

5 CONSTRUCTO QUE COMPRENDE UNA SECUENCIA DE ÁCIDO NUCLEICO DE UN GEN FUSR1 QUE CODIFICA UNA PROTEÍNA DE RESISTENCIA CONTRA FUSARIUM OXYSPOURUM RAZA 4 PARA SU EXPRESIÓN EN PLANTAS DE BANANO IDENTIFICACIÓN DE GENES DE RESISTENCIA A PARTIR DE PARIENTES SILVESTRES DEL PLÁTANO Y SUS USOS EN EL CONTROL DEL MAL DE PANAMÁ

REFERENCIA CRUZADA A SOLICITUDES RELACIONADAS

Esta solicitud reivindica el beneficio de la Solicitud de Patente Provisional de los Estados Unidos núm. 62/866,872, presentada el 26 de junio de 2019, y de la Solicitud de Patente Provisional de los Estados Unidos núm. 62/912,010, presentada el 7 de octubre de 2019, cuyos contenidos completos se incorporan en la presente descripción como referencia.

CAMPO DE LA INVENCION

La presente descripción se refiere generalmente al campo de la industria agrícola, especialmente a la producción de cultivos de consumo con resistencia a patógenos. Más particularmente, la presente descripción se refiere a composiciones y métodos para generar plantas que poseen rasgos resistentes a patógenos fúngicos tales como los hongos *Fusarium* procedentes del suelo y/o que muestran resistencia a enfermedades provocadas por dichos patógenos fúngicos.

DESCRIPCIÓN DEL ARCHIVO DE TEXTO ENVIADO ELECTRÓNICAMENTE

El listado de secuencias asociado con esta solicitud se proporciona en formato de texto en lugar de una copia impresa. El contenido del archivo de texto enviado electrónicamente adjunto se incorpora en la presente descripción como referencia en su totalidad: una copia en formato legible por computadora del listado de secuencias (nombre del archivo: EVOL_009_02WO_SeqList_ST25.txt, fecha de registro: 28 de mayo de 2020, tamaño del archivo: ≈ 26,9 kilobytes).

ANTECEDENTES DE LA DESCRIPCIÓN

Los plátanos son uno de los cultivos de frutas más grandes del mundo, con un total de más de 100 millones de toneladas métricas. Los plátanos son la fruta más popular en los países desarrollados y son una importante fuente de alimentos e ingresos para un gran porcentaje del mundo, lo que proporciona seguridad alimentaria en muchas naciones tropicales y subtropicales. De hecho, los plátanos son el cuarto cultivo

alimentario más importante en los países en desarrollo, donde la gran mayoría de los plátanos se producen y consumen localmente. Los principales países productores son India, China, Ecuador, Brasil y algunos países africanos.

Aproximadamente el 15 por ciento de la producción de plátano se comercializa en el mercado mundial, lo que genera aproximadamente \$ 8 mil millones al año. Los principales países exportadores son Ecuador, Filipinas, Costa Rica y Colombia.

Sin embargo, este importante cultivo ahora está severamente amenazado por la marchitez por *Fusarium*, también conocido como mal de Panamá, provocado por el hongo *Fusarium oxysporum* f. sp. *cubense* (Foc).

La mitad del cultivo comercial del plátano en todo el mundo e incluso hasta el 90 % de las exportaciones de plátano en algunos países consiste en un solo grupo de cultivares, los genotipos Cavendish, que se propagan clonalmente. Además, la mayoría de los plátanos comercializados y muchos de los plátanos consumidos localmente se cultivan clonalmente con un solo cultivo en un área determinada, lo que se conoce como "monocultivo". El monocultivo se ha practicado ampliamente por los agricultores para producir en masa cultivos altamente demandados tales como el plátano, que se ve fácilmente afectado por una gama de enfermedades fúngicas, virales, bacterianas y provocadas por nemátodos. Claramente, la expansión actual de la epidemia del mal de Panamá es particularmente destructiva debido al monocultivo masivo de plátanos Cavendish susceptibles.

Los plátanos Cavendish son el fruto de uno de varios cultivares de plátano pertenecientes al subgrupo Cavendish del grupo de cultivares de plátano AAA. El mismo término también se usa para describir las plantas en las que crecen los plátanos. Incluyen cultivares comercialmente importantes como "Cavendish enano" (1888) y "Grand Nain" (el plátano "Chiquita"). "Williams" es un cultivar del tipo "Cavendish gigante" en el subgrupo Cavendish. Es uno de los cultivares más ampliamente cultivados en plantaciones comerciales. "Formosana" es otro nombre para la variante somaclonal "GCTCV-218", que tiene cierta resistencia a la marchitez por *Fusarium* TR4. Otros cultivares comerciales representativos incluyen "Masak Hijau" y "Robusta". Desde la década de 1950, estos cultivares han sido los plátanos más comercializados internacionalmente. Reemplazaron el plátano Gros Michel (comúnmente conocido como plátano de *Kampala* en Kenia y *Bogoya* en Uganda) después de que fuera devastado por el mal de Panamá.

Por tanto, existe una necesidad urgente en la técnica de plátanos que sean resistentes a la marchitez por *Fusarium* o mal de Panamá.

Código de campo cambiado

Código de campo cambiado

Código de campo cambiado

Código de campo cambiado

Código de campo cambiado

Código de campo cambiado

Código de campo cambiado

Código de campo cambiado

Código de campo cambiado

Código de campo cambiado

BREVE DESCRIPCIÓN DE LA DESCRIPCIÓN

- 5 LA presente descripción resuelve el problema del mal de Panamá antes mencionado al identificar la arquitectura genética subyacente que da lugar a la resistencia. Además, la descripción enseña una metodología mediante la cual esta arquitectura genética de resistencia puede importarse en plátanos susceptibles a enfermedades y, por tanto, hacer que estos plátanos sean resistentes a la enfermedad. La importación de esta
- 10 arquitectura genética puede tomar muchas formas, como se detalla en la presente descripción, lo que incluye: fitomejoramiento tradicional, manipulación genética transgénica, fitomejoramiento de próxima generación (CRISPR, edición de bases, MAS, etc.) y otros métodos.
- En algunas modalidades, como se proporciona en la presente descripción, se aíslan
- 15 moléculas de ácido nucleico que comprenden la secuencia de ácido nucleico SEQ ID NO: 14 que codifica la susceptibilidad a *Fusarium oxysporum* raza 4 cuando se expresa en una planta, en donde la SEQ ID NO: 14 se modifica mediante uno, dos, tres o cuatro sustituciones de ácido nucleico de modo que la secuencia de ácido nucleico resultante codifique la resistencia a *Fusarium oxysporum* raza 4 cuando se expresa en una planta.
- 20 En algunas modalidades, la molécula de ácido nucleico aislada incluye sustituciones de ácido nucleico que comprenden reemplazar una T correspondiente a la posición 148 de la SEQ ID NO: 14 por una G (148T>G). En algunas modalidades, la molécula de ácido nucleico aislada incluye sustituciones de ácido nucleico que comprenden reemplazar una T correspondiente a la posición 323 de la SEQ ID NO: 14 por una A (323T>A). En
- 25 algunas modalidades, la molécula de ácido nucleico aislada incluye sustituciones de ácido nucleico que comprenden reemplazar una G correspondiente a la posición 344 de la SEQ ID NO: 14 por una C (344G>C). En algunas modalidades, la molécula de ácido nucleico aislada incluye sustituciones de ácido nucleico que comprenden reemplazar una A correspondiente a la posición 347 de la SEQ ID NO: 14 por una T (347A>T). En
- 30 algunas modalidades, la molécula de ácido nucleico aislada incluye sustituciones de ácido nucleico que comprenden reemplazar una T correspondiente a la posición 323 por una A (323T>A), reemplazar una G correspondiente a la posición 344 por una C (344G>C) y reemplazar una A correspondiente a la posición 347 por una T (347A>T), y en donde todas las posiciones se basan en la SEQ ID NO: 14. En algunas modalidades,
- 35 la molécula de ácido nucleico aislada de SEQ ID NO: 14 codifica una secuencia de

aminoácidos de SEQ ID NO: 15 y en donde las sustituciones de ácido nucleico dan como resultado el reemplazo de una leucina correspondiente a la posición 50 de la SEQ ID NO: 15 por una valina (50L>V). En algunas modalidades, la molécula de ácido nucleico aislada incluye la SEQ ID NO: 14 que codifica una secuencia de aminoácidos de la SEQ ID NO: 15 y en donde las sustituciones de ácido nucleico dan como resultado el reemplazo de una valina correspondiente a la posición 108 de la SEQ ID NO: 15 por un ácido glutámico (108V>E). En algunas modalidades, el ácido nucleico aislado incluye una SEQ ID NO: 14 que codifica una secuencia de aminoácidos de SEQ ID NO: 15 y en donde las sustituciones de ácido nucleico dan como resultado el reemplazo de una arginina correspondiente a la posición 115 de la SEQ ID NO: 15 por una prolina (115R>P). En algunas modalidades, la molécula de ácido nucleico aislada incluye una SEQ ID NO: 14 que codifica una secuencia de aminoácidos de SEQ ID NO: 15 y en donde las sustituciones de ácido nucleico dan como resultado el reemplazo de un ácido aspártico correspondiente a la posición 116 de la SEQ ID NO: 15 por una valina (116D>V). En algunas modalidades, la molécula de ácido nucleico aislada incluye una SEQ ID NO: 14 que codifica una secuencia de aminoácidos de SEQ ID NO: 15 y en donde las sustituciones de ácido nucleico dan como resultado el reemplazo de una valina correspondiente a la posición 108 de la SEQ ID NO: 15 por un ácido glutámico (108V>E), una arginina correspondiente a la posición 115 de la SEQ ID NO: 15 por una prolina (115R>P) y un ácido aspártico correspondiente a la posición 116 de la SEQ ID NO: 15 por una valina (116D>V).

En algunas modalidades, la expresión se produce en una célula vegetal, tejido vegetal, cultivo de células vegetales, cultivo de tejido vegetal o planta completa. En algunas modalidades, la expresión se produce en una célula, tejido, cultivo celular, cultivo de tejido o planta completa de *Musa*. En algunas modalidades, la expresión se produce en una célula, tejido, cultivo celular, cultivo de tejido o planta completa de *Musa acuminata*. En algunas modalidades, un constructo de ácido nucleico comprende las secuencias de ácido nucleico de la presente invención que están unidas operativamente a un promotor capaz de dirigir la expresión de la secuencia de ácido nucleico. En algunas modalidades, el promotor es un promotor vegetal. En algunas modalidades, el promotor es un promotor 35S. En algunas modalidades, el promotor está codificado por la SEQ ID NO: 31.

