o la presencia de una mutación comprende determinar una secuencia parcial o completa del genoma de dicho microorganismo bacteriano, en donde dicha secuencia parcial o completa del genoma comprende al menos una secuencia parcial de dicho al menos un gen.

De acuerdo con un aspecto opcional de la descripción, la muestra es una muestra de paciente (aislado clínico).

65 De acuerdo con un aspecto opcional de la descripción, la determinación de la información de la secuencia de ácido nucleico o la presencia de una mutación comprende el uso de un método de secuenciación de próxima generación o secuenciación de alto rendimiento. De acuerdo con un aspecto adicional preferido de este aspecto de la descripción, se

determina una secuencia parcial o completa del genoma del organismo bacteriano utilizando un método de secuenciación de próxima generación o de secuenciación de alto rendimiento.

5 De acuerdo con un aspecto opcional de la descripción, el método de la descripción comprende además determinar la resistencia a 2, 3, 4, 5 o 6 antibióticos.

En un aspecto adicional, la presente descripción está dirigida a un método de diagnóstico para determinar una infección por E. coli resistente a antibióticos en un paciente, que comprende las etapas de:

10 a) obtener o proporcionar una muestra que contenga o se sospeche que contiene E. coli del paciente;
b) determinar la presencia de al menos una mutación en al menos un gen como se describió anteriormente, en donde la presencia de dicha al menos una mutación es indicativa de una infección por E. coli resistente a los antibióticos en dicho paciente.

15 En un aspecto adicional más, la presente descripción está dirigida a un método para tratar a un paciente que padece una infección por E. coli resistente a los antibióticos en un paciente:

20 a) obtener o proporcionar una muestra que contenga o se sospeche que contiene E. coli del paciente;
b) determinar la presencia de al menos una mutación en al menos un gen como se describe anteriormente, en donde la presencia de dicha al menos una mutación es indicativa de una resistencia a uno o más fármacos antibióticos;
c) identificar dichos al menos uno o más antibióticos;
d) seleccionar uno o más antibióticos diferentes de los identificados en la etapa c) y ser adecuado para el tratamiento de una infección por E. coli; y
e) tratar al paciente con dicho uno o más antibióticos.

25 De acuerdo con una modalidad preferida de la presente descripción, el paciente es un vertebrado, con mayor preferencia un mamífero y el más preferido un paciente humano.

30 Con respecto a la dosis del antibiótico, se refiere a los principios establecidos de farmacología en medicina humana y veterinaria. Por ejemplo, Forth, Henschler, Rummel "Allgemeine und spezielle Pharmakologie und Toxikologie", novena edición, 2005 pueden usarse como una guía. Con respecto a la formulación de un fármaco listo para usar, se hace referencia a "Remington, The Science and Practice of Pharmacy", 22ª edición, 2013.

Definiciones

35 A menos que se defina de cualquier otra manera, todos los términos técnicos y científicos usados en la presente descripción tienen el mismo significado que el comprendido comúnmente por el experto en la técnica a la que pertenece esta descripción.

40 El término "molécula de ácido nucleico" se refiere a una molécula polinucleotídica que tiene una secuencia definida. Comprende moléculas de ADN, moléculas de ARN, moléculas de análogos de nucleótidos y combinaciones y derivados de las mismas, tales como moléculas de ADN o moléculas de ARN con análogos de nucleótidos incorporados o ADNc.

45 El término "información de secuencia de ácido nucleico" se refiere a una información que puede derivarse de la secuencia de una molécula de ácido nucleico, como la secuencia en sí misma o una variación en la secuencia en comparación con una secuencia de referencia.

50 El término "mutación" se refiere a una variación en la secuencia en comparación con una secuencia de referencia. Dicha secuencia de referencia puede ser una secuencia determinada en un organismo de tipo silvestre predominante o un organismo de referencia, por ejemplo, una cepa o subcepa bacteriana definida y conocida. Una mutación es, por ejemplo, una eliminación de uno o varios nucleótidos, una inserción de uno o múltiples nucleótidos, o la sustitución de uno o múltiples nucleótidos, la duplicación de uno o una secuencia de múltiples nucleótidos, la translocación de uno o una secuencia de múltiples nucleótidos, y en particular, un polimorfismo de un solo nucleótido (SNP).

55 En el contexto de la presente descripción, una "muestra" es una muestra que comprende una molécula de ácido nucleico de un microorganismo bacteriano. Ejemplos de muestras son: células, tejidos, fluidos corporales, muestras de biopsia, sangre, orina, saliva, esputo, plasma, suero, sobrenadante de cultivo celular, muestra de hisopo y otros.

60 Los nuevos y altamente eficientes métodos de secuenciación de ácidos nucleicos denominados secuenciación de próxima generación han abierto la posibilidad de un análisis genómico a gran escala. El término "secuenciación de próxima generación" o "secuenciación de alto rendimiento" se refiere a las tecnologías de secuenciación de alto rendimiento que paralelizan el proceso de secuenciación, produciendo miles o millones de secuencias a la vez. Los ejemplos incluyen la secuenciación masiva de firmas paralelas (MPSS) Polony, 454 pirosecuenciación, Illumina (Solexa), la secuenciación SOLiD, la secuenciación de semiconductores de iones, la secuenciación de nanoball de ADN, la secuenciación de moléculas únicas de Helioscope (TM), la secuenciación de una sola molécula SMRT (TM), la molécula única Secuenciación en tiempo real (RNAP), secuenciación del ADN de Nanopore.

65

Antes de que se describa con detalle ilustrativo, debe entenderse que esta descripción no se limita a las partes componentes particulares de las etapas del proceso de los métodos descritos en la presente descripción ya que dichos métodos pueden variar. Se debe también entender que la terminología usada en la presente descripción es para propósitos de describir las modalidades particulares de la presente descripción solamente, y no pretende ser limitativa. Debe señalarse que, tal como se usa en esta descripción y las reivindicaciones adjuntas, las formas singulares "un," "una" y "el/la" incluyen referentes en singular y/o plural a menos que el contexto claramente lo indique de cualquier otra manera. Por ejemplo, el término "un" tal como se usa en la presente descripción puede entenderse como una sola entidad o en el significado de "una o más" entidades. También debe entenderse que las formas plurales incluyen referentes singulares y/o plurales, a menos que el contexto indique claramente lo contrario. Además, debe entenderse que, en el caso de que se den intervalos de parámetros delimitados por valores numéricos, se considera que los intervalos incluyen estos valores de limitación.

Leyendas de las figuras

- Las Figuras 1-4 están relacionadas con el Ejemplo 2, más abajo.
- Figura 1: Una tabla de contingencia ilustrativa para el cálculo de la prueba exacta de Fisher y las mediciones de precisión, sensibilidad, especificidad, valor predictivo positivo (PPV) y valor predictivo negativo (NPV). Los números se dan para el cambio de aminoácidos S83L (GyrA) y Ciprofloxacina.
- Figura 2: Resumen de los valores medios de CIM para la ciprofloxacina para muestras que no tienen mutación en GyrA (S83, D87) y ParC (S80), ya sea una mutación en GyrA y no ParC, ambas mutaciones en GyrA y no ParC, o las tres mutaciones.
- Figura 3: Panel A: gráfico de barras de genes con mayor número de sitios significativos. El cuadro de barras del panel B. detalla los genes con mayor número de sitios correlacionados con al menos 3 fármacos. Panel C. Diagrama de dispersión que muestra para cada gen el número de sitios significativos correlacionados con al menos 3 fármacos en función del número total de sitios significativos en el gen. Los genes coloreados representan aquellos con números absolutos más altos (yfaL, yehL, YjgL), con una mayor frecuencia de resistencia correlacionada con al menos 3 fármacos (yjcN) y genes con sitios significativos más bajos en al menos 3 fármacos (fhuA, yeeJ). Panel D. A lo largo de la trama del gen para yjcN. Los sitios significativos a lo largo de la secuencia genética se presentan como puntos, el eje y muestra el número de clases de fármacos significativas para el sitio respectivo. Más abajo, se muestra un llamado diagrama de serpiente de la proteína transmembrana, se indican los aminoácidos afectados.
- Figura 4: Panel A: diagrama de red que muestra los fármacos como rectángulos y genes con mayor o menor cobertura si la resistencia para el fármaco respectivo se muestra como círculos. mmuP, mmuM, yiel, insN-1 corresponden a conteos más altos de lectura en el caso de aislamientos resistentes, mientras que los genes verdes corresponden a una cobertura más baja. Panel B y C: dos ejemplos de gráficos de cromosomas. Cada muestra está representada por una línea, las líneas negras corresponden a las líneas no resistentes y las líneas grises a los aislados resistentes.
- Figura 5: Diagrama de decisión para la ampicilina.
- Figura 6: Diagrama de decisión para la ampicilina sulbactam.
- Figura 7: Diagrama de decisión para el clavulánico amoxicilina.
- Figura 8: Diagrama de decisión para el aztreonam.
- Figura 9: Diagrama de decisión para la ceftriaxona.
- Figura 10: Diagrama de decisión para la ceftazidima.
- Figura 11: Diagrama de decisión para la cefotaxima.
- Figura 12: Diagrama de decisión para la ciprofloxacina.
- Figura 13: Diagrama de decisión para la cefepima.
- Figura 14: Diagrama de decisión para la cefuroxima.
- Figura 15: Diagrama de decisión para el ertapenem.
- Figura 16: Diagrama de decisión para la gentamicina.
- Figura 17: Diagrama de decisión para la levofloxacina.
- Figura 18: Diagrama de decisión para la piperacilina tazobactam.
- Figura 19: Diagrama de decisión para la tobramicina.
- Figura 20: Diagrama de decisión para trimetoprim sulfametoxazol.

EJEMPLOS

Ejemplo 1:

Aquí, se identificó una colección única de genes que permiten determinar la resistencia de un microorganismo bacteriano a los antibióticos de uso común.

Se secuenció una cohorte única de muestras bacterianas obtenidas de 150 aislados clínicos para comprender los mecanismos de resistencia genética mediante el uso de la secuenciación de alto rendimiento. Paralelamente, se aplicaron pruebas de resistencia clásicas con 21 fármacos o combinaciones de fármacos (Tabla 1)

Tabla 1: Fármacos antibióticos

65

ES 2 731 913 T3

	Medicamento	Drugbank ID	Abreviatura
	Amoxicilina	DB00766	
5	Clavulanato	DB01060	AUG
	Ampicilina	DB00415	AM
	Ampicilin Sulbactam	DB00415	A/S
10	Aztreonam	DB00355	AZT
	Cefazolina	DB01327	CFZ
	Cefepima	DB01413	CPE
15	Cefotaxima	DB00493	CFT
	Ceftazidima	DB00438	CAZ
	Ceftriaxona	DB01212	CAX
20	Cefuroxima	DB01112	CRM
	Cefalotina	DB00456	CF
	Ciprofloxacina	DB00537	CP
25	Gentamicina	DB00798	GM
	Imipenem	DB01598	IMP
	Levofloxacina	DB01137	LVX
30	Piperacilina	DB00319	
	Tazobactam	DB01606	P/T
	Tetraciclina	DB00759	TE
	Tobramicina	DB00684	TO
35	Trimetoprima	DB00440	
	Sulfametoxazol	DB01015	T/S
	Meropenem	DB00760	MER
40	Ertapenem	DB00303	ETP

Las cepas de *E. coli* que se analizaron se sembraron en placas de agar y se incubaron en condiciones de crecimiento durante 24 horas. Luego, las colonias se seleccionaron e incubaron en medio de crecimiento en presencia de un antibiótico dado en series de dilución en condiciones de crecimiento durante 16-20 horas. El crecimiento bacteriano se determinó observando la turbidez.

Se buscaron las siguientes mutaciones que están altamente correlacionadas con los resultados de la prueba de resistencia fenotípica.

Para la secuenciación, las muestras se prepararon usando una preparación de la biblioteca Nextera, seguido de la secuenciación multiplexada usando el sistema Illuminat HiSeq 2500, secuenciación final pareada. Los datos se mapearon con BWA (Li H. y Durbin R. (2010) Fast and accurate long-read alignment with Burrows-Wheeler Transform. *Bioinformatics*, Epub. [PMID: 20080505]) y los SNP se llamaron usando SNP samtools (Li H.*, Handsaker B.*, Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin R. y 1000 Genome Project Subgrupo de procesamiento de datos (2009) The Sequence alignment/map (SAM) format and SAMtools. *Bioinformatics*, 25, 2078-9. [PMID: 19505943]).

La secuencia de referencia se obtuvo de *Escherichia coli*. cepa K-12 subcepa DH10B:
 LOCUS CP000948 4686137 pb ADN circular BCT 05-JUN-2008
 DEFINICIÓN cepa de *Escherichia coli* K12 subcepa DH10B, genoma completo
 Número de acceso CP000948
 VERSIÓN CP000948.1 GI:169887498
 DBLINK BioProyecto: PRJNA20079
 PALABRAS CLAVE.
 FUENTE *Escherichia coli* cepa K-12 subcepa DH10B
 ORGANISMO *Escherichia coli* cepa K-12 subcepa DH10B
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; *Escherichia*.