En algunas modalidades, un vector de transformación comprende los constructos de ácido nucleico de la presente invención.

En algunas modalidades, se proporciona en la presente descripción un método para transformar una célula vegetal que comprende introducir los vectores de transformación de la presente invención en una célula vegetal, por lo cual la célula vegetal transformada expresa la secuencia de ácido nucleico que codifica la resistencia a *Fusarium oxysporum* raza 4. En algunas modalidades, el método usa una célula vegetal que es una célula vegetal de *Musa*. En algunas modalidades, el método usa una célula vegetal que es una célula vegetal de *Musa acuminata*.

En algunas modalidades, el tejido vegetal transformado se produce a partir de la célula vegetal transformada. En algunas modalidades, se produce una plántula transformada a partir del tejido vegetal transformado. En algunas modalidades, se produce un clon a partir de la plántula transformada. En algunas modalidades, el método comprende hacer crecer la plántula transformada o el clon de la plántula transformada en una planta madura transformada. En algunas modalidades, la planta madura transformada es una planta de *Musa* y la planta de *Musa* madura transformada es capaz de producir frutos.

En algunas modalidades, los métodos de la presente invención incluyen producir además clones de la planta de *Musa* madura transformada. En algunas modalidades, la planta de *Musa* madura transformada o el clon de la planta de *Musa* madura transformada se usan en métodos de mejoramiento.

En algunas modalidades, la presente invención proporciona una molécula de aminoácidos aislada que comprende una secuencia de aminoácidos de SEQ ID NO: 15 que codifica una proteína que cuando se produce en una planta da como resultado susceptibilidad a *Fusarium oxysporum* raza 4, en donde la SEQ ID NO: 15 se modifica mediante sustituciones de uno, dos, tres o cuatro aminoácidos de modo que codifique una proteína que cuando se produce en una planta da como resultado resistencia a *Fusarium oxysporum* raza 4. En algunas modalidades, las sustituciones de aminoácidos comprenden reemplazar una leucina correspondiente a la posición 50 de la SEQ ID NO: 15 por una valina (50L>V). En algunas modalidades, las sustituciones de aminoácidos comprenden reemplazar una valina correspondiente a la posición 108 de la SEQ ID NO: 15 por un ácido glutámico (108V>E). En algunas modalidades, las sustituciones de aminoácidos comprenden reemplazar una arginina correspondiente a la posición 115 de la SEQ ID NO: 15 por una prolina (115R>P). En algunas modalidades, las sustituciones de aminoácidos comprenden reemplazar un ácido aspártico correspondiente a la posición 116 de la SEQ ID NO: 15 por una valina (116D>V). En algunas modalidades, las sustituciones de aminoácidos comprenden reemplazar una valina correspondiente a la posición 108 de la SEQ ID NO: 15 por un ácido glutámico (108V>E), una arginina

correspondiente a la posición 115 de la SEQ ID NO: 15 por una prolina (115R>P) y un ácido aspártico correspondiente a la posición 116 de la SEQ ID NO: 15 por una valina (116D>V). En algunas modalidades, la producción de proteínas se produce en una célula vegetal, tejido vegetal, cultivo de células vegetales, cultivo de tejido vegetal o planta completa. En algunas modalidades, la producción de proteínas se produce en una célula, tejido, cultivo celular, cultivo de tejido o planta completa de *Musa*. En algunas modalidades, la producción de proteínas se produce en una célula, tejido, cultivo celular, cultivo de tejido o planta completa de *Musa acuminata*.

En algunas modalidades, los constructos de ácido nucleico de la presente invención comprenden una secuencia de ácido nucleico que codifica la resistencia a *Fusarium oxysporum* raza 4 cuando se expresa en una planta, en donde dicha secuencia de ácido nucleico se selecciona del grupo que consiste en SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 18, SEQ ID NO: 21, SEQ ID NO: 24 y SEQ ID NO: 29, y en donde la secuencia de ácido nucleico está unida operativamente a un promotor capaz de dirigir la expresión de la secuencia de ácido nucleico. En algunas modalidades, el promotor es un promotor vegetal. En algunas modalidades, el promotor es un promotor 35S. En algunas modalidades, el promotor está codificado por la SEQ ID NO: 31. En algunas modalidades, un vector de transformación comprende los constructos de ácido nucleico de la presente invención. En algunas modalidades, la presente invención proporciona métodos para transformar una célula vegetal que comprenden introducir el vector de transformación en una célula vegetal, con lo que la célula vegetal transformada expresa la secuencia de ácido nucleico que codifica la resistencia a *Fusarium oxysporum* raza 4. En algunas modalidades, la célula vegetal es una célula vegetal de *Musa*. En algunas modalidades, la célula vegetal es una célula vegetal de *Musa acuminata*. En algunas modalidades, los métodos comprenden además producir tejido vegetal transformado a partir de la célula vegetal transformada. En algunas modalidades, se produce una plántula transformada a partir del tejido vegetal transformado. En algunas modalidades, los métodos comprenden además producir un clon de la plántula transformada. En algunas modalidades, los métodos comprenden además hacer crecer la plántula transformada o el clon de la plántula transformada en una planta madura transformada. En algunas modalidades, la planta madura transformada es una planta de *Musa* y la planta de *Musa* madura transformada es capaz de producir frutos. En algunas modalidades, los métodos comprenden además producir clones de la planta de *Musa* madura transformada. En algunas modalidades, la planta

de *Musa* madura transformada o el clon de la planta de *Musa* madura transformada se usa en un método de mejoramiento.

En algunas modalidades, la invención proporciona un método de mejoramiento del plátano que comprende cruzar una primera planta de *Musa* que comprende una
5 secuencia de ácido nucleico que codifica la resistencia a *Fusarium oxysporum* raza 4 con una segunda planta de *Musa* que es susceptible a *Fusarium oxysporum* raza 4 y seleccionar la progenie resultante del cruce basado en su resistencia a *Fusarium oxysporum* raza 4, en donde dicha secuencia de ácido nucleico que codifica la
10 resistencia a *Fusarium oxysporum* raza 4 se selecciona del grupo que consiste en SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 18, SEQ ID NO: 21, SEQ ID NO: 24 y SEQ ID NO: 29. En algunas modalidades, los métodos de mejoramiento del plátano comprenden además producir clones de la progenie resultante del cruce en donde los clones se seleccionan basado en su resistencia a *Fusarium oxysporum* raza 4. En algunas modalidades, la primera y segunda plantas de *Musa* son
15 de diferentes especies de *Musa*. En algunas modalidades, la primera y segunda plantas de *Musa* son de la misma especie de *Musa*. En algunas modalidades, la primera y/o segunda planta de *Musa* es una planta de *Musa acuminata*. En algunas modalidades, la progenie del cruce que muestra resistencia a *Fusarium oxysporum* raza 4 se selecciona mediante el uso de marcadores moleculares que se diseñan basado en la
20 secuencia de ácido nucleico que codifica la resistencia a *Fusarium oxysporum* raza 4 que está presente en la primera planta de *Musa* usada en el cruzamiento.

En algunas modalidades, la presente invención proporciona métodos para obtener una célula vegetal de *Musa acuminata* con un gen endógeno silenciado que codifica la
susceptibilidad a *Fusarium oxysporum* raza 4, donde el método comprende la
25 introducción de una ruptura de doble cadena en al menos un sitio en un gen endógeno codificado por la SEQ ID NO: 14 para producir una célula vegetal de *Musa acuminata* con un gen endógeno silenciado que codifica la susceptibilidad a *Fusarium oxysporum* raza 4. En algunas modalidades, los métodos comprenden además generar una planta de *Musa acuminata* a partir de la célula vegetal de *Musa acuminata* con un gen
30 endógeno silenciado que codifica la susceptibilidad a *Fusarium oxysporum* raza 4 para producir una planta de *Musa acuminata* con un gen endógeno silenciado que codifica la susceptibilidad a *Fusarium oxysporum* raza 4. En algunas modalidades, los métodos comprenden además el uso de la planta de *Musa acuminata* con un gen endógeno silenciado que codifica la susceptibilidad a *Fusarium oxysporum* raza 4 en un programa
35 de mejoramiento del plátano. En algunas modalidades, los métodos de la presente

invencción utilizan una célula vegetal que es la célula vegetal de *Musa acuminata* con un gen endógeno silenciado que codifica la susceptibilidad a *Fusarium oxysporum* raza 4. En algunas modalidades, la ruptura de doble cadena es inducida por una nucleasa seleccionada del grupo que consiste en una TALEN, una meganucleasa, una nucleasa con dedos de zinc y una nucleasa asociada a CRISPR. En algunas modalidades, la ruptura de doble cadena es inducida por una nucleasa asociada a CRISPR y donde se proporciona un ARN guía.

En algunas modalidades, la presente invencción proporciona métodos para producir una célula vegetal resistente a *Fusarium oxysporum* raza 4 que comprenden introducir al menos una modificación genética en una o más secuencias de ácido nucleico endógenas que codifican la susceptibilidad a *Fusarium oxysporum* raza 4, en donde la modificación genética confiere resistencia a *Fusarium oxysporum* raza 4 a la célula vegetal. En algunas modalidades, se introduce al menos una modificación genética mediante una TALEN, una meganucleasa, una nucleasa con dedos de zinc o una nucleasa asociada a CRISPR. En algunas modalidades, la al menos una modificación genética se introduce mediante una nucleasa asociada a CRISPR y un ARN guía asociado. En algunas modalidades, la al menos una modificación genética se selecciona de la lista que consiste en reemplazar una T correspondiente a la posición 148 de la SEQ ID NO: 14 por una G (148T>G), reemplazar una T correspondiente a la posición 323 de la SEQ ID NO: 14 por una A (323T>A), reemplazar una G correspondiente a la posición 344 de la SEQ ID NO: 14 por una C (344G>C), y reemplazar una A correspondiente a la posición 347 de la SEQ ID NO: 14 por una T (347A>T). En algunas modalidades, la al menos una modificación genética da como resultado un cambio en un aminoácido seleccionado del grupo que consiste en reemplazar una leucina correspondiente a la posición 50 de la SEQ ID NO: 15 por una valina (50L>V), reemplazar una valina correspondiente a la posición 108 de la SEQ ID NO: 15 por un ácido glutámico (108V>E), reemplazar una arginina correspondiente a la posición 115 de la SEQ ID NO: 15 por una prolina (115R>P), y reemplazar un ácido aspártico correspondiente a la posición 116 de la SEQ ID NO: 15 por una valina (116D>V). En algunas modalidades, la célula vegetal es una célula vegetal de *Musa*. En algunas modalidades, la célula vegetal es una célula vegetal de *Musa acuminata*. En algunas modalidades, los métodos comprenden además producir tejido vegetal transformado a partir de la célula vegetal transformada. En algunas modalidades, los métodos comprenden además producir una plántula transformada a partir del tejido vegetal transformado. En algunas modalidades, los métodos comprenden además producir un

clon de la plántula transformada. En algunas modalidades, los métodos comprenden además hacer crecer la plántula transformada o el clon de la plántula transformada en una planta madura transformada. En algunas modalidades, la planta madura transformada es una planta de *Musa* y la planta de *Musa* madura transformada es capaz de producir frutos. En algunas modalidades, los métodos comprenden además producir clones de la planta de *Musa* madura transformada. En algunas modalidades, los métodos comprenden además el uso de la planta de *Musa* madura transformada o el clon de la planta de *Musa* madura transformada en un método de mejoramiento.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

La **Figura 1** ilustra secuencias codificantes de *FusR1* de plátano alineadas. Los codones de iniciación (inicio) y de terminación (parada) están subrayados.

Las sustituciones de bases de nucleótidos de *FusR1* entre especies de *Musa* están en **negritas**. Las sustituciones que codifican el reemplazo de residuos de aminoácidos (es decir, no son sinónimos) se muestran en **negritas** con un asterisco (*); las sustituciones silenciosas se muestran en **negritas** con un punto (•). Las primeras 96 bases codifican un péptido líder (mostrado en minúsculas) que se escinde de la proteína madura. Se conoce que esto es común para las proteínas Bowman-Birk (Barbosa y otros, 2007). El inventor confirmó la extensión de la secuencia líder mediante el uso de dos herramientas bioinformáticas diferentes, SignalP-5.0 (Armenteros y otros, 2019) y PrediSi (Hiller y otros, 2004), que identificaron ambas el mismo péptido líder. Mediante el uso de la herramienta bioinformática DeepLoc-1.0 (Armenteros y otros, 2017), el inventor estableció después que la proteína FUSR1 madura está localizada en el citoplasma celular (probabilidad de 0,9732).

Las bases se muestran en un código en MAYÚSCULAS para la proteína madura.

Una base faltante, mostrada como un guion (-), en la secuencia de *FusR1* de *M. balbisiana* da como resultado un codón de parada prematuro (mostrado en minúsculas subrayadas, en cursiva), con respecto a las otras secuencias de *FusR1*. Como se describe en el texto, los ARNm de *FusR1* de todas las accesiones de *M. balbisiana* que el inventor examinó tienen un intrón sin empalmar (es decir, expresado); para mayor claridad en la Figura y para centrarse en las similitudes/diferencias de secuencia en las secuencias codificantes de *FusR1* de diferentes especies de plátano, la secuencia del intrón se ha eliminado aquí de *M. balbisiana*, aunque el inventor no ha visto que eso suceda. Por tanto, la SEQ ID NO: 27 es una secuencia codificante "hipotética".

La secuencia de *FusR1* de *M. itinerans* se obtuvo de múltiples accesiones (ITC1526, ITC1571 y PT-BA-00223), todos los cuales son resistentes a FW. La secuencia de *FusR1* de *M. acuminata* etiquetada como "resistente a FW" se obtuvo de múltiples accesiones resistentes a FW, lo que incluye ITC0896 (*M. a.* subespecie *banksii*) y PT-BA-00281 (Pisang Bangkahulu). La secuencia de *M. acuminata* etiquetada como "sensible" proviene de las accesiones sensibles a FW ITC0507, ITC0685, PT-BA-00304, PT-BA-00310 y PT-BA-00315. Estas accesiones incluyen múltiples muestras de cultivares de plátano tales como Pisang Madu, Pisang Pipit y Pisang Rojo Uter, todos los cuales se han caracterizado bien como sensibles a FW (Chen y otros, 2019). La secuencia de *M. balbisiana* incluida aquí se obtuvo de ITC1016. *FusR1* de *M. basjoo* es de accesiones resistentes a FW (ITC0061 y PD # 3064).

El examen de la Figura 1 revela que nuestras secuencias de *FusR1* de plátano están bien conservadas en la región que codifica el péptido líder, como se esperaba. Sin embargo, la secuencia de *FusR1* que codifica la proteína FUSR1 madura muestra un número inusualmente alto de sustituciones no sinónimas. Este es el resultado de una fuerte presión selectiva sobre estas proteínas, que se refleja en las elevadas relaciones Ka/Ks observadas para estos genes. (Ver más abajo) El inventor encontró 2 alelos resistentes a FW para *FusR1* de *M. itinerans*. Éstos difieren muy levemente y por simplicidad, solo se muestra el alelo 1 (SEQ ID NO: 2) de *M. itinerans* en la Figura 1. La secuencia codificante del alelo 2 de *M. itinerans* se incluye en el listado de secuencias como SEQ ID NO: 5. De manera similar, el inventor encontró 2 alelos resistentes a FW para *FusR1* en *M. acuminata*. Estos se diferencian solo por una única sustitución de base silenciosa. Nuevamente, por simplicidad, la Figura 1 muestra solo uno de estos alelos (SEQ ID NO: 9). El segundo alelo, no mostrado en la Figura 1, se registra en el listado de secuencias como SEQ ID NO: 11.

La **Figura 2** ilustra las secuencias de proteínas FUSR1 de plátano alineadas. Los residuos de aminoácidos que difieren entre las secuencias de proteínas FUSR1 de plátano están subrayados. Los primeros 32 residuos constituyen un péptido líder que se escinde de la proteína madura. Los residuos de la secuencia líder se muestran en minúsculas y los residuos de la proteína madura en MAYÚSCULAS.

La proteína FUSR1 de plátano plegada funcional consiste en dos subdominios: el subdominio 1 se indica mediante un sombreado gris claro; el subdominio 2 se indica con un sombreado gris oscuro. Como en otras proteínas Bowman-Birk, la estructura de FUSR1 de plátano se mantiene mediante 14 enlaces disulfuro. Los residuos de cisteína que forman estos enlaces disulfuro se muestran en **negritas**. Cada subdominio contiene

un sitio reactivo, que se muestra en *cursiva*. Los residuos que son específicos para la tripsina (subdominio 1) y la quimotripsina (subdominio 2) se indican mediante un asterisco (*). Para *M. acuminata*, los residuos que difieren entre el alelo FusR1 sensible a Foc4 y los alelos resistentes a Foc-4 se muestran mediante un punto (•), con el residuo de arginina (número 115) que explica la sensibilidad a Foc4 mostrado en **negritas** con un punto (•).

La **Figura 3** proporciona un árbol filogenético para varias especies de plátano, basado en las secuencias de nucleótidos del gen *C2H2*.

La topología en árbol que se muestra aquí se recuperó del análisis de nuestras secuencias de nucleótidos *C2H2* de plátano. Esta topología es idéntica a la recuperada a partir del análisis de la secuencia de la proteína *C2H2*. El mismo árbol se recuperó a partir de nuestras secuencias de nucleótidos y proteínas *TOPO6*. La topología que se muestra aquí también es similar a la de las referencias.

Es importante señalar que, por el contrario, las topologías recuperadas a partir de las secuencias de proteínas de FusR1 y las regiones codificantes de la proteína del gen FusR1 dan una topología diferente, lo que es claramente el resultado de las presiones selectivas impuestas sobre *FusR1* durante la adaptación debido al desafío por *Fusarium*. Las regiones no codificantes de *FusR1* tienen la misma topología que los árboles filogenéticos de *C2H2* y *TOPO6*.

La historia evolutiva se infirió mediante el uso del método de máxima parsimonia. Se muestra el único árbol con más parsimonia. El índice de consistencia es 1,000000, el índice de retención es 1,000000 y el índice compuesto es 1,000000 para todos los sitios. El árbol de MP se obtuvo mediante el algoritmo Subtree-Pruning-Regrafting (SPR) con nivel de búsqueda 0 en el que los árboles iniciales se obtuvieron mediante la adición aleatoria de secuencias (10 réplicas). Este análisis involucró 5 secuencias de nucleótidos. Las posiciones de codón incluidas fueron 1^a+2^a+3^a+No codificante. Se eliminaron todas las posiciones con menos del 95 % de cobertura del sitio, es decir, se permitieron menos del 5 % de huecos de alineamiento, datos faltantes y bases ambiguas en cualquier posición (opción de delección parcial). Existió un total de 218 posiciones en el conjunto de datos final. Los análisis evolutivos se realizaron en MEGA X (Kumar y otros, 2018).

La **Figura 4** proporciona un árbol filogenético para varias especies de plátano, basado en secuencias de proteínas de FUSR1. Tenga en cuenta que este árbol une *Musa acuminata* y *M. basjoo*, por el contrario a su relación filogenética real. *M. acuminata* está más estrechamente relacionado con *M. balbisiana*, con *M. basjoo* como taxón hermano

de estas 2 especies. Sin embargo, debido a los fuertes efectos de la selección positiva, la secuencia de las proteínas de FusR1 de *M. acuminata* y *M. basjoo* se agrupan. (De hecho, estas secuencias de proteínas son idénticas.)

La historia evolutiva se infirió mediante el uso del método de máxima parsimonia. Se muestra el único árbol con más parsimonia con una longitud = 55. El índice de consistencia es 0,963636, el índice de retención es 0,875000 y el índice compuesto es 0,843182 para todos los sitios. El árbol de MP se obtuvo mediante el algoritmo Subtree-Pruning-Regrafting (SPR) con nivel de búsqueda 0 en el que los árboles iniciales se obtuvieron mediante la adición aleatoria de secuencias (10 réplicas). Los análisis evolutivos se realizaron en MEGA X.

La **Figura 5** proporciona el alineamiento de secuencias de ARNm de *FusR1* de accesiones de *Musa balbisiana* sensibles a FW. Las secuencias incluidas aquí se obtuvieron de muchas accesiones de *M. balbisiana*, lo que incluye ITC1016, ITC0545, ITC0080, ITC1527, ITC0565, ITC1781, ITC1580 y varias otras.