Referencia 1 (bases 1 a 4686137)

AUTORES Durfee,T., Nelson,R., Baldwin,S., Plunkett,G. III, Burland,V., Mau,B., Petrosino,J.F., Qin,X., Muzny,D.M., Ayele,M., Gibbs,R.A., Csorgo,B., Posfai,G., The inventors-in-stock,G.M. y Blattner,F.R.
 TÍTULO The complete genome sequence of Escherichia coli DH10B: insights into the biology of a laboratory workhorse
 JOURNAL J. Bacteriol. 190 (7), 2597-2606 (2008)
 PUBMED 18245285

Referencia 2 (bases 1 a 4686137)

AUTORES Plunkett,G. III.
 TÍTULO Entrega Directa
 Entrega a la revista (20-FEB-2008) Department of Genetics and Biotechnology, University of Wisconsin, 425G Henry Mall, Madison, WI 53706, USA
 COMENTARIO DH10B y DH10B-T1R están disponibles en Invitrogen Corporation (<http://www.invitrogen.com>).

Las mutaciones se ajustaron a los genes y se calcularon los cambios de aminoácidos. Usando diferentes algoritmos (SVM, modelos de homología), se calcularon mutaciones que conducen a cambios de aminoácidos con probable patogenicidad/resistencia. Se excluyeron las variantes conocidas de la base de datos swissprot y se seleccionaron todas las variantes en los genes respectivos.

Como se indicó anteriormente, para E.coli se identificaron 86 pares de posiciones genéticas de resistencia a fármacos extremadamente importantes (tabla 2). Las 86 combinaciones corresponden a 35 posiciones genéticas, ya que los sitios suelen ser significativos para más de un solo fármaco. Lo más importante es que los sitios respectivos están ubicados en 9 genes: hofB, allA, mukB, ymdC, potB, ycgK, ycgB, valS, yjjJ. Por lo tanto, estos genes parecen ser críticos para la resistencia/susceptibilidad a los antibióticos. Todas las mutaciones identificadas conducen a alteraciones de aminoácidos, ya sea a un cambio de aminoácidos en la posición respectiva o la aparición de un nuevo codón de parada. Por lo tanto, se detectaron variantes relacionadas con la resistencia para los siguientes 6 antibióticos: CP, LVX, TE, CFZ, CRM, GM.

Tabla 2: Mutaciones identificadas

Posición genómica	Terapia	Valor-p	Posición del gen	Ref	Alt	AA	Alt AA	Gen	Cambio
90064	CP	2,6527E-22	1140	C	T	W	C	hofB	W380C
90064	LVX	5,2992E-20	1140	C	T	W	C	hofB	W380C
471036	CP	2,6527E-22	31	G	A	E	K	allA	E11K
471036	LVX	5,2992E-20	31	G	A	E	K	allA	E11K
1030161	CP	2,6527E-22	685	C	T	R	C	mukB	R229C
1030161	LVX	5,2992E-20	685	C	T	R	C	mukB	R229C
1161719	CP	4,7722E-13	697	A	C	N	H	ymdC	N233H
1161719	LVX	8,8353E-13	697	A	C	N	H	ymdC	N233H
1161764	CP	2,6527E-22	742	C	T	R	C	ymdC	R248C
1161764	LVX	5,2992E-20	742	C	T	R	C	ymdC	R248C
1238314	CP	3,4979E-18	799	A	G	L	V	potB	L267V
1238314	LVX	7,91E-17	799	A	G	L	V	potB	L267V

ES 2 731 913 T3

5	1238314	TE	2,5459E-07	799	A	G	L	V	potB	L267V
	1239076	CP	2,6527E-22	37	C	T	V	F	potB	V13F
10	1239076	LVX	5,2992E-20	37	C	T	V	F	potB	V13F
	1266748	CP	8,9009E-16	189	A	G	H	Q	ycgK	H63Q
15	1266748	LVX	1,2635E-14	189	A	G	H	Q	ycgK	H63Q
	1266748	CFZ	4,4331E-10	189	A	G	H	Q	ycgK	H63Q
20	1266748	TE	6,1345E-10	189	A	G	H	Q	ycgK	H63Q
	1266748	CRM	6,0841E-07	189	A	G	H	Q	ycgK	H63Q
25	1266829	CP	2,6527E-22	108	G	A	S	R	ycgK	S36R
	1266829	LVX	5,2992E-20	108	G	A	S	R	ycgK	S36R
30	1275217	LVX	1,2048E-15	1489	T	C	I	L	ycgB	I497L
	1275217	CP	1,5912E-15	1489	T	C	I	L	ycgB	I497L
35	1275217	TE	1,8029E-07	1489	T	C	I	L	ycgB	I497L
	1275217	CFZ	6,7512E-07	1489	T	C	I	L	ycgB	I497L
40	1275307	CP	2,6527E-22	1399	G	A	L	M	ycgB	L467M
	1275307	LVX	5,2992E-20	1399	G	A	L	M	ycgB	L467M
45	4582262	CP	3,7422E-11	1407	G	A	D	E	valS	D469E
	4582262	LVX	7,4232E-10	1407	G	A	D	E	valS	D469E
50	4582262	CFZ	7,6762E-07	1407	G	A	D	E	valS	D469E
	4582280	CP	1,1193E-12	1389	A	G	D	E	valS	D463E
55	4582280	LVX	1,2386E-12	1389	A	G	D	E	valS	D463E
	4582301	LVX	1,6858E-16	1368	T	C	K	N	valS	K456N
60	4582301	CP	4,0168E-16	1368	T	C	K	N	valS	K456N
	4582313	LVX	3,6594E-16	1356	G	A	D	E	valS	D452E
65	4582313	CP	2,3782E-15	1356	G	A	D	E	valS	D452E

ES 2 731 913 T3

5
10
15
20
25
30
35
40
45
50
55
60
65

4582337	CP	3,3749E-11	1332	G	A	N	K	valS	N444K
4582337	LVX	1,2081E-10	1332	G	A	N	K	valS	N444K
4582337	CFZ	3,6214E-08	1332	G	A	N	K	valS	N444K
4582352	CP	4,0453E-16	1317	A	G	Y	*	valS	Y439*
4582352	LVX	5,235E-16	1317	A	G	Y	*	valS	Y439*
4582352	CFZ	3,2243E-07	1317	A	G	Y	*	valS	Y439*
4582382	CP	3,7422E-11	1287	C	T	L	F	valS	L429F
4582382	LVX	7,4232E-10	1287	C	T	L	F	valS	L429F
4582430	CP	1,4858E-11	1239	G	A	Y	*	valS	Y413*
4582430	LVX	4,4757E-11	1239	G	A	Y	*	valS	Y413*
4582430	CFZ	2,3492E-07	1239	G	A	Y	*	valS	Y413*
4582838	LVX	1,1423E-14	831	T	C	K	N	valS	K277N
4582838	CP	1,2258E-14	831	T	C	K	N	valS	K277N
4582838	CFZ	7,2547E-08	831	T	C	K	N	valS	K277N
4582937	LVX	6,011E-18	732	A	G	D	E	valS	D244E
4582937	CP	2,912E-17	732	A	G	D	E	valS	D244E
4582937	GM	7,6578E-07	732	A	G	D	E	valS	D244E
4582943	CP	1,7589E-13	726	G	A	Y	*	valS	Y242*
4582943	LVX	1,3467E-11	726	G	A	Y	*	valS	Y242*
4582943	CFZ	1,2817E-09	726	G	A	Y	*	valS	Y242*
4582943	TE	1,3866E-07	726	G	A	Y	*	valS	Y242*
4582987	CP	1,7363E-22	682	G	A	L	M	valS	L228M
4582987	LVX	3,2E-21	682	G	A	L	M	valS	L228M
4583141	CP	3,6032E-18	528	A	G	D	E	valS	D176E
4583141	LVX	1,1539E-17	528	A	G	D	E	valS	D176E
4583141	CFZ	7,2389E-08	528	A	G	D	E	valS	D176E
4666362	CP	5,4173E-19	109	G	A	D	N	yjjJ	D37N

ES 2 731 913 T3

5	4666362	LVX	2,3023E-18	109	G	A	D	N	yjjJ	D37N
	4666405	CP	5,4173E-19	152	C	T	A	V	yjjJ	A51V
10	4666405	LVX	2,3023E-18	152	C	T	A	V	yjjJ	A51V
	4666461	CP	9,4295E-08	208	A	G	T	A	yjjJ	T70A
15	4666461	LVX	8,4496E-07	208	A	G	T	A	yjjJ	T70A
	4666768	CP	9,4295E-08	515	A	G	H	R	yjjJ	H172R
20	4666768	LVX	8,4496E-07	515	A	G	H	R	yjjJ	H172R
	4666804	CP	1,6956E-22	551	A	G	H	R	yjjJ	H184R
25	4666804	LVX	7,7611E-22	551	A	G	H	R	yjjJ	H184R
	4666804	CFZ	4,5083E-07	551	A	G	H	R	yjjJ	H184R
30	4666885	CP	9,4295E-08	632	A	G	Y	C	yjjJ	Y211C
	4666885	LVX	8,4496E-07	632	A	G	Y	C	yjjJ	Y211C
35	4667178	CP	9,4295E-08	925	C	G	Q	E	yjjJ	Q309E
	4667178	LVX	8,4496E-07	925	C	G	Q	E	yjjJ	Q309E
40	4667191	CP	2,6527E-22	938	G	A	R	H	yjjJ	R313H
	4667191	LVX	5,2992E-20	938	G	A	R	H	yjjJ	R313H
45	4667359	CP	9,4295E-08	1106	T	C	V	A	yjjJ	V369A
	4667359	LVX	8,4496E-07	1106	T	C	V	A	yjjJ	V369A
50	4667424	CP	2,6527E-22	1171	G	A	V	I	yjjJ	V391I
	4667424	LVX	5,2992E-20	1171	G	A	V	I	yjjJ	V391I
55	4667568	CP	1,2838E-17	1315	G	A	A	T	yjjJ	A439T
	4667568	LVX	2,3051E-17	1315	G	A	A	T	yjjJ	A439T

60

En la Tabla 2 las columnas se designan de la siguiente manera:

Pos genoma: posición genómica del SNP/variante en el genoma de referencia de E. coli (ver más abajo);

Terapia: la terapia con la cual la mutación está significativamente correlacionada, las terapias múltiples están en filas separadas (si un SNP está correlacionado con, por ejemplo, 4 terapias, esto lleva a 4 filas individuales);

65

Valor P: valor significativo calculado utilizando la prueba exacta de los pescadores;

Pos gen: posición de la mutación en el gen;

Ref: Referencia base, A, C, T, G;
 Alt: Base alternativa asociada a la resistencia;
 AA: aminoácido original;
 Alt A: aminoácido cambiado;
 Gen: gen afectado;
 Cambio: cambio de aminoácidos en la nomenclatura estándar;

El valor de P se calculó utilizando la prueba exacta de Fisher basada en la tabla de contingencia con 4 campos: #muestras tipo Resistente/silvestre; #muestras Resistente/mutante; #muestras no Resistentes/tipo silvestre; #muestras no Resistentes/mutantes

En la tabla 3, los genes y productos génicos identificados se enumeran e identifican por ID de gen del gen y (NCBI) Número de acceso dado de la proteína correspondiente

Tabla 3: Nombre de gen e identificador

Nombre del gen	Gen ID	Número de acceso
hofB	6061494	ACB01286.1
allA	6059827	ACB01630.1
mukB	6060547	ACB02124.1
ymdC	6059214	ACB02240.1
potB	6058608	ACB02318.1
ycgK	6058586	ACB02348.1
valS	6060190	ACB05239.1
yjjJ	6058313	ACB05313.1
ycgB	6058539	ACB02358.1

La prueba se basa en la distribución de las muestras en los 4 campos. La distribución uniforme no indica importancia, mientras que la agrupación en dos campos indica importancia.

Usando este enfoque, se identificaron 35 nuevas posiciones o mutaciones genéticas altamente significativas en 9 genes (hofB, allA, mukB, ymdC, potB, ycgK, ycgB, valS, yjjJ) que pueden usarse para y permiten la determinación de resistencia a los antibióticos de uso común. Todas las mutaciones altamente significativas descritas en este documento y enumeradas en la tabla 2 son mutaciones no conservativas que conducen a un cambio de aminoácidos o un nuevo codón de parada (designado con un símbolo "*" en la tabla 2), y por lo tanto a una proteína alterada. Por lo tanto, es probable que los 9 genes identificados desempeñen un papel importante en la resistencia a los antibióticos y sean objetivos putativos para el desarrollo de nuevos fármacos candidatos.