Las sustituciones de bases de nucleótidos de *FusR1* entre accesiones de *Musa balbisiana* están en *cursiva*. Los codones de inicio y de terminación (parada) se muestran en minúsculas. Las inserciones, con respecto a otras accesiones de *M. balbisiana* (así como también a las secuencias de *FusR1* de todas las demás plantas analizadas por el inventor), están en **negritas**. Las deleciones de nucleótidos se muestran mediante el símbolo de dos puntos (:). La delección de 85 pares de bases en *FusR1* de las accesiones ITC0545 e ITC1781 es exclusiva de *M. balbisiana*. Como la secuencia de *FusR1* de ITC1781 es idéntica a la de ITC0545, ITC1781 no se presenta en la Figura 5. De manera similar, la delección de un solo par de bases encontrada en estas accesiones de *M. balbisiana* sensibles a FW no se ha encontrado en ninguna otra secuencia de *FusR1*. Sin embargo, existe en todas las accesiones de *M. balbisiana* analizados por el inventor. Esta delección de un solo par de bases da como resultado un codón de terminación prematuro con respecto a las secuencias de *FusR1* de las accesiones de plátano resistentes a FW.

Todas las accesiones de *M. balbisiana* que el inventor examinó tenían uno de los 4 tipos de alelos mostrados aquí. Varias accesiones compartieron alelos de *FusR1* idénticos. Por tanto, por simplicidad, en esta figura solo se muestran 4 accesiones. Estos 4 alelos de *FusR1* son todos muy similares en secuencia de nucleótidos. Existen variantes transcripcionales entre accesiones, pero todas estas variantes tienen el intrón expresado no empalmado. Todas las accesiones también tienen la delección de un par

de bases de un solo par de bases. Tres accesiones también tienen una delección de 85 pares de bases y varias tienen una inserción de 4 pares de bases.

Por tanto, todas estas secuencias de *FusR1* están "rotas" y todas codifican proteínas *FusR1* no funcionales. Significativamente, todas estas accesiones de *M. balbisiana* son sensibles a FW.

DESCRIPCIÓN DETALLADA DE LA DESCRIPCIÓN

LA presente descripción proporciona una solución para enfermedades fúngicas, virales, bacterianas y/o provocadas por nemátodos al inducir una respuesta de defensa a muchos patógenos invasores. La presente descripción proporciona métodos para identificar materiales genéticos que pueden impulsar la resistencia a enfermedades y/o resistencia a hongos en plantas, lo que incluye el plátano, y en plantas y partes de plantas. Además, la presente descripción proporciona métodos para transferir materiales genéticos a cultivares de plátano susceptibles para dar lugar a rasgos de resistencia a enfermedades y/o hongos. Además, la presente descripción enseña componentes genéticos recientemente identificados y métodos para generar plantas, células vegetales, tejidos y semillas modificados genéticamente, que tienen una resistencia a enfermedades modificada.

I. Definiciones

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 entiende comúnmente un experto en la técnica a la que pertenece la descripción. Aunque se cree que los siguientes términos son bien entendidos por un experto en la técnica, las siguientes definiciones se establecen para facilitar la explicación de la materia objeto descrita en la presente. Aunque pueden usarse cualesquiera métodos y materiales similares o equivalentes a los descritos en la presente descripción en la práctica o en pruebas de la presente descripción, se describen métodos y materiales preferidos. Los siguientes términos se definen más abajo. Estas definiciones tienen propósitos ilustrativos y no pretenden limitar el significado común en la técnica de los términos definidos.

El término "un" o "una" se refiere a uno o más de esa entidad, es decir, puede referirse a un referente plural. Como tal, los términos "un" o "una", "uno o más" y "al menos uno" pueden usarse indistintamente en la presente descripción. Además, la referencia a "un elemento" en el artículo indefinido "a" o "una" no excluye la posibilidad de que esté

presente más de uno de los elementos, a menos que el contexto claramente requiera que exista uno y solo uno de los elementos.

Como se usa en esta memoria descriptiva, el término "y/o" se usa en esta descripción para significar ya sea "y" u "o", a menos que se indique de cualquier otra manera.

- 5 A lo largo de esta memoria descriptiva, a menos que el contexto lo requiera de cualquier otra manera, las palabras "comprende", o variaciones tales como "comprenden" o "que comprende", se entenderá que implican la inclusión de un elemento o número entero o grupo de elementos o números enteros establecidos pero no la exclusión de cualquier otro elemento o número entero o grupo de elementos o números enteros.
- 10 Como se usa en esta solicitud, los términos "aproximadamente" y "alrededor de " se usan como equivalentes. Cualquier número usado en esta solicitud con o sin aproximadamente/de manera aproximada está destinado a cubrir cualquier fluctuación normal apreciada por un experto en la técnica relevante. En determinadas modalidades, el término "aproximadamente" o "alrededor de" se refiere a un intervalo de valores que
- 15 se encuentran dentro del 25 %, 20 %, 19 %, 18 %, 17 %, 16 %, 15 %, 14 %, 13 %, 12 %, 11 %, 10 %, 9 %, 8 %, 7 %, 6 %, 5 %, 4 %, 3 %, 2 %, 1 % o menos en cualquier dirección (mayor o menor que) del valor de referencia establecido a menos que se indique de cualquier otra manera o sea evidente por el contexto (excepto donde dicho número excedería el 100 % de un valor posible).
- 20 Como se usa en la presente descripción, el término "al menos una porción" o "fragmento" de un ácido nucleico o polipéptido significa una porción que tiene las características de tamaño mínimo de dichas secuencias, o cualquier fragmento más grande de la molécula de longitud completa, hasta y que incluye la molécula de longitud completa. Un fragmento de un polinucleótido de la descripción puede codificar una
- 25 porción biológicamente activa de un elemento regulador genético. Puede prepararse una porción biológicamente activa de un elemento regulador genético al aislar una porción de uno de los polinucleótidos de la descripción que comprende el elemento regulador genético y evaluar la actividad como se describe en la presente descripción. De manera similar, una porción de un polipéptido puede ser de 4 aminoácidos, 5
- 30 aminoácidos, 6 aminoácidos, 7 aminoácidos, etc., con aumento hasta el polipéptido de longitud completa. La longitud de la porción a usar dependerá de la aplicación particular. Una porción de un ácido nucleico útil como sonda de hibridación puede ser tan corta como 12 nucleótidos; en algunas modalidades, es de 20 nucleótidos. Una porción de un polipéptido útil como epítipo puede ser tan corta como 4 aminoácidos. Una porción de
- 35 un polipéptido que realiza la función del polipéptido de longitud completa generalmente

tendría más de 4 aminoácidos. En algunas modalidades, un fragmento de un polipéptido o polinucleótido comprende al menos 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 95 %, 96 %, 97 %, 98 % o 99 % de la longitud total del polipéptido o polinucleótido de referencia. En algunas modalidades, un polipéptido o fragmento de polinucleótido
5 puede contener 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000 o más nucleótidos o aminoácidos.

Como se usa en la presente descripción, el término "optimización de codones" implica que el uso de codones de un ADN o ARN se adapta al de una célula u organismo de interés para mejorar la tasa de transcripción de dicho ácido nucleico recombinante en la
10 célula u organismo de interés. El experto en la técnica es muy consciente del hecho de que un ácido nucleico objetivo puede modificarse en una posición debido a la degeneración del codón, mientras que esta modificación todavía conducirá a la misma secuencia de aminoácidos en esa posición después de la traducción, lo que se logra mediante la optimización del codón al tener en cuenta el uso de codones específicos de
15 la especie de una célula u organismo objetivo.

Como se usa en la presente descripción, el término "endógeno" o "gen endógeno" se refiere al gen de origen natural, en la localización en la que se encuentra naturalmente dentro del genoma de la célula huésped. "Gen endógeno" es sinónimo de "gen nativo" como se usa en la presente descripción. Un gen endógeno como se describe en la
20 presente descripción puede incluir alelos de genes de origen natural que se han mutado de acuerdo con cualquiera de los métodos de la presente descripción, es decir, un gen endógeno podría haberse modificado en algún momento mediante métodos tradicionales de fitomejoramiento y/o métodos de fitomejoramiento de próxima generación.

Como se usa en la presente descripción, el término "exógeno" se refiere a una sustancia que proviene de alguna fuente distinta a su fuente nativa. Por ejemplo, los términos "proteína exógena" o "gen exógeno" se refieren a una proteína o gen de una fuente no nativa y que se ha suministrado artificialmente a un sistema biológico. Como se usa en la presente descripción, el término "exógeno" se usa indistintamente con el término
30 "heterólogo" y se refiere a una sustancia que proviene de alguna fuente distinta a su fuente nativa.

Los términos "célula huésped manipulada genéticamente", "célula huésped recombinante" y "cepa recombinante" se usan indistintamente en la presente descripción y se refieren a células huésped que se han modificado genéticamente
35 mediante los métodos de la presente descripción. Por tanto, los términos incluyen una

célula huésped (por ejemplo, bacteria, célula de levadura, célula fúngica, CHO, célula humana, célula vegetal, protoplasto derivado de una planta, callo, etc.) que se ha alterado, modificado o manipulado genéticamente, de manera que presenta un genotipo y/o fenotipo alterado, modificado o diferente (por ejemplo, cuando la modificación genética afecta a secuencias de ácido nucleico codificantes), en comparación con la célula huésped de origen natural de la que se deriva. Se entiende que los términos se refieren no solo a la célula huésped recombinante particular en cuestión, sino también a la progenie o la progenie potencial de dicha célula huésped.

Como se usa en la presente descripción, el término "heterólogo" se refiere a una sustancia que proviene de alguna fuente o localización diferente a su fuente o localización nativa. En algunas modalidades, el término "ácido nucleico heterólogo" se refiere a una secuencia de ácido nucleico que no se encuentra naturalmente en el organismo en particular. Por ejemplo, el término "promotor heterólogo" puede referirse a un promotor que se ha tomado de un organismo fuente y se ha utilizado en otro organismo, en el que el promotor no se encuentra naturalmente. Sin embargo, el término "promotor heterólogo" también puede referirse a un promotor que proviene del mismo organismo fuente, pero que simplemente se ha movido a una nueva localización, en la que dicho promotor no se localiza normalmente.

Pueden introducirse secuencias de genes heterólogos en una célula objetivo mediante el uso de un "vector de expresión", que puede ser un vector de expresión eucariota, por ejemplo, un vector de expresión vegetal. Los métodos usados para construir vectores se conocen bien por los expertos en la técnica y se describen en diversas publicaciones. En particular, las técnicas para construir vectores adecuados, lo que incluye una descripción de los componentes funcionales tales como promotores, potenciadores, señales de terminación y poliadenilación, marcadores de selección, orígenes de replicación y señales de corte y empalme, se revisan en la técnica anterior. Los vectores pueden incluir, pero no se limitan a, vectores plasmídicos, fagémidos, cósmidos, cromosomas artificiales/minicromosomas (por ejemplo, ACE) o vectores virales tales como baculovirus, retrovirus, adenovirus, virus adenoasociados, virus del herpes simple, retrovirus, bacteriófagos. Los vectores de expresión eucariotas contendrán típicamente también secuencias procariotas que faciliten la propagación del vector en bacterias tales como genes de origen de replicación y de resistencia a antibióticos para la selección en bacterias. Una variedad de vectores de expresión eucariotas, que contienen un sitio de clonación en el que puede unirse operativamente un polinucleótido, se conocen bien en la técnica y algunos están disponibles comercialmente en compañías tales como

Stratagene, La Jolla, California; Invitrogen, Carlsbad, California; Promega, Madison, Wis. o BD Biosciences Clontech, Palo Alto, California. En una modalidad, el vector de expresión comprende al menos una secuencia de ácido nucleico que es una secuencia reguladora necesaria para la transcripción y traducción de secuencias de nucleótidos que codifican un péptido/polipéptido/proteína de interés.

Como se usa en la presente descripción, el término "de origen natural" aplicado a un ácido nucleico, un polipéptido, una célula o un organismo, se refiere a un ácido nucleico, polipéptido, célula u organismo que se encuentra en la naturaleza. El término "de origen natural" puede referirse a un gen o secuencia derivada de una fuente de origen natural.

Por tanto, para los propósitos de esta descripción, una secuencia "que no es de origen natural" es una secuencia que se ha sintetizado, mutado, manipulado, editado o modificado de cualquier otra manera para tener una secuencia diferente de las secuencias naturales conocidas. En algunas modalidades, la modificación puede ser a nivel de proteína (por ejemplo, sustituciones de aminoácidos). En otras modalidades, la modificación puede ser a nivel de ADN (por ejemplo, sustituciones de nucleótidos).

Como se usa en la presente descripción, el término "cambio de nucleótidos" o "modificación de nucleótidos" se refiere, por ejemplo, a la sustitución, delección y/o inserción de nucleótidos, como se entiende bien en la técnica. Por ejemplo, dichos cambios/modificaciones de nucleótidos incluyen mutaciones que contienen alteraciones que producen sustituciones, adiciones o delecciones silenciosas, pero no alteran las propiedades o actividades de la proteína codificada o cómo se fabrican las proteínas. Como otro ejemplo, dichos cambios/modificaciones de nucleótidos incluyen mutaciones que contienen alteraciones que producen sustituciones, adiciones o delecciones de reemplazo, que alteran las propiedades o actividades de la proteína codificada o cómo se fabrican las proteínas.

Como se usa en la presente descripción, el término "modificación de proteínas" se refiere, por ejemplo, a la sustitución de aminoácidos, modificación, delección y/o inserción de aminoácidos, como se entiende bien en la técnica.

El término "fitomejoramiento de próxima generación" se refiere a una serie de herramientas y metodologías de mejoramiento de plantas que están disponibles para los obtentores de hoy. Una característica distintiva clave del fitomejoramiento de próxima generación es que el obtentor ya no se limita a depender de la variación fenotípica observada para inferir las causas genéticas subyacentes de un rasgo determinado. Más bien, el fitomejoramiento de próxima generación puede incluir la utilización de marcadores moleculares y la selección asistida por marcadores (MAS), de

manera que el obtentor pueda observar directamente el movimiento de alelos y elementos genéticos de interés de una planta en la población en mejoramiento, y no está limitado a la mera observación del fenotipo. Además, los métodos de fitomejoramiento de próxima generación no se limitan a utilizar la variación genética natural que se encuentra dentro de una población de plantas. Más bien, el obtentor que utiliza la metodología de fitomejoramiento de próxima generación puede acceder a una serie de herramientas modernas de manipulación genética que alteran/cambian/editan directamente la arquitectura genética subyacente de la planta de una manera dirigida, para generar un rasgo fenotípico de interés. En algunos aspectos, las plantas obtenidas con una metodología de fitomejoramiento de próxima generación son indistinguibles de una planta que se obtuvo de manera tradicional, ya que la planta de producto final resultante podría teóricamente ser desarrollada mediante cualquiera de los métodos. En aspectos particulares, una metodología de fitomejoramiento de próxima generación puede dar como resultado una planta que comprende: una modificación genética que es una delección o inserción de cualquier tamaño; una modificación genética que es una sustitución de uno o más pares de bases; una modificación genética que es una introducción de secuencias de ácido nucleico desde dentro del acervo genético natural de la planta (por ejemplo, cualquier planta que pueda cruzarse o mejorarse con una planta de interés) o a partir de la edición de secuencias de ácido nucleico en una planta que corresponden a una secuencia que se conoce se produce en el acervo genético natural de la planta; y descendencia de dichas plantas.

Como se usa en la presente descripción, el término "unido operativamente" se refiere a la asociación de secuencias de ácido nucleico en un solo fragmento de ácido nucleico de modo que la función de una está regulada por la otra. Por ejemplo, un promotor está unido operativamente a una secuencia codificante cuando es capaz de regular la expresión de esa secuencia codificante (es decir, que la secuencia codificante está bajo el control transcripcional del promotor). Las secuencias codificantes pueden unirse operativamente a secuencias reguladoras en una orientación sentido o antisentido. En otro ejemplo, las regiones de ARN complementarias de la descripción pueden unirse operativamente, ya sea directa o indirectamente, en dirección 5' con respecto al ARNm objetivo, o en dirección 3' con respecto al ARNm objetivo, o dentro del ARNm objetivo, o una primera región complementaria está en dirección 5' y su complemento está en dirección 3' con respecto al ARNm objetivo.

Los términos "polinucleótido", "ácido nucleico" y "secuencia de nucleótidos", usados indistintamente en la presente descripción, se refieren a una forma polimérica de

nucleótidos de cualquier longitud, ya sean ribonucleótidos o desoxirribonucleótidos, o análogos de los mismos. Este término se refiere a la estructura primaria de la molécula y, por lo tanto, incluye ADN bicatenario y monocatenario, así como también ARN bicatenario y monocatenario. Este término incluye, pero no se limita a, ADN o ARN de

5 cadena simple, doble o múltiple, ADN genómico, ADNc, híbridos de ADN-ARN, o un polímero que comprende bases de purina y pirimidina u otras bases de nucleótidos naturales, modificadas química o bioquímicamente, no naturales o derivatizadas. También incluye ácidos nucleicos modificados tales como ácidos nucleicos metilados y/o con caperuza, ácidos nucleicos que contienen bases modificadas, modificaciones

10 de la cadena principal, y similares. "Oligonucleótido" se refiere generalmente a polinucleótidos de entre aproximadamente 5 y aproximadamente 100 nucleótidos de ADN monocatenario o bicatenario. Sin embargo, para los propósitos de esta descripción, no existe un límite superior para la longitud de un oligonucleótido. Los oligonucleótidos también se conocen como "oligómeros" u "oligos" y pueden aislarse a partir de genes o

15 sintetizarse químicamente mediante métodos conocidos en la técnica. Debe entenderse que los términos "polinucleótido", "ácido nucleico" y "secuencia de nucleótidos" incluyen, según sea aplicable a las modalidades que se describen, polinucleótidos monocatenarios (tales como sentido o antisentido) y bicatenarios.

Los términos "péptido", "polipéptido" y "proteína" se usan indistintamente en la presente

20 descripción y se refieren a una forma polimérica de aminoácidos de cualquier longitud, que puede incluir aminoácidos codificados y no codificados, aminoácidos modificados química o bioquímicamente o derivatizados y polipéptidos que tienen cadenas principales peptídicas modificadas.

Como se usa en la presente descripción, las frases "constructo recombinante",

25 "constructo de expresión", "constructo quimérico", "constructo" y "constructo de ADN recombinante" se usan indistintamente en la presente descripción. Un constructo recombinante comprende una combinación artificial de fragmentos de ácido nucleico, por ejemplo, secuencias reguladoras y codificantes que no se encuentran juntas en la naturaleza. Por ejemplo, un constructo quimérica puede comprender secuencias

30 reguladoras y secuencias codificantes que se derivan de diferentes fuentes, o secuencias reguladoras y secuencias codificantes derivadas de la misma fuente, pero dispuestas de una manera diferente a la que se encuentra en la naturaleza. Tal constructo puede usarse por sí mismo o puede usarse junto con un vector. Si se usa un vector, entonces la elección del vector depende del método que se usará para

35 transformar las células huésped, como se conoce bien por los expertos en la técnica.

Por ejemplo, puede usarse un vector plasmídico. El experto en la técnica conoce bien los elementos genéticos que deben estar presentes en el vector con el fin de transformar, seleccionar y propagar con éxito las células huésped que comprenden cualquiera de los fragmentos de ácido nucleico aislados de la descripción. El experto en la técnica también reconocerá que diferentes eventos de transformación independientes darán como resultado diferentes niveles y patrones de expresión (Jones y otros, (1985) EMBO J. 4:2411-2418; De Almeida y otros, (1989) Mol. Gen. Genetics 218:78-86) y, por tanto, deben cribarse múltiples eventos para obtener líneas que muestren el nivel y patrón de expresión deseados. Dicho cribado puede realizarse mediante análisis Southern de ADN, análisis Northern de expresión de ARNm, análisis de inmunotransferencia de expresión de proteínas o análisis fenotípico, entre otros. Los vectores pueden ser plásmidos, virus, bacteriófagos, pro-virus, fagémidos, transposones, cromosomas artificiales y similares, que se replican de forma autónoma o pueden integrarse en un cromosoma de una célula huésped. Un vector también puede ser un polinucleótido de ARN desnudo, un polinucleótido de ADN desnudo, un polinucleótido compuesto de ADN y ARN dentro de la misma cadena, un ADN o ARN conjugado con polilisina, un ADN o ARN conjugado con péptido, un conjugado con liposoma ADN, o similares, que no se replica de forma autónoma. Como se usa en la presente descripción, el término "expresión" se refiere a la producción de un producto final funcional, por ejemplo, un ARNm o una proteína (precursora o madura).

El término "fitomejoramiento tradicional" se refiere a la utilización de la variación natural que se encuentra dentro de una población de plantas como una fuente de alelos y variantes genéticas que imparten un rasgo de interés a una planta dada. Los métodos de mejoramiento tradicionales hacen uso de procedimientos de cruzamiento que se basan en gran medida en la variación fenotípica observada para inferir la asociación de alelos causales. Es decir, el fitomejoramiento tradicional se basa en observaciones del fenotipo expresado de una planta determinada para inferir la causa genética subyacente. Estas observaciones se utilizan para informar el procedimiento de mejoramiento con el fin de mover la variación alélica al germoplasma de interés.