Ejemplo 2

En este ejemplo, los inventores evalúan la susceptibilidad genética de E. coli a 21 fármacos diferentes de cinco clases de fármacos (ver más abajo).

Métodos: Se realizó una prueba de susceptibilidad antimicrobiana (AST) para 1,162 aislamientos clínicos de E. coli con diferentes espectros de resistencia a 21 fármacos aprobados por la FDA y se secuenciaron genomas de todos los aislamientos. Las variantes genéticas se correlacionaron con los datos AST.

Resultados: Los inventores reportan que 25,744 sitios en el genoma de E. coli se correlacionan significativamente con la resistencia a los fármacos. Se alcanzó la mayor significancia para los fármacos Ciprofloxacina y Levofloxacina con respecto al cambio de aminoácidos S83L en GyrA ($p_{\text{Ciprofloxacina}}=10^{-235}$, precisión, especificidad y sensibilidad: 98%, 99% y 94%; $p_{\text{Levofloxacina}}=10^{-209}$, 97%, 98%, 93%), un objetivo para las quinolonas. La segunda asociación más significativa se observó para ParC, un segundo objetivo de quinolonas (cambio de AA S80I, $p_{\text{Ciprofloxacina}}=10^{-196}$ and $p_{\text{Levofloxacina}}=10^{-194}$). En particular, muchos cambios de AA asociados significativamente con la resistencia a múltiples fármacos se descubrieron en YigN. Al analizar la secuencia de cobertura en el nivel del genoma, los inventores identificaron una dependencia de la dosis de genes de varios genes, incluyendo *mmuP* y *mmuM*, que codifican un transportador putativo de S-metilmetionina y una homocisteína S-metiltransferasa. Ambos loci están asociados con resistencia contra las β -lactamas y quinolonas.

Conclusión: Los inventores aquí presentan una línea de análisis y selección de alto rendimiento para investigar la resistencia a los antibióticos en cepas de E. coli. Los resultados demuestran el potencial de las pruebas basadas en la

genética para predecir la susceptibilidad frente a los fármacos antimicrobianos. Además, se reportan nuevas correlaciones entre la dosis de genes y la resistencia.

5 Los inventores llevaron a cabo una evaluación sistemática utilizando *E. coli*. Específicamente, los inventores recolectaron 1,162 muestras de *E. coli* durante 22 años (1991-2013) en más de 60 institutos diferentes. Para estos aislados, los inventores llevaron a cabo AST estándar para 21 fármacos aprobados por la FDA y realizaron la secuenciación del genoma completo (WGS) para los mismos 1,162 aislamientos para construir una base de datos que revelaba sitios genéticos para predecir el AST a partir de datos genéticos.

10 Métodos

Cepas bacterianas

15 Los inventores seleccionaron 1,162 cepas de *E. coli* de la colección de cepas de microbiología en Siemens Healthcare Diagnostics (West Sacramento, CA) para las pruebas de susceptibilidad y la secuenciación del genoma completo.

Paneles de prueba de sensibilidad antimicrobiana

20 Se prepararon paneles AST de referencia congelados siguiendo las recomendaciones del Instituto de Normas de Laboratorio Clínico (CLSI) ¹⁶. Los siguientes agentes antimicrobianos (con concentraciones de µg/ml mostradas entre paréntesis) se incluyeron en los paneles: Amoxicilina/Clavulanato K (0.5/0.25-64/32), Ampicilina (0.25-128), Ampicilina/Sulbactam (0.5/0.25-64/32), Aztreonam (0.25-64), Cefazolina (0.5-32), Cefepime (0.25-64), Cefotaxima (0.25-128), Ceftazidima (0.25-64), Ceftriaxona (0.25-128), Cefuroxima (1-64), Cefalotina (1-64), Ciprofloxacina (0.015-8), Ertepenem (0.12-32), gentamicina (0.12-32), imipenem (0.25-32), levofloxacina (0.25-16), meropenem (0.12-32), piperacilina/tazobactam (0.25/4-256/4), tetraciclina (0.5-64), Tobramicina (0.12-32) y trimetoprim/sulfametoxazol (0.25/4.7-32/608). Antes de su uso con aislados clínicos, los paneles de AST se probaron con cepas de control de calidad. Los paneles de AST se consideraron aceptables para las pruebas con aislados clínicos cuando los resultados del control de calidad cumplieron con los rangos de control de calidad descritos por CLSI16.

30 Preparación del inóculo

Los aislados se cultivaron en agar de tripticasa de soja con 5% de sangre de oveja (BBL, Cockeysville, Md.) y se incubaron en aire ambiental a $35 \pm 1^\circ\text{C}$ durante 18-24 h. Las colonias aisladas (4-5 colonias grandes o 5-10 colonias pequeñas) se transfirieron a un agua estéril de inóculo de 3 ml (Siemens) y se emulsionaron a una turbidez final de un estándar de 0.5 McFarland. Se agregaron 2 ml de esta suspensión a 25 ml de agua de inóculo con Pluronic-F (Siemens). Utilizando el Inoculator (Siemens) específico para paneles de AST congelados, se transfirieron 5 µl de la suspensión celular a cada pocillo del panel de AST. Los paneles de AST inoculados se incubaron en aire ambiental a $35 \pm 1^\circ\text{C}$ durante 16-20 h. Los resultados del panel se leyeron visualmente y se determinaron las concentraciones inhibitorias mínimas (MIC).

40 Extracción de ADN

Se hicieron cuatro rayas de cada aislado bacteriano Gram-negativo cultivado en agar de soja tripticasa que contenía 5% de sangre de oveja y suspensiones celulares en tubos de recolección estériles de 1,5 ml que contenían 50 µl de Agua libre de nucleasas (AM9930, Life Technologies). Las muestras de aislamientos bacterianos se almacenaron a -20°C hasta la extracción del ácido nucleico. El estuche del Sistema de Preparación de Tejidos (TPS) (096D0382-02_01_B, Siemens) y el estuche de Reactivos para la Preparación de Tejidos VERSANT® (TPR) (10632404B, Siemens) se utilizaron para extraer el ADN de estos aislados bacterianos. Antes de la extracción, los aislados bacterianos se descongelaron a temperatura ambiente y se sedimentaron a 2000 G durante 5 segundos. Se utilizó el protocolo de extracción de ADN DNAext para la extracción completa de ácido nucleico total de 48 muestras aisladas y eluatos, 50 µl cada uno, en 4 horas. Los eluatos de ácido nucleico totales se transfirieron luego a placas de detección de qPCR de 96 pocillos (401341, Agilent Technologies) para los procesos de digestión con ARNasa A, cuantificación de ADN y estandarización de la concentración de ADN en placa. La ARNasa A (AM2271, Life Technologies), que se diluyó en agua libre de nucleasas siguiendo las instrucciones del fabricante, se agregó a 50 µl del eluido total de ácido nucleico para una concentración de trabajo final de 20 µg/ml. La enzima de digestión y la mezcla de eluato se incubaron a 37°C durante 30 minutos utilizando el instrumento de amplificación y detección Siemens VERSANT®. El ADN del eluato digerido con ARNasa se cuantificó utilizando el ensayo Quantico D™ PicoGreen dsDNA (P11496, Life Technologies) siguiendo las instrucciones del estuche de ensayo, y se determinó la fluorescencia en el instrumento de detección y amplificación Siemens VERSANT®. El análisis de los datos se realizó con Microsoft® Excel 2007. Se transfirieron 25 µl de los eluatos de ADN cuantificados a una nueva placa de PCR de 96 pocillos para la estandarización de la concentración de ADN en placa antes de la preparación de la biblioteca. Se utilizó tampón de elución del estuche TPR para ajustar la concentración de ADN. La placa de eluato de ADN estandarizada se almacenó luego a -80°C hasta la preparación de la biblioteca.

Secuenciación de nueva generación

65 Antes de la preparación de la biblioteca, se llevó a cabo un control de calidad del ADN bacteriano aislado utilizando un fluorómetro Qubit 2.0 (estuche de ensayo Qubit dsDNA BR, Life Technologies) y una TapeStation Agilent 2200 (cinta de

pantalla Genomic DNA, Agilent Technologies). Las bibliotecas de NGS se prepararon en formato de 96 pocillos utilizando el estuche de preparación de muestras de ADN NexteraXT y el estuche de índices NexteraXT para 96 índices (Illumina) según el protocolo del fabricante. Las bibliotecas de secuenciación resultantes se cuantificaron en un enfoque basado en qPCR utilizando el estuche MasterMix qPCR KAPA SYBR FAST (Peqlab) en un sistema de PCR en tiempo real ViiA 7 (Life Technologies). Se agruparon 96 muestras por línea para la secuenciación de extremos pareados (2x 100 pb) en los secuenciadores Illumina HiSeq2000 o HiSeq2500 utilizando TruSeq PE Cluster v3 y TruSeq SBS v3 sequencing chemistry (Illumina). Los parámetros de calidad de la secuenciación básica se determinaron utilizando la herramienta de control de calidad FastQC para datos de secuencia de alto rendimiento (Babraham Bioinformatics Institute).

10 Análisis de Datos

Los datos de la secuencia de extremos pareados sin procesar para las 1,162 muestras de *E. coli* se mapearon con la referencia DH10B de *E. coli* (NC_010473) (ver también arriba en el Ejemplo 1) con BWA 0.6.1.20. Los archivos SAM resultantes se clasificaron, se convirtieron a los archivos BAM y los duplicados de PCR se marcaron utilizando el paquete de herramientas Picard 1.104 (<http://picard.sourceforge.net/>). Se utilizó el Genome Analysis Toolkit 3.1.1 (GATK) 21 para llamar los SNP e indeles para bloques de 200 muestras de *E. coli* (parámetros: -ploidy 1 -glm BOTH -stand_call_conf 30 -stand_emit_conf 10). Los archivos VCF se combinaron en un solo archivo y se llevó a cabo un filtrado de calidad para los SNP (QD <2.0 || FS> 60.0 || MQ <40.0) e indeles (QD <2.0 || FS> 200.0). Las variantes detectadas se anotaron con SnpEff22 para predecir los efectos de codificación. Para cada posición anotada, se consideraron los genotipos de todas las muestras de *E. coli*. Las muestras de *E. coli* se dividieron en dos grupos, un grupo de baja resistencia (que tenía una concentración de MIC más baja para el fármaco considerado) y un grupo de alta resistencia (que tenía concentraciones de MIC más altas) con respecto a una cierta concentración de MIC (punto de ruptura). Para encontrar el mejor punto de ruptura, se evaluaron todos los umbrales y se calcularon los valores de p con la prueba exacta de Fisher basándose en una tabla de contingencia de 2x2 (número de muestras de *E. coli* que tienen el genotipo de referencia o variante frente al número de muestras que pertenecen a la resistencia alta y baja grupo). El mejor punto de interrupción calculado fue el umbral que produjo el valor p más bajo para una determinada posición genómica y un fármaco. Para análisis adicionales se consideraron las posiciones con alteraciones no sinónimos y valor de p <10⁻⁹. Sobre la base de la tabla de contingencia, se calcularon la precisión (ACC), la sensibilidad (SENS), la especificidad (SPEC) y los valores predictivos positivos/negativos (PPV/NPV) (Figura 1).

Dado que una posible razón para la resistencia del fármaco es la duplicación de genes, se evaluó la dependencia de la dosis de genes. Para cada muestra, la cobertura genómica para cada posición se determinó utilizando las herramientas BED. Se extrajeron 23 rangos de genes del conjunto de referencia NC_010473.gff y se calculó la cobertura mediana normalizada por gen. Para comparar los aislamientos de baja y alta resistencia, se calculó el valor del mejor área bajo la curva (AUC). Se consideraron grupos de al menos el 20% de todas las muestras que tenían una cobertura mediana mayor que cero para ese gen y que contenían más de 15 muestras por grupo para excluir los artefactos y los casos con AUC > 0.75 se evaluaron más a fondo.

40 Resultados

El objetivo de nuestro estudio fue demostrar la viabilidad de las pruebas genéticas de susceptibilidad antimicrobiana (GAST), verificar nuestro método para los mecanismos de resistencia conocidos y descubrir nuevos mecanismos. Los inventores realizaron AST basada en cultivo para 1,162 aislamientos de *E. coli* y 21 fármacos antimicrobianos pertenecientes a 5 clases de fármacos diferentes: β -lactamas, fluoroquinolonas, aminoglicósidos, tetraciclinas e inhibidores de la síntesis de folato. La lista completa de fármacos se muestra en la Tabla 1. Para los mismos 1,162 aislamientos de *E. coli*, se llevó a cabo la secuenciación del genoma completo utilizando el instrumento HiSeq2500 de Illumina.

50 Sitios más significativos en el genoma de *E. coli*

Para calcular las puntuaciones de significación de todo el genoma, los inventores mapearon los 1,162 genomas de *E. coli* con respecto a la cepa de referencia DH10B. Para cada posición genómica, los inventores determinaron la base para cada muestra y descubrieron 973,226 sitios que pasaron el filtrado de calidad y en los que al menos una muestra tenía una base no de referencia. Los sitios respectivos se correlacionaron con los datos de AST para los 21 fármacos mediante la prueba exacta de Fisher. Nuestro análisis reveló 25,744 sitios donde una mutación genética se correlacionó significativamente con al menos un fármaco (valor p <10⁻⁹) y dio lugar a un cambio en la secuencia de AA, incluida la mutación puntual y pequeñas inserciones y eliminaciones. Se alcanzó el mayor significado para el cambio de AA S83L en GyrA y el fármaco Ciprofloxacina (p=10⁻²³⁵). Cabe destacar que GyrA es uno de los objetivos de la ciprofloxacina. Para esta posición, tres cambios AA, S83L, S83W, S83A, se anotan en UniProt como una resistencia a las quinolonas. Para este sitio, solo se detectaron 5 falsos positivos (0,4%) y 18 falsos negativos (1,6%), mientras que 1,139 se identificaron correctamente, lo que corresponde a la precisión, especificidad y sensibilidad del 98,0%, 99,4% y 93,8%, respectivamente (Figura 1). De manera similar, el segundo sitio más significativo en GyrA, D87N/D87Y reveló solo 12 falsos positivos y 10 falsos negativos, el valor p respectivo fue 10⁻²⁰⁶ y la precisión del 98.1%. Nuevamente, para este sitio, el cambio D87N está anotado como que confiere resistencia a la quinolona en UniProt. Para el tercer y cuarto sitios más significativos, ubicados en el segundo objetivo de la Ciprofloxacina, ParC, (S80I, E84G), también se describieron variantes relacionadas con la resistencia. En la Figura 2, los inventores presentan los promedios y las desviaciones estándar de las MIC para

5 Ciprofloxacina para muestras que no tienen variante en GyrA (S83/D87) y ParC (S80), muestras que tienen solo una mutación en GyrA S83 o D87 y no en ParC, muestras que tienen ambas mutaciones en GyrA y no en ParC, y muestras que tienen las tres mutaciones. Curiosamente, los valores medios de MIC aumentan desde menos de 1.0 para mutantes nulos o simples hasta más de 7.8 para mutantes dobles o triples, lo que demuestra que es necesaria una combinación de mutaciones para alcanzar un nivel más alto de resistencia contra Ciprofloxacina en este caso.

10 Además de las mutaciones en los objetivos del fármaco de topoisomerasa tipo II (GyrA/ParC), mutaciones en los genes *ygiF* (A110T, $p=10^{-67}$, acc=86%, spec=89.5%, sens=69.9%) y *ygiM* (A68V, $p=10^{-63}$, acc=89.9%, spec=94.4%, sens=67.1%) también tienen un alto significado. En comparación con los cambios de AA descritos anteriormente, estos dos sitios demuestran una sensibilidad sustancialmente reducida y valores predictivos positivos (PPV). Mientras que el PPV para los cuatro cambios de AA en GyrA y ParC estuvo entre 94.8% y 98.2%, el PPV de estos dos cambios disminuye a 59.0% y 70.8%. Esto significa que la probabilidad de ser resistente dado el AA intercambiado es casi tan alta como la probabilidad de ser susceptible dado el AA intercambiado, limitando la probabilidad de que los respectivos cambios de AA sean causales.

15 Para descubrir otros cambios de AA que son potencialmente causantes de resistencia a los fármacos, los inventores filtraron la lista de todos los 25,744 sitios (al menos 150 aislamientos de *E. coli* resistentes llevan el cambio de AA, NPV > 50%, VPP > 75%). Este filtrado reveló 127 sitios candidatos (ver también la Tabla 4). Además de los cambios ya descritos en GyrA y ParC, los inventores descubrieron cambios de AA en YdjO asociados con la resistencia predicha a diferentes β -lactamas (V121E, S120C, V118F, I114V, K111E y D112N). Del mismo modo, para las lactamas, los inventores informan los cambios de AA en YcbS (E848Q, E848*), RhsC (R717Q, W492C), YcbQ (T86I), YagR (S274T) y YeaU (N293K). Finalmente, los inventores descubrieron cambios de AA relacionados con quinolonas, tetraciclina y lactamas en YhaL (en total 23 sitios diferentes). Además, los inventores calcularon el cambio de AA no sinónimos más significativo para cada fármaco (umbral de valor $p < 10^{-9}$). De 21 fármacos probados, solo dos (Imipenem, Meropenem) no se encontraron asociados con un cambio de AA con un valor de p tan bajo. Curiosamente, la mutación S83L en GyrA es el cambio predominante en 15 fármacos. Para los fármacos Ciprofloxacina y Levofloxacina, de los cuales GyrA es un objetivo, los valores p fueron sin embargo mucho más bajos que los valores p para esta mutación en asociación con los 13 fármacos restantes ($> 10^{-62}$ versus $< 10^{-209}$). Además, los inventores observaron nuevamente una disminución significativa en la sensibilidad y/o PPV en estos casos: o la sensibilidad o PPV es inferior al 55% para los fármacos, de los cuales GyrA no es el objetivo, lo que demuestra que estas medidas son efectivas para separar mutaciones Objetivos verdaderos de los demás.

Mutaciones en objetivos conocidos de fármacos

35 En 9 casos, los inventores detectaron mutaciones asociadas con fármacos en genes que también están codificando los objetivos de los respectivos fármacos. Esto incluye las mutaciones asociadas con Ciprofloxacina y Levofloxacina en GyrA (S83L, D87N, D87Y, D678E, E574D) y ParC (S80I, E84G, E84V, E84A, A192V, Q481H, A471G, T718A, Q198H), mutaciones asociadas con cefalotina en AmpC (K40R, I300V, T335I, A210P, Q196H, A236T, R248C), con Sulfametoxazol en FolC (A319T, R88C, G217S), con Cefazolina en MrcB (D839E, QQQP815Q, R556C) y PbpC (L357V, V348A, A15T, A217V, Q495L, V768F, A701E, K766R, K766T, T764S, T764A, R602L, E446G, R669H, A202T) y con Ceftazidima en PbpG (A28V).

Genes más afectados y los sitios resistentes a múltiples fármacos

45 Las mutaciones no se distribuyen uniformemente entre los genes de *E. coli*: por ejemplo, *yfaL*, *fhuA*, *yehI*, *yjgL* y *yeeJ* tienen más de 120 variantes no sinónimos por gen (Figura 3A); en *yfaL*, se descubrieron 182 cambios significativos. Para descubrir sitios que son relevantes para la resistencia a múltiples fármacos, los inventores calcularon el número de cambios de AA significativos en asociación con al menos 3 clases de fármacos (Figura 3B) y trazaron los recuentos de sitios respectivos para cada gen en la Figura 3C. En promedio, el 35% de todos los sitios significativos se asociaron con al menos tres fármacos. Mientras que tres genes, *yfaL*, *yehI* y *yjgL*, tuvieron el mayor número de cambios de AA, *yjgN* tuvo un número sustancialmente mayor de sitios asociados con resistencia a múltiples fármacos (53 de 64 sitios, 83%), mientras que *yeeJ* (15 de 122 sitios, 12%) y *fhuA* (12 de 166 sitios, 7%) tienen menos sitios relevantes para múltiples clases de fármacos de lo esperado. En *yjgN*, las posiciones significativamente asociadas con múltiples clases de fármacos se concentraron en las regiones terminales del gen (Figura 3D).

55 Análisis de Cobertura

60 Una razón potencial para la resistencia a los fármacos es la duplicación o eliminación de genes, que puede observarse en nuestro conjunto de datos mediante la inspección de la cobertura de lectura de diferentes genes en los grupos de aislados resistentes y susceptibles. Para estimar la diferencia en la cobertura, los inventores calcularon los valores de AUC para la cobertura mediana normalizada por gen en los dos grupos. En conjunto, los inventores descubrieron 23 casos de diferencias anormales en la cobertura de genes entre bacterias resistentes y susceptibles que resultaron en un AUC > 0.75 (Figura 4A). Los inventores reportan conexiones para tres β -lactamas y dos quinolonas. Los genes centrales son *mmuP* y *mmuM*, que codifican un putativo transportador de S-metilmetionina y una homocisteína S-metiltransferasa, respectivamente, para los cuales la cobertura es sustancialmente mayor en bacterias resistentes a los 5 fármacos. En las cepas resistentes a levofloxacina y ciprofloxacina, la proteína de membrana interna Yiel e InsN-1, un regulador del

elemento de inserción, también fueron más abundantes. En contraste, los genes que codifican las glucosiltransferasas YaiP, YaiO, la proteína de membrana externa NmpC y el represor transcripcional de unión a ADN MngR tenían menor cobertura en cepas resistentes a estos fármacos. Las Figuras 4B y 4C muestran un ejemplo de gráfico de cobertura para la cobertura menos abundante yaiP y la cobertura más abundante mmuP en cepas resistentes a la ciprofloxacina. Se alcanzó la mejor precisión diagnóstica para la ciprofloxacina y el gen mmuP, con un valor de AUC de 0.923, lo que demuestra que esta información cuantitativa permite una separación precisa entre cepas resistentes y susceptibles.

Discusión

El aumento considerable y continuo de infecciones causadas por patógenos resistentes a múltiples fármacos representa una gran amenaza para los pacientes, especialmente en entornos hospitalarios. El desarrollo de nuevos fármacos es una empresa larga y costosa, y se estancó en los últimos años a pesar del aumento de las inversiones en investigación y desarrollo. El anuncio de la FDA en septiembre de 2012 para formar un grupo de trabajo interno para apoyar el desarrollo de nuevos fármacos antimicrobianos enfatiza la importancia de este tema. Hasta que estos fármacos estén disponibles, los inventores deben aprender a aplicar los disponibles de la manera más eficiente. La prescripción abundante de antibióticos de amplio espectro promueve el desarrollo de resistencia a múltiples fármacos, por lo que se necesita una selección más cuidadosa de los fármacos. Por lo tanto, los inventores necesitan métodos que puedan estratificar rápidamente a los pacientes y proporcionarles la terapia óptima. Identificar los loci genéticos en el agente infeccioso que son predominantemente responsables de una resistencia o susceptibilidad observada es un punto crucial para esto.

Aquí, los inventores analizaron 1,162 aislados clínicos de *E. coli* en cuanto a su susceptibilidad a 21 fármacos aprobados por la FDA y combinaron esta información con datos de NGS del genoma completo para identificar posibles variantes que podrían ser causantes de los patrones de resistencia observados. En total, los inventores encontraron 25,744 sitios significativos (valor $p < 10^{-9}$). El método identificó correctamente los objetivos farmacológicos ya conocidos en nueve combinaciones de gen/fármaco: *gyrA* (Ciprofloxacina, Levofloxacina), *parC* (Ciprofloxacina, Levofloxacina), *ampC* (Cefalotina), *folC* (Trimetoprim-sulfametoxazol), *mrcB* (Cefazolina), *pbpC* (Cefazolina), y *pbpG* (Ceftazidima). Para identificar otros sitios potenciales que podrían ser objetivos secundarios de fármacos, los inventores aplicaron criterios de filtrado utilizando las medidas NPV/PPV que permitieron a los inventores reducir el número de sitios potencialmente relevantes de 25,744 a 127 sitios.

Considerando las mejores combinaciones de fármaco-objetivo de acuerdo con los valores de p calculados, los inventores encontraron que el cambio de AA S83L en GyrA es la mutación predominante para 15 fármacos. Dado que solo la ciprofloxacina o la levofloxacina son fármacos aprobados para GyrA, las otras asociaciones a esta proteína podrían ser un efecto secundario de la resistencia a múltiples fármacos. Los inventores demostraron que el empleo de medidas adicionales, como la sensibilidad, el VPP y el VPN, facilita la separación de objetivos farmacológicos causales de otras variantes, como se ejemplifica en este caso.

En lugar de utilizar solo variantes individuales, una combinación de varias posiciones de variante puede mejorar la precisión de la predicción y reducir aún más los resultados falsos positivos que están influenciados por otros factores.

Dado que la duplicación y/o eliminación de genes también podría jugar un papel en los mecanismos de desarrollo de resistencia, los inventores analizaron la cobertura de genes combinada con los datos de resistencia y descubrieron 23 casos de diferencias anormales en la cobertura de genes entre bacterias resistentes y susceptibles. Curiosamente, los inventores no solo encuentran un aumento de material genético en bacterias resistentes, por ejemplo, para los genes *mmuP*, *mmuM*, y *yiel*, pero también una disminución en ciertos genes como *mngB* y *mngR*. Mientras que para las proteínas de membrana o transportadoras, tanto un aumento como una disminución de la dosificación del gen pueden influir en la susceptibilidad del fármaco al no permitir que un fármaco impregne las membranas o transportarlo más eficientemente fuera de la célula, una disminución de la cantidad de enzimas metabólicas o factores de transcripción. no es tan fácil de interpretar en este contexto, y podría estar más o menos directamente relacionado con la aptitud de los aislamientos.