Además, el fitomejoramiento tradicional también se ha caracterizado por comprender técnicas de mutagénesis aleatoria, que pueden usarse para introducir una variación genética en un germoplasma dado. Estas técnicas de mutagénesis aleatoria pueden incluir procedimientos de mutagénesis basados en radiación y/o químicos. En consecuencia, una característica clave del fitomejoramiento tradicional es que el obtentor no utiliza una herramienta de ingeniería genética que altera/cambia/edita

directamente la arquitectura genética subyacente de la planta de una manera dirigida, con el fin de introducir diversidad genética y producir un rasgo fenotípico de interés.

Un "efector asociado a CRISPR", como se usa en la presente descripción, puede definirse como cualquier nucleasa, nickasa o recombinasa asociada con CRISPR (repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas), que tiene la capacidad de introducir una escisión de una o dos cadenas en un sitio objetivo genómico, o que tiene la capacidad de introducir una modificación dirigida, que incluye una mutación puntual, una inserción o una delección, en un sitio objetivo genómico de interés. Al menos un efector asociado a CRISPR puede actuar por sí solo o en combinación con otras moléculas como parte de un complejo molecular. El efector asociado a CRISPR puede estar presente como molécula de fusión, o como moléculas individuales que se asocian o están asociadas por al menos una de una interacción covalente o no covalente con ARNg y/o sitio objetivo de modo que los componentes del complejo asociado a CRISPR se ponen en estrecha proximidad física.

Un "editor de base", como se usa en la presente descripción, se refiere a una proteína o un fragmento de la misma que tiene la misma actividad catalítica que la proteína de la que se deriva, donde dicha proteína o fragmento de la misma, solo o cuando se proporciona como complejo molecular, denominado complejo de edición de bases en la presente descripción, tiene la capacidad de mediar en una modificación de base dirigida, es decir, la conversión de una base de interés que resulta en una mutación puntual de interés, que a su vez puede resultar en una mutación dirigida, si la conversión de base no causa una mutación silenciosa, pero más bien una conversión de un aminoácido codificado por el codón que comprende la posición a convertir con el editor de base. Al menos un editor de base de acuerdo con la presente descripción enlazado temporal o permanentemente a al menos un efector asociado a CRISPR, u opcionalmente a un componente de al menos un complejo efector asociado a CRISPR.

El término "nucleasa Cas9" y "Cas9" pueden usarse indistintamente en la presente descripción, que se refieren a una enzima endonucleasa de ADN guiada por ARN asociada con CRISPR (Repeticiones Palindrómicas Cortas Agrupadas Regularmente Interespaciadas), que incluye la proteína Cas9 o fragmentos de la misma (como una proteína que comprende un dominio de escisión de ADN activo de Cas9 y/o un dominio de unión de ARNg de Cas9). Cas9 es un componente del sistema de edición del genoma CRISPR/Cas, que se dirige y escinde una secuencia objetivo de ADN para formar rupturas de doble cadena de ADN (DSB) bajo la guía de un ARN guía.

El término "ARN CRISPR" o "ARNcr" se refiere a la cadena de ARN responsable de hibridar con secuencias de ADN objetivo y de reclutar endonucleasas CRISPR y/o efectores asociados a CRISPR. Los ARNcr pueden ser de origen natural o pueden sintetizarse de acuerdo con cualquier método conocido de producción de RNA.

- 5 El término "ARNtracr" se refiere a un pequeño ARN codificado en trans. El ARNtracr es complementario y pares de bases con ARNcr para formar un híbrido de ARNcr/ARNtracr, capaz de reclutar endonucleasas CRISPR y/o efectores asociados a CRISPR para secuencias objetivo.

- El término "ARN guía" o "ARNg", como se usa en la presente descripción, se refiere a
10 una secuencia de ARN o combinación de secuencias capaces de reclutar una endonucleasa CRISPR y/o efectores asociados a CRISPR en una secuencia objetivo. Por lo general, el ARNg se compone de moléculas de ARNcr y ARNtracr que forman complejos a través del complemento parcial, en donde el ARNcr comprende una secuencia que es suficientemente complementaria a una secuencia objetivo para la
15 hibridación y dirige al complejo CRISPR (es decir, híbrido Cas9-ARNcr/ARNtracr) para que se una específicamente a la secuencia objetivo. Además, puede diseñarse un ARN guía único (ARNsg), que comprende las características tanto del ARNcr como del ARNtracr. Por lo tanto, como se usa en la presente descripción, un ARN guía puede ser un ARNcr natural o sintético (por ejemplo, para Cpf1), un híbrido de ARNcr/ARNtracr
20 natural o sintético (por ejemplo, para Cas9) o un ARN guía único (ARNg).

El término "secuencia guía" o "secuencia espaciadora" se refiere a la porción de un ARNcr o ARN guía (ARNg) que es responsable de hibridar con el ADN objetivo.

- El término "protoespaciador" se refiere a la secuencia de ADN dirigida por una secuencia guía de ARNcr o ARNg. En algunas modalidades, la secuencia protoespaciadora se
25 hibrida con la secuencia guía (espaciadora) de ARNcr o ARNg de un complejo CRISPR. El término "sitio de aterrizaje CRISPR" como se usa en la presente descripción, se refiere a una secuencia de ADN capaz de ser dirigida por un complejo CRISPR-Cas. En algunas modalidades, un sitio de aterrizaje de CRISPR comprende una secuencia de combinación de protoespaciador/motivo adyacente de protoespaciador colocada
30 próximamente que es capaz de ser escindida por un complejo CRISPR.

- El término "complejo CRISPR", "complejo de endonucleasa CRISPR", "complejo CRISPR Cas" o "complejo CRISPR-ARNg" se usan indistintamente en la presente descripción. "Complejo CRISPR" se refiere a una nucleasa Cas9 y/o efectores asociados a CRISPR que forman un complejo con un ARN guía (ARNg). Por tanto, el
35 término "complejo CRISPR" se refiere a una combinación de endonucleasa CRISPR y

ARN guía capaz de inducir una ruptura de doble cadena en un sitio de aterrizaje CRISPR. En algunas modalidades, el "complejo CRISPR" de la presente descripción se refiere a una combinación de proteína Cas9 muerta catalíticamente y ARN guía capaz de dirigirse a una secuencia objetivo, pero no capaz de inducir una ruptura de doble

5 cadena en un sitio de aterrizaje CRISPR porque pierde una actividad nucleasa. En otras modalidades, el "complejo CRISPR" de la presente descripción se refiere a una combinación de nickasa Cas9 y ARN guía capaz de introducir rupturas de una cadena dirigidas a ARNg en el ADN en lugar de rupturas de doble cadena creadas por enzimas Cas de tipo silvestre.

10 Como se usa en la presente descripción, el término "unión específica de secuencia de dirección" en el contexto de complejos CRISPR se refiere a la capacidad de un ARN guía para reclutar una endonucleasa CRISPR y/o efectores asociados a CRISPR en un sitio de aterrizaje CRISPR.

Como se usa en la presente descripción, el término "desaminasa" se refiere a una

15 enzima que cataliza la reacción de desaminación. En algunas modalidades de la presente descripción, la desaminasa se refiere a una citidina desaminasa, que cataliza la desaminación de una citidina o una desoxicitidina a un uracilo o una desoxiuridina, respectivamente. En otras modalidades de la presente descripción, la desaminasa se refiere a una adenosina desaminasa, que cataliza la desaminación de una adenina para

20 formar hipoxantina (en forma de su nucleósido inosina), que la ADN polimerasa lee como guanina.

Como se usa en la presente descripción, el término "glicosilasa" se refiere a una familia de enzimas involucradas en la reparación por escisión de bases, clasificadas bajo el número EC: EC 3.2.2. La reparación por escisión de bases es el mecanismo por el cual

25 las bases dañadas en el ADN se eliminan y reemplazan. Las ADN glicosilasas catalizan la primera etapa de este proceso. Eliminan la base nitrogenada dañada mientras dejan intacta la estructura de azúcar-fosfato, lo que crea un sitio apurínico/apirimidínico, comúnmente conocido como sitio AP. Esto se logra al voltear la base dañada fuera de la doble hélice seguido de la escisión del enlace N-glicosídico. En algunas modalidades

30 de la presente descripción, con la expectativa de proporcionar una tendencia de introducción de mutación diferente a la de la desaminasa y similares, una reacción de escisión de base por hidrólisis del enlace N-glicosídico del ADN, y después se usa la inducción de la introducción de mutación en un proceso de reparación de células. En aspectos, se usa una enzima que tiene actividad citosina-ADN glicosilasa (CDG) o

35 timina-ADN glicosilasa (TDG). En aspectos, se usa un mutante de uracilo-ADN

glicosilasa mitocondrial de levadura (UNG 1) como enzima que realiza dicha reacción de escisión de bases. El documento US 2017/0321210 A1 de Nishida y otros, publicado el 9 de noviembre de 2017, se incorpora en la presente descripción como referencia.

Como se usa en la presente descripción, el término "dirigido" se refiere a la expectativa de que un elemento o molécula interactuará con otro elemento o molécula con un grado de especificidad, para excluir elementos o moléculas no dirigidos. Por ejemplo, un primer polinucleótido que se dirige a un segundo polinucleótido, de acuerdo con la presente descripción, se ha diseñado para hibridar con el segundo polinucleótido de una manera específica de secuencia (por ejemplo, mediante emparejamiento de bases Watson-Crick). En algunas modalidades, la región de hibridación seleccionada se diseña para hacer que la hibridación sea única para una o más regiones objetivo. Un segundo polinucleótido puede dejar de ser un objetivo de un primer polinucleótido de dirección, si su secuencia de dirección (región de hibridación) está mutada, o se elimina/separa de otro modo del segundo polinucleótido. Además, "dirigido" puede usarse indistintamente con "sitio específico" o "dirigido al sitio", que se refiere a una acción de biología molecular que usa información sobre la secuencia de una región genómica de interés que va a modificarse, y que además se basa en información del mecanismo de acción de las herramientas moleculares, por ejemplo, nucleasas, que incluye las nucleasas CRISPR y variantes de las mismas, TALEN, ZFN, meganucleasas o recombinasas, enzimas modificadoras del ADN, que incluye las enzimas modificadoras de bases como las enzimas citidina desaminasas, enzimas modificadoras de histonas y similares, ADN-proteínas de unión, ARNcr/ARNtracr, ARN guía y similares.