Otra fuente de información que podría mejorar la precisión de nuestro análisis son los plásmidos específicos de la cepa. El mapeo de los datos de secuenciación contra esos plásmidos ampliará nuestro conocimiento sobre mecanismos de resistencia adicionales. En un primer enfoque, los inventores mapearon un subconjunto de datos de secuenciación a aproximadamente 300 plásmidos de *E. coli*. Entre los genes que tienen los sitios variantes más significativos se encuentran, por ejemplo, *repA1*, *trbI*, *psiB*, y *traG* que están directamente involucrados en la replicación, la transferencia de plásmidos y el mantenimiento y pueden desempeñar un papel indirecto en el desarrollo de la resistencia al dar a su huésped la capacidad de facilitar la propagación de genes de resistencia.

En comparación con los enfoques que utilizan MALDI-TOF MS, el enfoque actual tiene la ventaja de que cubre casi todo el genoma y, por lo tanto, nos permite identificar los sitios genómicos potenciales que podrían estar relacionados con la resistencia. Si bien la MALDI-TOF MS también puede usarse para identificar mutaciones puntuales en proteínas bacterianas³³, esta tecnología solo detecta un subconjunto de proteínas y de éstas no todas están igualmente bien cubiertas. Además, la identificación y diferenciación de ciertas cepas relacionadas no siempre es factible.

El presente método permite calcular un mejor punto de interrupción para la separación de aislados en grupos resistentes y susceptibles. Los inventores diseñaron una herramienta de software flexible que permite considerar, además de los

mejores puntos de interrupción, valores definidos por diferentes directrices (por ejemplo, directrices europeas y estadounidenses), que se preparan para una aplicación del GAST en diferentes países.

5 Otro punto crítico de este estudio es que el análisis solo incluyó cepas de bacterias cultivadas. Varios estudios utilizaron muestras de cultivo independientes de orina, muestras fecales o frotis vaginal y aplicaron NGS para identificar o caracterizar los patógenos directamente. El avance de la tecnología NGS, incluido el desarrollo de nuevos secuenciadores de lectura larga como PacBio y Oxford Nanopore, mejorará y acelerará nuestro procedimiento en el futuro para desarrollar una prueba de diagnóstico independiente de la cultura basada en datos de NGS.

10 Los inventores demuestran que el enfoque actual es capaz de identificar mutaciones en genes que ya se conocen como objetivos farmacológicos, así como también detectar posibles nuevos sitios objetivo.

Referencias

15 1. Didelot X, Bowden R, Wilson DJ, Peto TE, Crook DW. Transforming clinical microbiology with bacterial genome sequencing. *Nat Rev Genet* 2012;13:601-12.

20 2. Beerenwinkel N, Schmidt B, Walter H, y otros. Diversity and complexity of HIV-1 drug resistance: a bioinformatics approach to predicting phenotype from genotype. *Proc Natl Acad Sci U S A* 2002;99:8271-6.

3. Reuter S, Ellington MJ, Cartwright EJ, y otros. Rapid bacterial whole-genome sequencing to enhance diagnostic and public health microbiology. *JAMA internal medicine* 2013;173:1397-404.

25 4. Wozniak M, Tiuryn J, Wong L. An approach to identifying drug resistance associated mutations in bacterial strains. *BMC Genomics* 2012;13 Suppl 7:S23.

5. Stoesser N, Batty EM, Eyre DW, y otros. Predicting antimicrobial susceptibilities for *Escherichia coli* and *Klebsiella pneumoniae* isolates using whole genomic sequence data. *The Journal of antimicrobial chemotherapy* 2013;68:2234-44.

30 6. Liu YF, Yan JJ, Lei HY, y otros. Loss of outer membrane protein C in *Escherichia coli* contributes to both antibiotic resistance and escaping antibody-dependent bactericidal activity. *Infection and immunity* 2012;80:1815-22.

35 7. Johnson TJ, Siek KE, Johnson SJ, Nolan LK. DNA sequence and comparative genomics of pAPEC-O2-R, an avian pathogenic *Escherichia coli* transmissible R plasmid. *Antimicrobial agents and chemotherapy* 2005;49:4681-8.

REIVINDICACIONES

1. Un método para determinar un perfil de resistencia a antibióticos para un microorganismo bacteriano que pertenece a la especie *E. coli* que comprende:
 5 determinar la presencia de una mutación en un gen del microorganismo bacteriano contenido o sospechoso de estar contenido en una muestra dada, en donde el gen se selecciona del grupo que consiste en potB, ycgK, ycgB, valS, yjjJ, yigN, yfaL, yjgN, yehl, yjgL, yfdF, yehB, yach, yeil, yncG, ygcQ, rhsD, rem, yehM, pgaA, zraS, stfR, mnmC, hypF, yegE, yphG, fhuA y yeeJ,
 10 en donde la presencia de una mutación en el gen es indicativa de una resistencia a los antibióticos que pertenecen a diferentes grupos de antibióticos.
2. El método de la reivindicación 1, en donde se determina la presencia de una mutación en un gen seleccionado del grupo que consiste en potB, ycgK, ycgB, valS y yjjJ.
- 15 3. El método de la reivindicación 1 o 2, en donde el método comprende determinar la presencia de una mutación en más de un gen seleccionado de dicho grupo.
4. El método de una cualquiera de las reivindicaciones 1 a 3, en donde la mutación en el gen potB es L267V.
- 20 5. El método de la reivindicación 4, en donde la mutación es indicativa de una resistencia a ciprofloxacina (CP), levofloxacina (LVX) y tetraciclina (TE).
6. El método de una cualquiera de las reivindicaciones 1 a 3, en donde la mutación en el gen ycgK es H63Q.
- 25 7. El método de la reivindicación 6, en donde la mutación es indicativa de una resistencia a ciprofloxacina (CP), levofloxacina (LVX), cefazolina (CFZ), tetraciclina (TE) y cefuroxima (CRM).
8. El método de una cualquiera de las reivindicaciones 1 a 3, en donde la mutación en el gen ycgB es I497L.
- 30 9. El método de la reivindicación 8, en donde la mutación es indicativa de una resistencia a la ciprofloxacina (CP), levofloxacina (LVX), cefazolina (CFZ) y tetraciclina (TE).
10. El método de una cualquiera de las reivindicaciones 1 a 3, en donde la mutación en el gen valS se selecciona del grupo que consiste en D469E, N444K, Y439*, Y413*, K277N, D244E, Y242* y D176E, y en donde el asterisco designa una mutación no conservativa que conduce a un nuevo codón de parada.
 35
11. El método de la reivindicación 10, en donde una mutación seleccionada del grupo que consiste en D469E, N444K, Y439*, Y413*, K277N y D176E es indicativa de una resistencia a la ciprofloxacina (CP), levofloxacina (LVX) y cefazolina (CFZ).
 40
12. El método de la reivindicación 10, en donde la mutación D244E es indicativa de una resistencia a ciprofloxacina (CP), levofloxacina (LVX) y gentamicina (GM).
- 45 13. El método de la reivindicación 10, en donde la mutación Y242* es indicativa de una resistencia a ciprofloxacina (CP), levofloxacina (LVX), cefazolina (CFZ) y tetraciclina (TE).
14. El método de una cualquiera de las reivindicaciones 1 a 3, en donde la mutación en el gen yjjJ es H184R.
- 50 15. El método de la reivindicación 14, en donde la mutación es indicativa de una resistencia a la ciprofloxacina (CP), levofloxacina (LVX) y cefazolina (CFZ).
16. El método de una cualquiera de las reivindicaciones 1 a 15, en donde la presencia de una mutación en un gen seleccionado de dicho grupo se prueba en relación con la cepa de referencia *E. coli* K12 subcepa DH10B.

Figura 1

	Aislados Resistentes	Aislados Susceptibles	
Prueba genética predice resistencia	TP ₂₇₂	FP ₅	$PPV = \frac{TP}{TP+FP}$ 98.2%
Prueba genética predice susceptibilidad	FN ₁₈	TN ₈₆₇	$NPV = \frac{TN}{TN+FN}$ 98.0%
	Sensibilidad $\frac{TP}{TP+FN}$ 93.8%	Especificidad $\frac{TN}{TN+FP}$ 99.4%	Exactitud $\frac{(TP+TN)}{(TN+FP+FN+TP)}$ 98.0%