El término "región semilla" se refiere a la porción crítica de una secuencia guía de ARNcr o ARN guía que es más susceptible a error de apareamiento con sus objetivos. En algunas modalidades, un único desajuste en la región semilla de un ARNcr/ARNg puede convertir un complejo CRISPR en inactivo en ese sitio de unión. En algunas modalidades, las regiones semilla para las endonucleasas Cas9 están ubicadas a lo largo de los últimos ~ 12 nt de la porción 3' de la secuencia guía, que corresponden (hibridan) a la porción de la secuencia objetivo protoespaciadora que es adyacente a la PAM. En algunas modalidades, las regiones semilla para las endonucleasas Cpf1 están ubicadas a lo largo de los primeros ~ 5 nt de la porción 5' de la secuencia guía, que corresponden (hibridan) a la porción de la secuencia objetivo protoespaciadora adyacente a la PAM.

El término "identidad de secuencia" se refiere al porcentaje de bases o aminoácidos entre dos secuencias de polinucleótidos o polipéptidos que son iguales y en la misma

posición relativa. Como tal, una secuencia de polinucleótidos o polipéptidos tiene un cierto porcentaje de identidad de secuencia en comparación con otra secuencia de polinucleótidos o polipéptidos. Para la comparación de secuencias, normalmente una secuencia actúa como secuencia de referencia, con la que se comparan las secuencias de prueba. El término "secuencia de referencia" se refiere a una molécula con la que se compara una secuencia de prueba. Cuando se usa el porcentaje de identidad de secuencia en referencia a proteínas, se reconoce que las posiciones de los residuos que no son idénticos a menudo difieren por sustituciones conservadoras de aminoácidos, donde los residuos de aminoácidos se sustituyen por otros residuos de aminoácidos con propiedades químicas similares (por ejemplo, carga o hidrofobicidad) y, por lo tanto, no modifican las propiedades funcionales de la molécula. Cuando las secuencias difieren en sustituciones conservadoras, el porcentaje de identidad de secuencia puede ajustarse hacia arriba para corregir la naturaleza conservadora de la sustitución. Se dice que las secuencias que se diferencian por dichas sustituciones conservadoras tienen "similitud de secuencia" o "similitud". Los expertos en la técnica conocen bien los medios para realizar este ajuste. Normalmente, esto implica puntuar una sustitución conservadora como una falta de coincidencia parcial en lugar de completa, lo que aumenta de esta manera el porcentaje de identidad de secuencia. Por tanto, por ejemplo, cuando un aminoácido idéntico recibe una puntuación de 1 y una sustitución no conservadora recibe una puntuación de cero, una sustitución conservadora recibe una puntuación entre cero y 1. La puntuación de las sustituciones conservadoras se calcula, por ejemplo, de acuerdo con el algoritmo de Meyers y Miller, *Computer Applic. Biol. Sci.* 4:11-17 (1988).

"Complementario" se refiere a la capacidad de emparejar, a través del apilamiento de bases y el enlace de hidrógeno específico, entre dos secuencias que comprenden bases naturales o no naturales o análogas de las mismas. Por ejemplo, si una base en una posición de un ácido nucleico es capaz de formar puentes de hidrógeno con una base en la posición correspondiente de un objetivo, entonces se considera que las bases son complementarias entre sí en esa posición. Los ácidos nucleicos pueden comprender bases universales o espaciadores abásicos inertes que no proporcionan ninguna contribución positiva o negativa al enlace de hidrógeno. Los emparejamientos de bases pueden incluir emparejamientos de bases de Watson-Crick canónicos y emparejamientos de bases que no son de Watson-Crick (por ejemplo, emparejamientos de bases de Wobble y emparejamientos de bases de Hoogsteen). Se entiende que para los pares de bases complementarias, las bases de tipo adenosina (A) son

complementarias a las bases de tipo timidina (T) o las bases de tipo uracilo (U), que las bases de tipo citosina (C) son complementarias a las bases de tipo guanosina (G), y que bases universales tales como 3-nitropirrol o 5-nitroindol pueden hibridar y se consideran complementarias a cualquier A, C, U o T. Nichols y otros, *Nature*, 1994; 369:492-493 y
5 Loakes y otros, *Nucleic Acids Res.*, 1994; 22:4039-4043. La inosina (I) también se ha considerado en la técnica como una base universal y se considera complementaria a cualquier A, C, U o T. Ver Watkins y Santa Lucia, *Nucl. Acids Research*, 2005; 33(19):6258-6267.

Como se menciona en la presente descripción, una "secuencia de ácido nucleico
10 complementaria" es una secuencia de ácido nucleico que comprende una secuencia de nucleótidos que le permite unirse de forma no covalente a otro ácido nucleico de una manera antiparalela específica de secuencia (es decir, un ácido nucleico específicamente se une a un ácido nucleico complementario) en las condiciones apropiadas *in vitro* y/o *in vivo* de temperatura y fuerza iónica de la solución.

15 Los métodos de alineamiento de secuencias para la comparación y determinación del porcentaje de identidad de secuencia y el porcentaje de complementariedad se conocen bien en la técnica. El alineamiento óptimo de secuencias para comparación puede realizarse, por ejemplo, mediante el algoritmo de alineamiento por homología de Needleman y Wunsch, (1970) *J. Mol. Biol.* 48:443, mediante el método de búsqueda de similitud de Pearson y Lipman, (1988) *Proc. Nat'l. Acad. Sci. USA* 85:2444, mediante
20 implementaciones computarizadas de estos algoritmos (GAP, BESTFIT, FASTA y TFASTA en Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), mediante alineamiento manual e inspección visual (ver, por ejemplo, Brent y otros, (2003) *Current Protocols in Molecular Biology*), mediante el uso
25 de algoritmos conocidos en la técnica que incluyen los algoritmos BLAST y BLAST 2.0, que se describen en Altschul y otros, (1977) *Nuc. Acids Res.* 25:3389-3402; y Altschul y otros, (1990) *J. Mol. Biol.* 215:403-410, respectivamente. El software para realizar análisis BLAST está disponible públicamente a través del National Center for Biotechnology Information. Algunos programas de alineamiento son MacVector (Oxford
30 Molecular Ltd, Oxford, Reino Unido), ALIGN Plus (Software científico y educativo, Pensilvania) y AlignX (Vector NTI, Invitrogen, Carlsbad, CA). Otro programa de alineamiento es Sequencher (Gene Codes, Ann Arbor, Michigan), mediante el uso de parámetros predeterminados, y MUSCLE (Comparación de Secuencias Múltiples mediante Excepción de Registro; un software informático con licencia de dominio
35 público).

En la presente descripción, el término "hibridar" se refiere al emparejamiento entre bases de nucleótidos complementarias (por ejemplo, la adenina (A) forma un par de bases con timina (T) en una molécula de ADN y con uracilo (U) en una molécula de ARN, y guanina (G) forma un par de bases con citosina (C) en moléculas de ADN y ARN) para formar una molécula de ácido nucleico bicatenario. (Ver, por ejemplo, Wahl y Berger (1987) *Methods Enzymol.* 152:399; Kimmel, (1987) *Methods Enzymol.* 152:507). Además, en la técnica también se conoce que para la hibridación entre dos moléculas de ARN (por ejemplo, ARNs), la guanina (G) forma un par de bases con uracilo (U). Por ejemplo, el emparejamiento de bases G/U es parcialmente responsable de la degeneración (es decir, la redundancia) del código genético en el contexto del emparejamiento de bases anti-codón de ARNt con codones en ARNm. En el contexto de esta descripción, una guanina (G) de un segmento de unión a proteínas (ARNds dúplex) de una molécula de ARN guía se considera complementaria a un uracilo (U) y viceversa. Como tal, cuando puede formarse un par de bases G/U en una posición de nucleótido dada, un segmento de unión a proteínas (ARNds dúplex) de una molécula de ARN guía, la posición no se considera no complementaria, sino que se considera que es complementaria. Se entiende en la técnica que la secuencia de polinucleótido no necesita ser 100 % complementaria a la de su ácido nucleico objetivo para ser específicamente hibridable. Además, un polinucleótido puede hibridar sobre uno o más segmentos de modo que los segmentos intermedios o adyacentes no estén implicados en el evento de hibridación (por ejemplo, una estructura de bucle o estructura de horquilla). Un polinucleótido puede comprender al menos 70 %, al menos 80 %, al menos 90 %, al menos 95 %, al menos 99 % o 100 % de complementariedad de secuencia con una región objetivo dentro de la secuencia de ácido nucleico objetivo a la que se dirigen.

El término "modificado" se refiere a una sustancia o compuesto (por ejemplo, una célula, una secuencia polinucleotídica y/o una secuencia polipeptídica) que se ha alterado o cambiado en comparación con la sustancia o compuesto no modificado correspondiente.

"Aislado" se refiere a un material que está libre en diversos grados de los componentes que normalmente lo acompañan como se encuentra en su estado nativo.

El término "planta, parte o célula con genes editados", como se usa en la presente descripción, se refiere a una planta, parte o célula que comprende uno o más genes endógenos que se editan mediante un sistema de edición de genes. El sistema de edición de genes de la presente descripción comprende un elemento de dirección y/o

un elemento de edición. El elemento de dirección es capaz de reconocer una secuencia genómica objetivo. El elemento de edición es capaz de modificar la secuencia genómica objetivo, por ejemplo, mediante sustitución o inserción de uno o más nucleótidos en la secuencia genómica, delección de uno o más nucleótidos en la secuencia genómica, alteración de secuencias genómicas para incluir secuencias reguladoras, inserción de transgenes en un sitio genómico de puerto seguro u otra ubicación específica en el genoma, o cualquier combinación de los mismos. El elemento de dirección y el elemento de edición pueden estar en la misma molécula de ácido nucleico o en diferentes moléculas de ácido nucleico. En algunas modalidades, el elemento de edición es capaz de editar el genoma con precisión mediante la sustitución de un solo nucleótido mediante el uso de un editor de base, como el editor de base de citosina (CBE) y/o el editor de base de adenina (ABE), que se fusiona directa o indirectamente a una proteína efectora asociada a CRISPR.

El término "planta" se refiere a plantas completas. El término "parte de la planta" incluye tejidos diferenciados e indiferenciados que incluyen, pero no se limitan a: órganos vegetales, tejidos vegetales, raíces, tallos, yemas, portainjertos, vástagos, estípulas, pétalos, hojas, flores, óvulos, pólenes, brácteas, pecíolos, entrenudos, corteza, pubescencia, macollos, rizomas, frondes, láminas, estambres, frutos, semillas, tejido tumoral y células vegetales (por ejemplo, células individuales, protoplastos, embriones y tejido calloso). Las células vegetales incluyen, sin limitación, células de semillas, cultivos en suspensión, embriones, regiones meristemáticas, tejido de callos, hojas, raíces, brotes, gametofitos, esporofitos, polen y micrósporas. El tejido vegetal puede estar en una planta o en un órgano vegetal, tejido o cultivo celular.