Figura 2

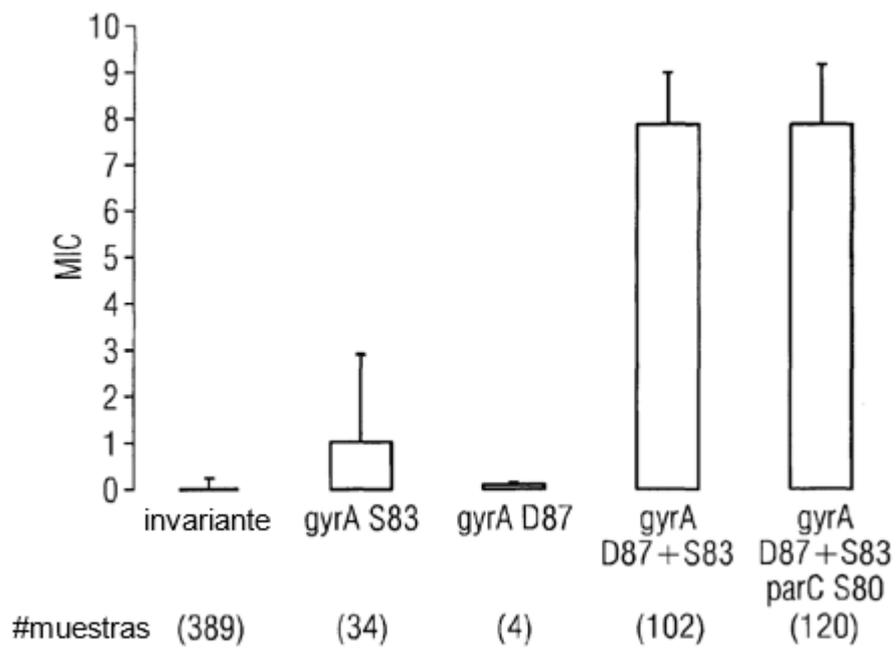


Figura 3

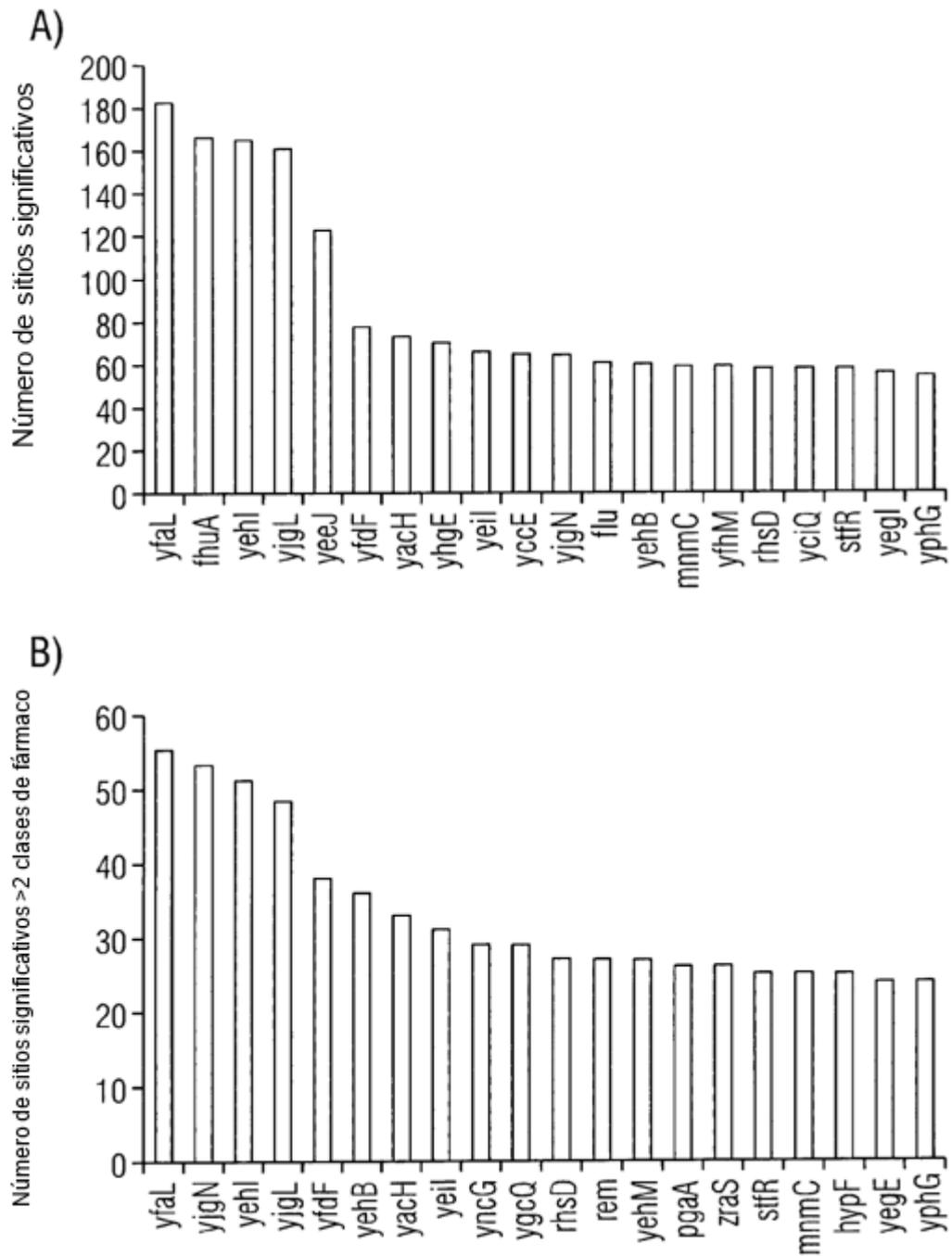


Figura 3

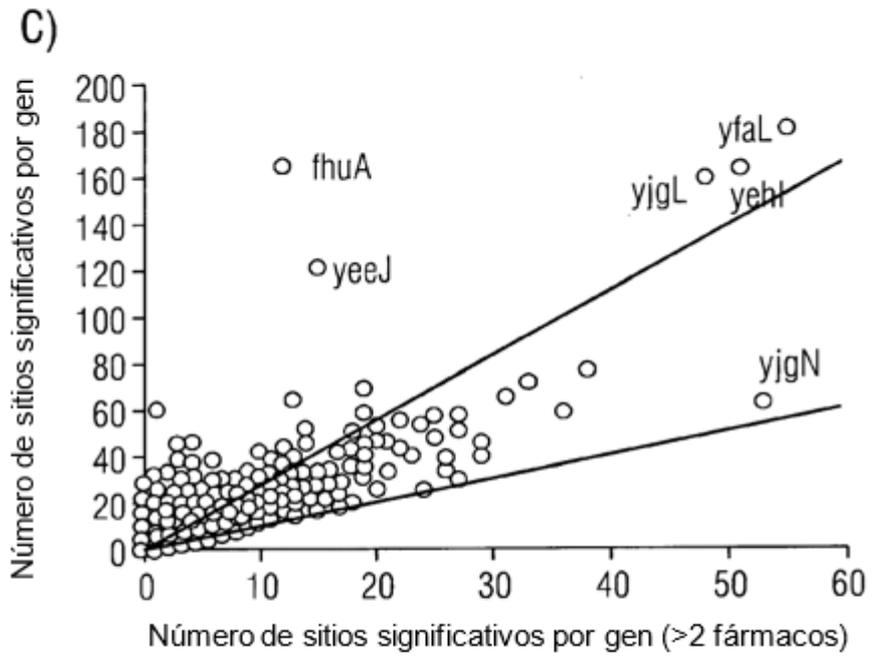


Figura 3

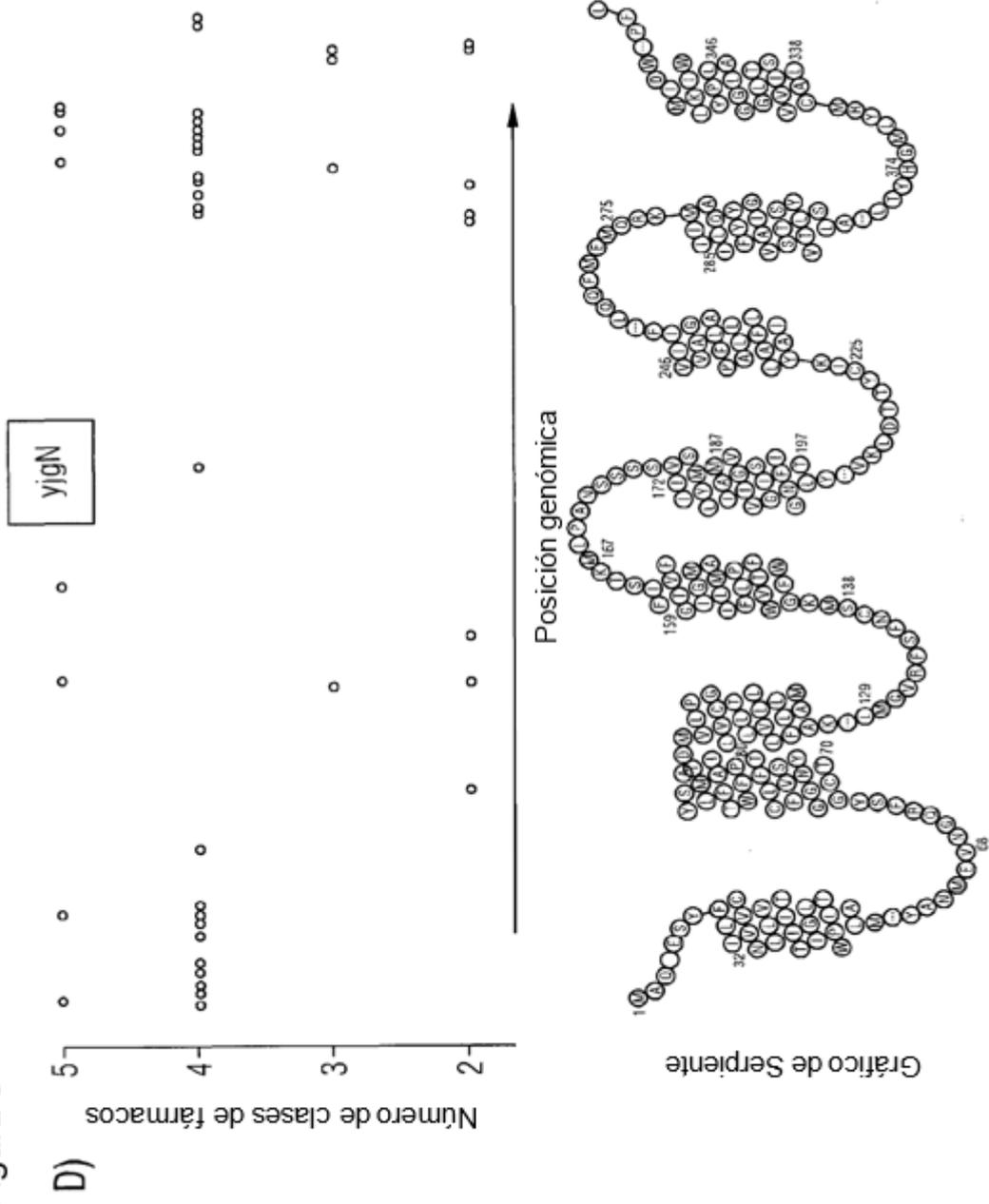


Figura 4

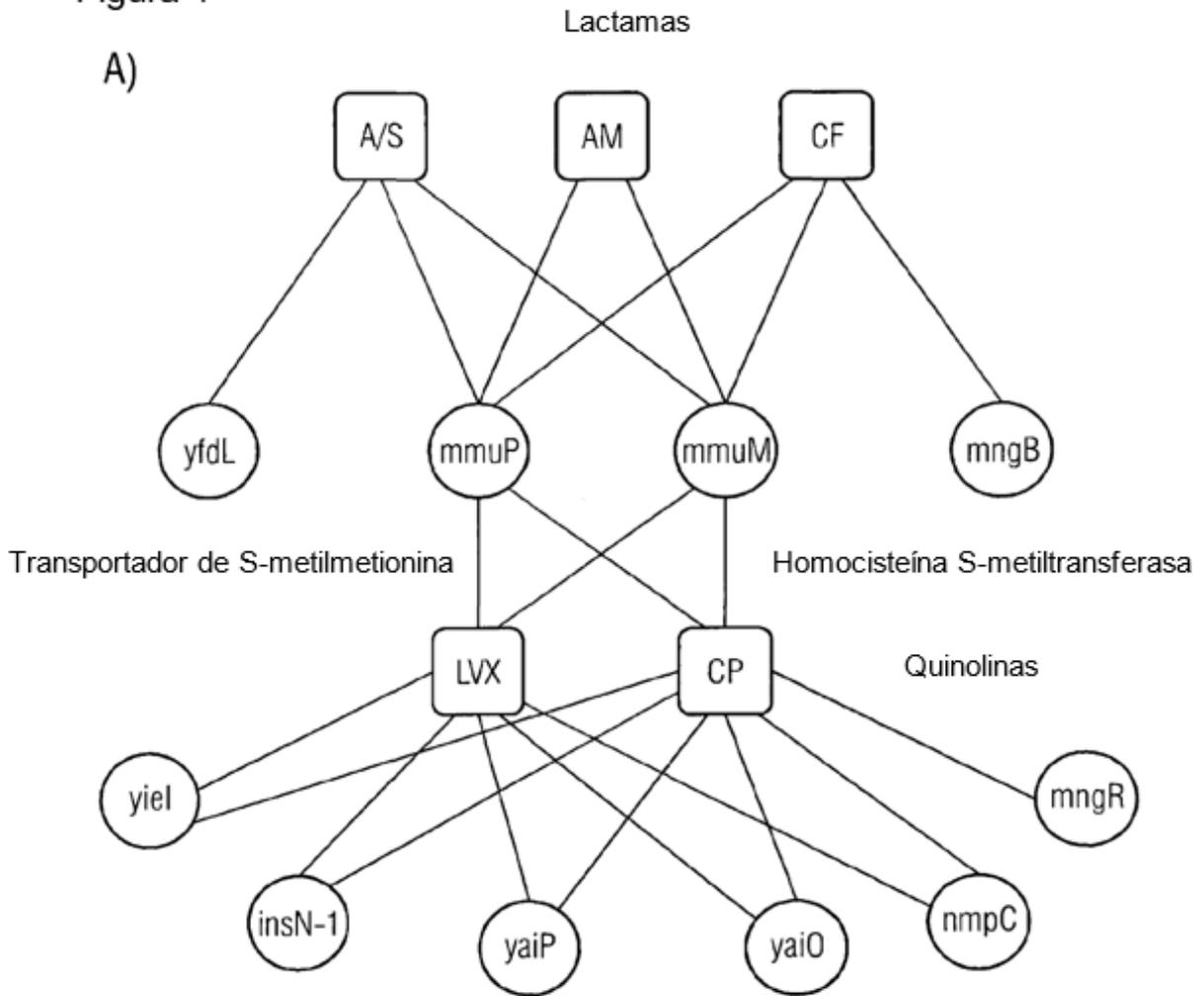


Figura 4

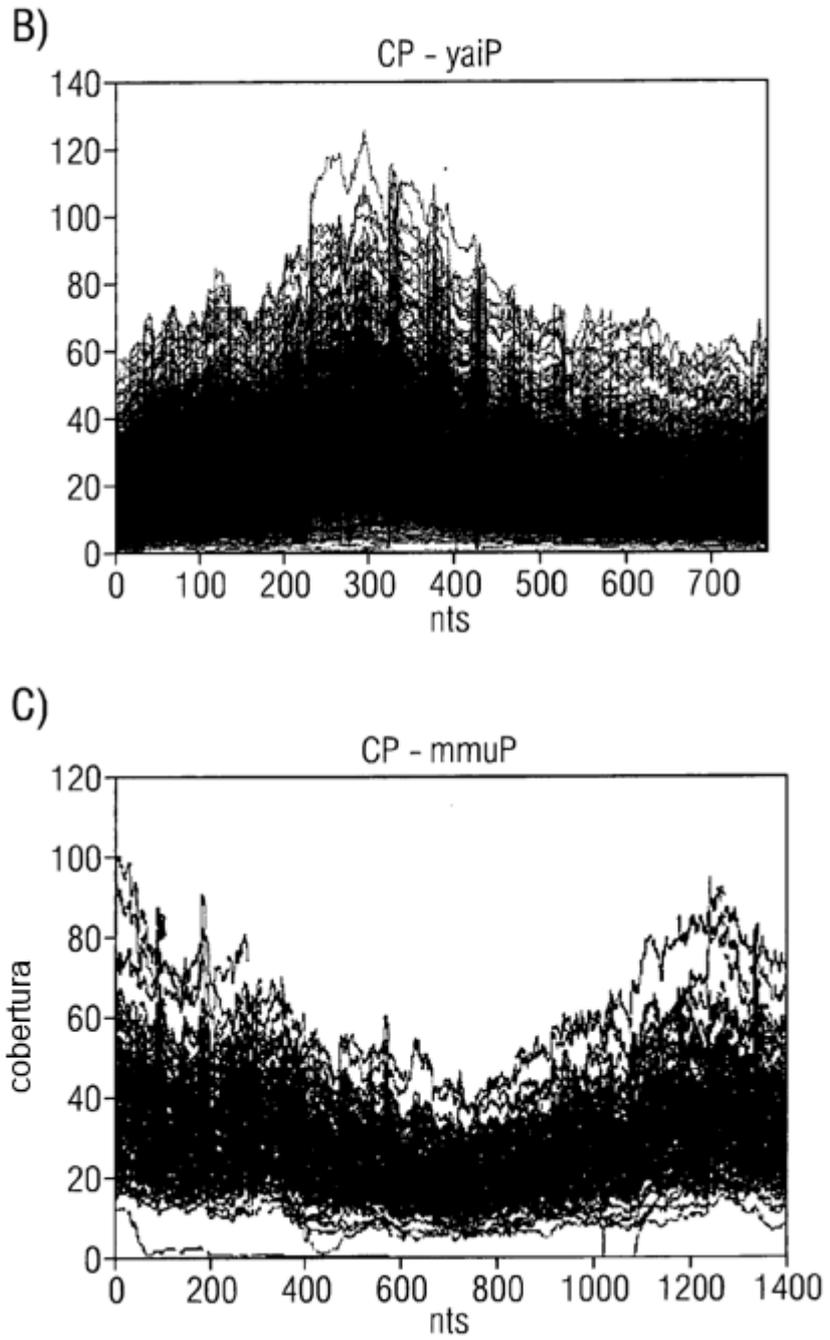


Figura 5

Modelo final para el fármaco AM

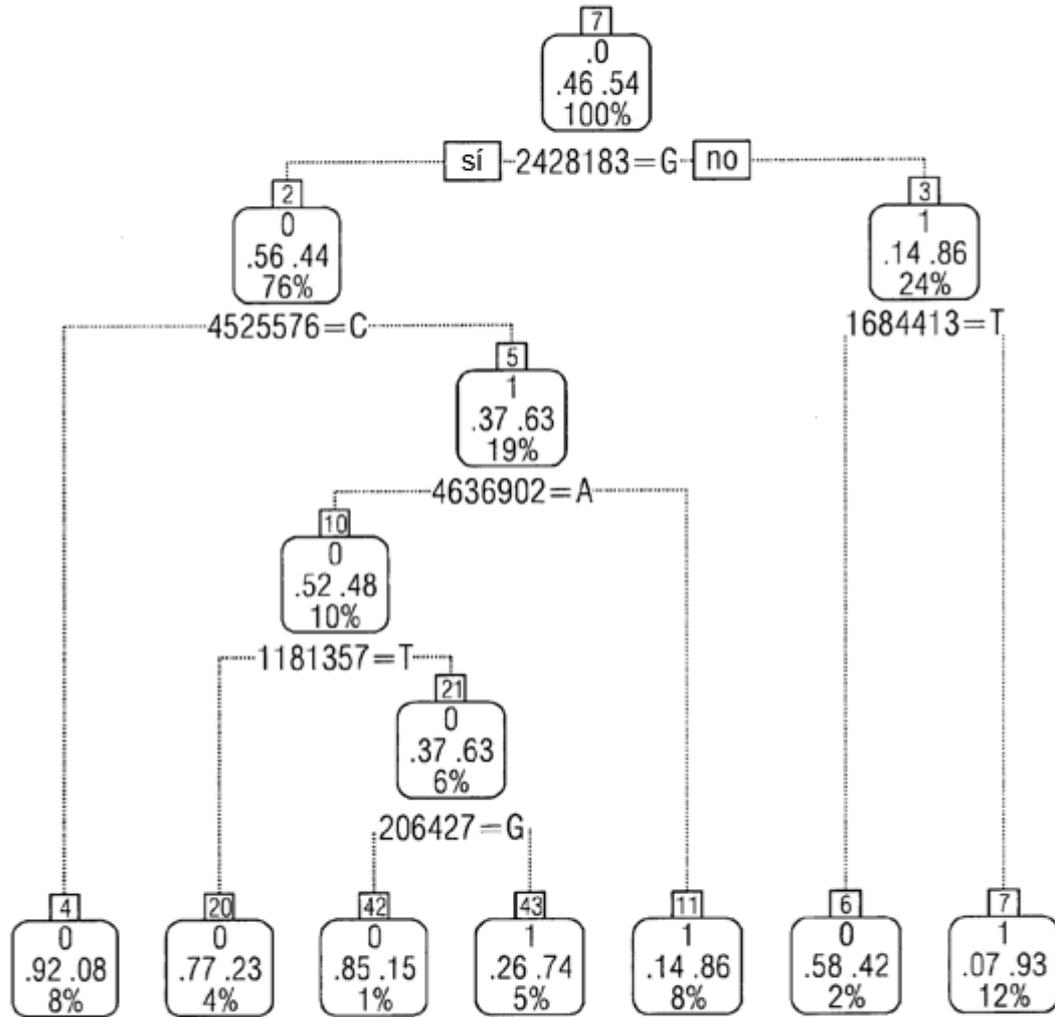


Figura 6

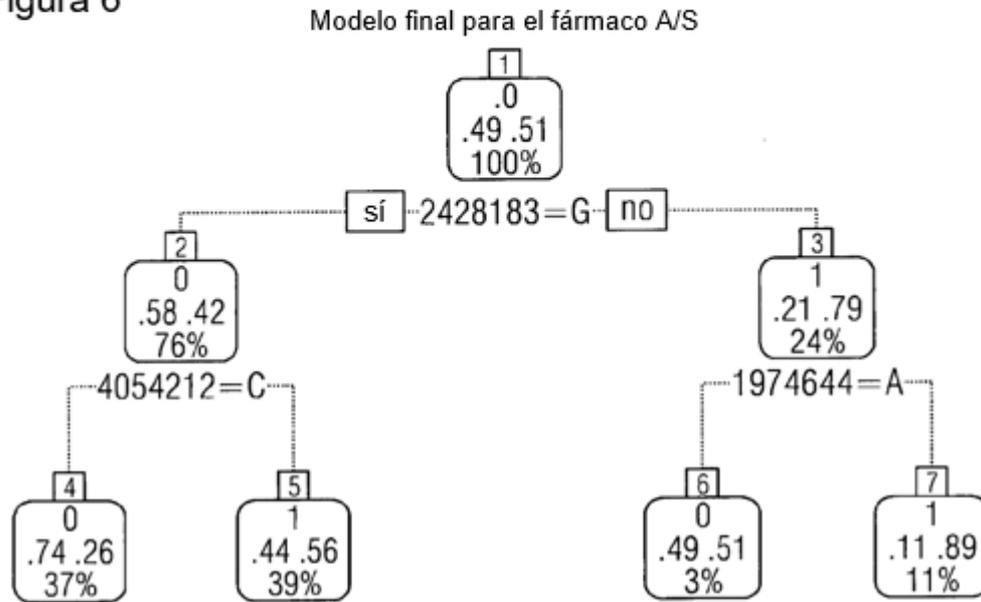


Figura 7

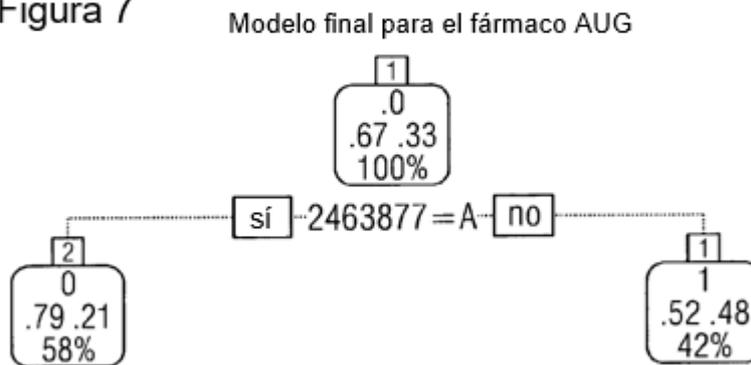


Figura 8

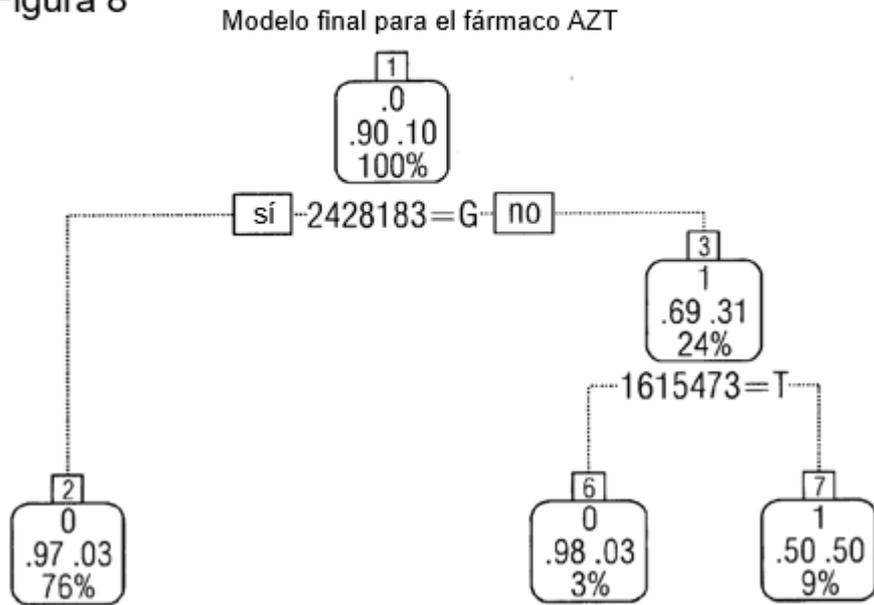


Figura 9

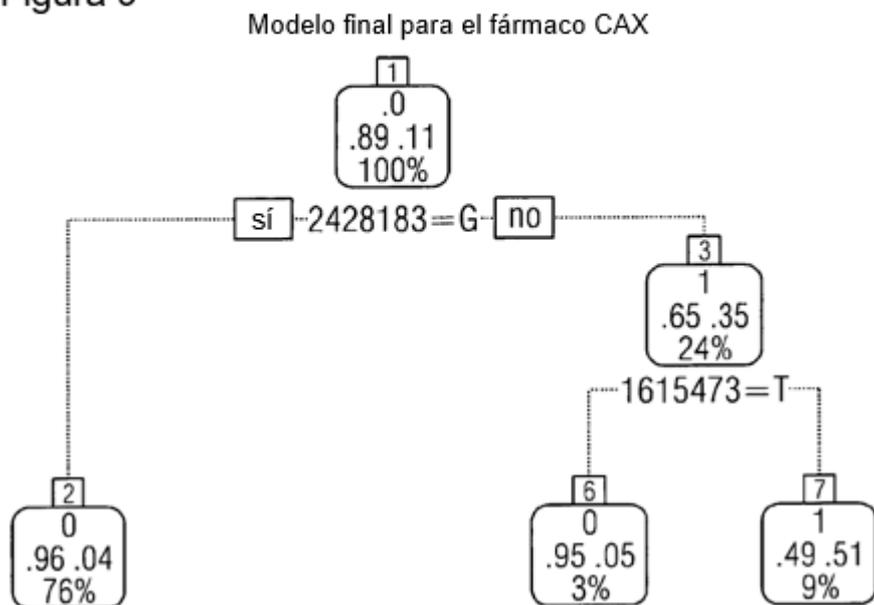


Figura 10

Modelo final para el fármaco CAZ

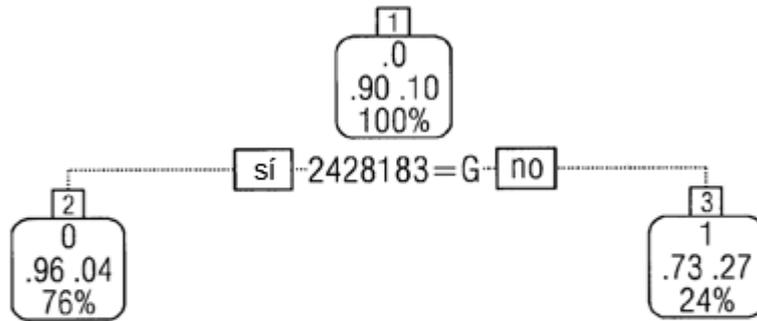


Figura 11

Modelo final para el fármaco CP

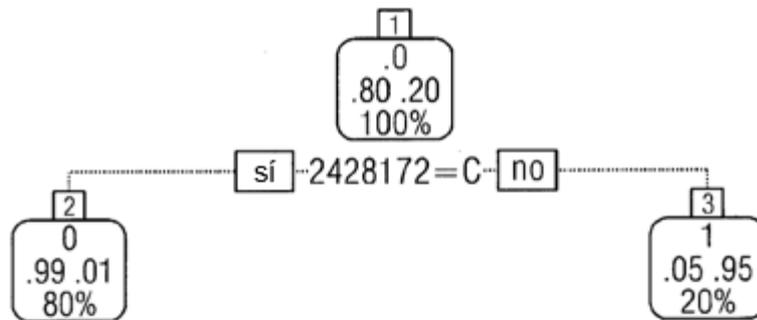


Figura 12

Modelo final para el fármaco CP

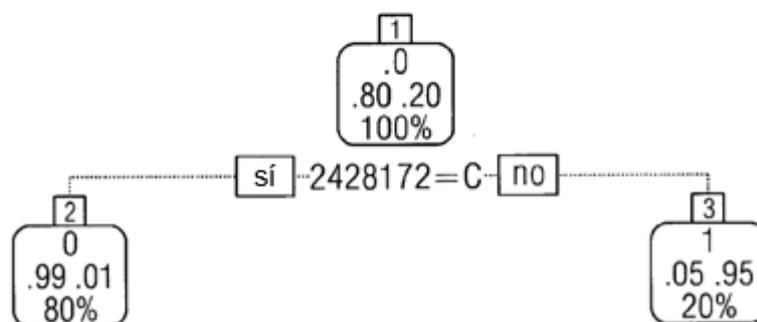


Figura 13

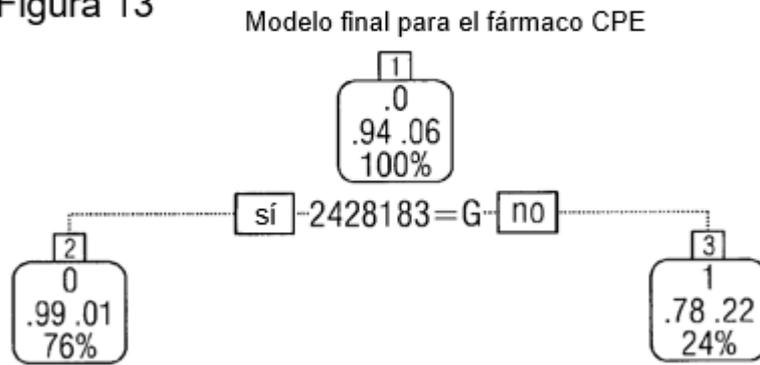


Figura 14

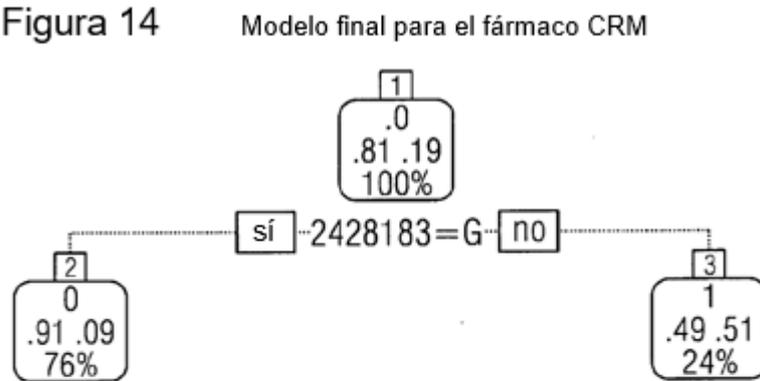


Figura 15

Modelo final para el fármaco ETP

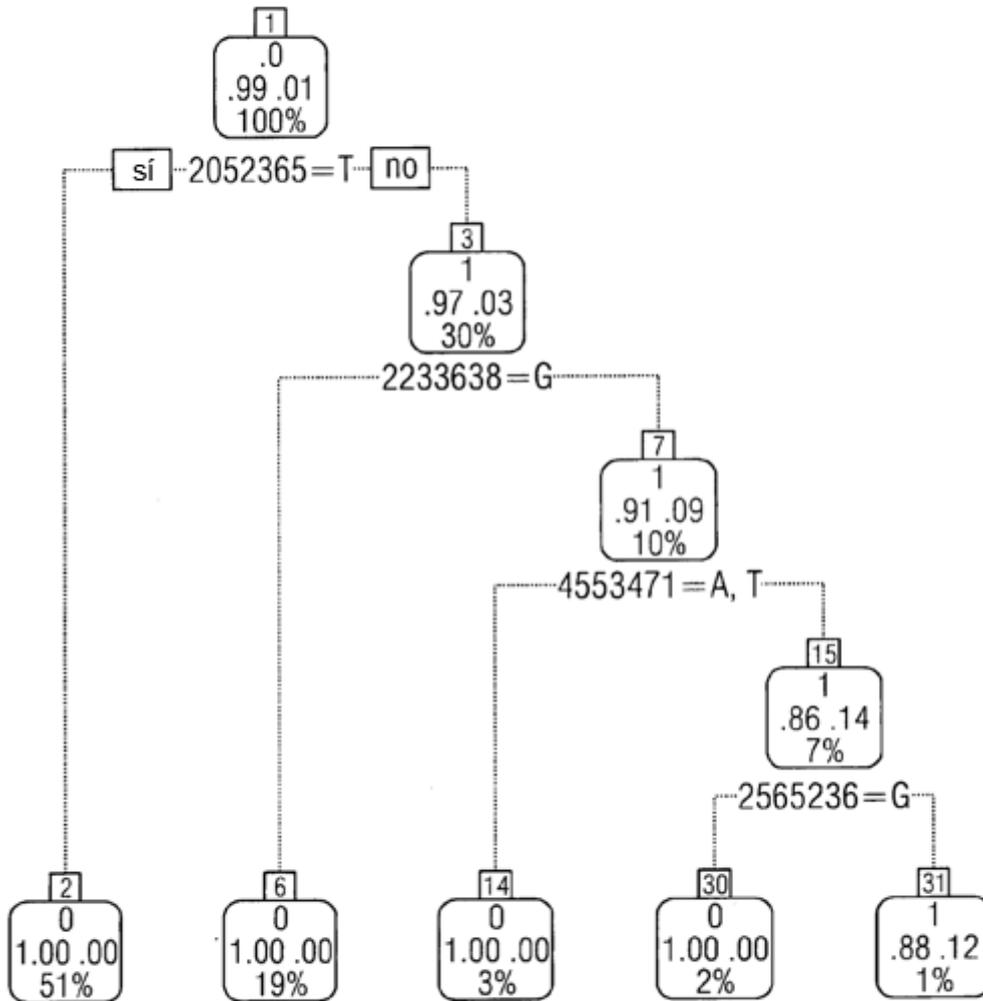


Figura 16

Modelo final para el fármaco GM

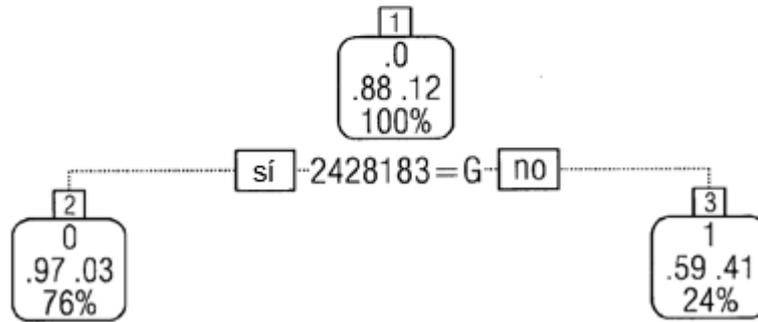


Figura 17

Modelo final para el fármaco LVX

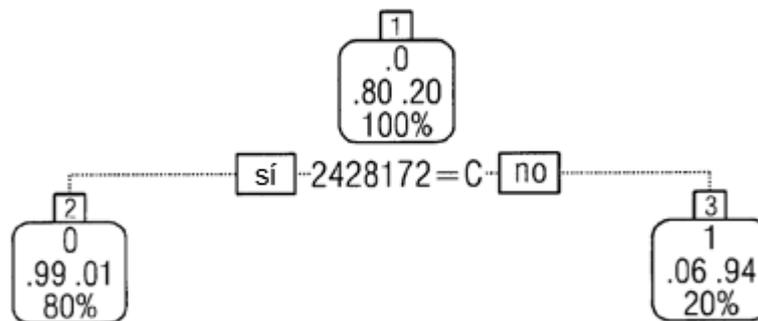


Figura 18 Modelo final para el fármaco P/T

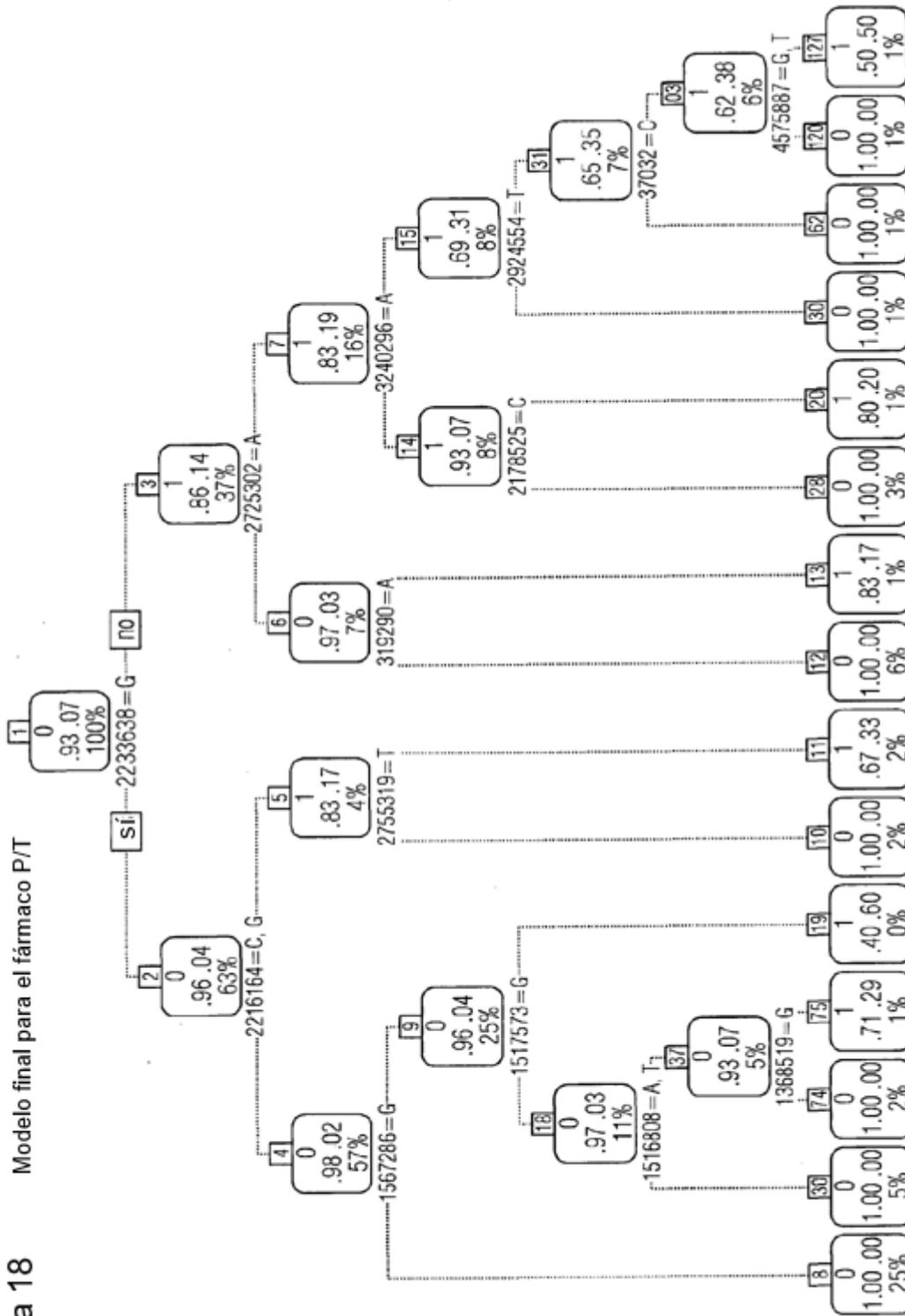


Figura 19 Modelo final para el fármaco TO

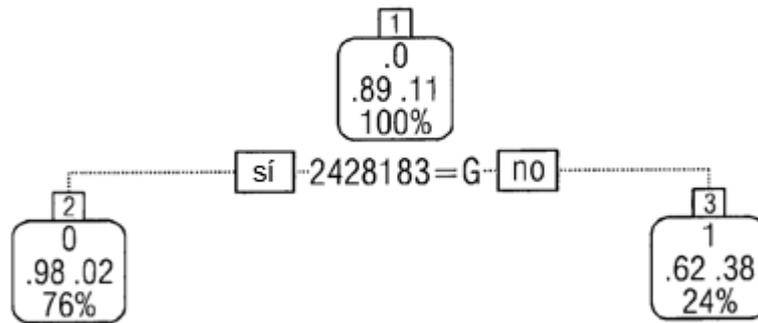


Figura 20 Modelo final para el fármaco T/S