Como se usa en la presente descripción, cuando se habla de plantas, el término "óvulo" se refiere al gametofito femenino, mientras que el término "polen" significa el gametofito masculino.

Como se usa en la presente descripción, el término "tejido vegetal" se refiere a cualquier parte de una planta. Los ejemplos de órganos vegetales incluyen, pero no se limitan a, la hoja, tallo, raíz, tubérculo, semilla, rama, pubescencia, nódulo, axila de la hoja, flor, polen, estambre, pistilo, pétalo, pedúnculo, tallo, estigma, estilo, bráctea, fruto, tronco, carpelo, sépalo, antera, óvulo, pedicelo, aguja, cono, rizoma, estolón, brote, pericarpio, endospermo, placenta, baya, estambre y vaina foliar.

Como se usa en la presente descripción, el término "fenotipo" se refiere a los caracteres observables de una célula individual, cultivo celular, organismo (por ejemplo, una planta)

o grupo de organismos que resulta de la interacción entre la composición genética de ese individuo (es decir, genotipo) y el entorno.

Los términos "transgén" o "transgénico" como se usan en la presente descripción se refieren a al menos una secuencia de ácido nucleico que se toma del genoma de un organismo, o se produce sintéticamente, y que después se introduce en una célula u organismo o tejido hospedero de interés y que posteriormente se integra en el genoma del hospedero mediante enfoques de transformación o transfección "estable". Por el contrario, el término transformación o transfección o introducción "transitoria" se refiere a una forma de introducir herramientas moleculares que incluyen al menos un ácido nucleico (ADN, ARN, monocatenario o bicatenario o una mezcla de los mismos) y/o al menos una secuencia de aminoácidos, que opcionalmente comprende agentes químicos o biológicos adecuados, para lograr una transferencia a al menos un compartimento de interés de una célula, que incluye, pero no se restringe a, el citoplasma, un orgánulo, que incluye el núcleo, una mitocondria, una vacuola, un cloroplasto, o en una membrana, lo que da como resultado la transcripción y/o traducción y/o asociación y/o actividad de la al menos una molécula introducida sin lograr una integración o incorporación estable y, por tanto, la herencia de la respectiva al menos una molécula introducida en el genoma de una célula. El término "libre de transgén" se refiere a una condición en la que el transgén no está presente o no se encuentra en el genoma de una célula o tejido u organismo hospedero de interés.

Como se usa en la presente descripción, el término "cultivo de tejidos" indica una composición que comprende células aisladas del mismo tipo o de un tipo diferente o una colección de tales células organizadas en partes de una planta. Los tipos ilustrativos de cultivos de tejidos son protoplastos, callos, grupos de plantas y células vegetales que pueden generar cultivos de tejidos que están intactos en plantas o partes de plantas, como embriones, polen, flores, semillas, hojas, tallos, raíces, ápices radicales, anteras, pistilos, células meristemáticas, yemas axilares, ovarios, cubierta de la semilla, endospermo, hipocótilos, cotiledones y similares. El término "órgano vegetal" se refiere a tejido vegetal o un grupo de tejidos que constituyen una parte morfológica y funcionalmente distinta de una planta. "Progenie" comprende cualquier generación subsecuente de una planta.

Los métodos generales de bioquímica molecular y celular pueden encontrarse en libros de texto estándar como Molecular Cloning: A Laboratory Manual, 3ra Ed. (Sambrook y otros, HarBor Laboratory Press 2001);, Short Protocols in Molecular Biology 4ta Ed. (Ausubel y otros, eds., John Wiley & Sons 1999); Protein Methods (Bollag y otros, John

Wiley & Sons 1996); Nonviral Vectors for Gene Therapy (Wagner y otros, eds., Academic Press 1999); Viral Vectors (Kaplift & Loewy eds., Academic Press 1995); Immunology Methods Manual (I. Lefkovits ed., Academic Press 1997); y Cell and Tissue Culture: Laboratory Procedures in Biotechnology (Doyle & Griffiths, John Wiley & Sons 1998), cuyas descripciones se incorporan en la presente descripción como referencia.

Como se usa en la presente descripción, el término "Factor de Transcripción del Clado AGAMOUS" o "Factor de Transcripción del Clado AG" es un miembro de la subfamilia AGAMOUS (AG) de genes de caja MADS de tipo MIKC. Las proteínas "tipo MIKC" representan una clase de factores de transcripción del dominio MADS y están definidas por una estructura de dominio única: (1) 'M' - un dominio MADS de unión al ADN altamente conservado, (2) 'I' - un dominio, (3) 'K' - un dominio K similar a la queratina, y (4) 'C' - un dominio C-terminal. En algunas modalidades, el "Factor de Transcripción del Clado AGAMOUS" o "Factor de Transcripción del Clado AG" comprende además una región N-terminal. En otras modalidades, "Factor de Transcripción del Clado AGAMOUS" o "Factor de Transcripción del Clado AG" comprende los genes AG, SHP1, SHP2 y STK en plantas de la presente descripción, cada uno de los cuales tiene un motivo NN en el dominio M, un motivo YQQ en el dominio K y/o un R/Q (R o Q) en el dominio C.

Por "porción biológicamente activa" se entiende una porción de un péptido o polipéptido parental de longitud completa cuya porción retiene una actividad de la molécula parental. Por ejemplo, una porción biológicamente activa del polipéptido de la descripción conservará la capacidad de conferir resistencia a enfermedades, especialmente resistencia a patógenos fúngicos tales como *Fusarium*. Como se usa en la presente descripción, el término "porción biológicamente activa" incluye mutantes de delección y péptidos, por ejemplo de al menos aproximadamente 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 40, 50, 60, 70, 80, 90, 100, 120, 150, 300, 400, 500, 600, 700, 800, 900 o 1000 aminoácidos contiguos, que comprenden una actividad de una molécula parental. Pueden obtenerse porciones de este tipo mediante la aplicación de técnicas estándar de ácido nucleico recombinante o sintetizarse mediante el uso de técnicas convencionales de síntesis en fase líquida o sólida. Por ejemplo, puede hacerse referencia a la síntesis en solución o síntesis en fase sólida como se describe, por ejemplo, en el Capítulo 9 titulado "Síntesis de péptidos" por Atherton y Shephard, que se incluye en una publicación titulada " Synthetic Vaccines" editada por Nicholson y publicada por Blackwell Scientific Publications. Alternativamente, los péptidos pueden producirse mediante la digestión de un péptido o polipéptido de la descripción con

proteinasas tales como endoLys-C, endoArg-C, endoGlu-C y estafilococo V8-proteasa. Los fragmentos digeridos pueden purificarse, por ejemplo, mediante técnicas de cromatografía líquida de alta resolución (HPLC). También pueden usarse técnicas de ácidos nucleicos recombinantes para producir dichas porciones.

- 5 Por "corresponde a" o "correspondiente a" se entiende un polinucleótido (a) que tiene una secuencia de nucleótidos que es sustancialmente idéntica o complementaria a toda o una parte de una secuencia de polinucleótidos de referencia o (b) que codifica una secuencia de aminoácidos idéntica a una secuencia de aminoácidos en un péptido o proteína. Esta frase también incluye dentro de su alcance un péptido o polipéptido que
- 10 tiene una secuencia de aminoácidos que es sustancialmente idéntica a una secuencia de aminoácidos en un péptido o proteína de referencia.

Los términos "crecimiento" o "regeneración", como se usan en la presente descripción, significan hacer crecer una planta completa y diferenciada a partir de una célula vegetal, un grupo de células vegetales, una parte de la planta (que incluye las semillas) o una

15 parte de la planta (por ejemplo, de un protoplasto, callo o parte de tejido).

Como se usa en la presente descripción, el término "derivado de" se refiere al origen o fuente, y puede incluir moléculas de origen natural, recombinantes, no purificadas o purificadas. Un ácido nucleico o un aminoácido derivado de un origen o fuente puede tener todo tipo de cambios de nucleótidos o modificación de proteínas como se define

20 en otra parte de la presente descripción.

Por "obtenido de" se entiende que una muestra tal como, por ejemplo, un extracto de ácido nucleico o un extracto de polipéptido se aísla de, o se deriva de, una fuente particular. Por ejemplo, el extracto puede aislarse directamente de plantas, especialmente plantas monocotiledóneas y más especialmente plantas

25 monocotiledóneas no gramíneas como el plátano.

El término "patógeno" se usa en la presente descripción en su sentido más amplio para referirse a un organismo o agente infeccioso cuya infección de células de tejido vegetal viable provoca una respuesta a la enfermedad.

Por polipéptido "variante" se entiende un polipéptido derivado de la proteína nativa por delección (denominado truncamiento) o adición de uno o más aminoácidos al extremo N-terminal y/o C-terminal de la proteína nativa; delección o adición de uno o más aminoácidos en uno o más sitios en la proteína nativa; o sustitución de uno o más aminoácidos en uno o más sitios de la proteína nativa. Las proteínas variantes abarcadas por la presente descripción son biológicamente activas, es decir, aún poseen

30 la actividad biológica deseada de la proteína nativa, es decir, la actividad moduladora o

35

reguladora como se describe en la presente descripción. Dichas variantes pueden resultar de, por ejemplo, polimorfismo genético o de manipulación humana. Las variantes biológicamente activas de una proteína R nativa de la descripción tendrán al menos 40 %, 50 %, 60 %, 70 %, generalmente al menos 75 %, 80 %, 85 %, preferentemente aproximadamente 90 % a 95 % o más, y con mayor preferencia aproximadamente 98 % o más de identidad de secuencia con la secuencia de aminoácidos para la proteína nativa según se determina mediante programas de alineamiento de secuencia descritos en otra parte de la presente descripción mediante el uso de parámetros predeterminados. Una variante biológicamente activa de una proteína de la descripción puede diferir de esa proteína en tan solo 1-15 residuos de aminoácidos, tan solo 1-10, como 6-10, tan solo 5, tan solo 4, 3, 2 o incluso 1 residuo de aminoácido.

Las proteínas de la descripción pueden alterarse de diversas formas, lo que incluye sustituciones, deleciones, truncamientos e inserciones de aminoácidos. Los métodos para dichas manipulaciones generalmente se conocen en la técnica. Por ejemplo, las variantes de la secuencia de aminoácidos de las proteínas R pueden prepararse mediante mutaciones en el ADN. Los métodos para mutagénesis y alteraciones de la secuencia de nucleótidos se conocen bien en la técnica. Ver, por ejemplo, Kunkel (1985) Proc. Natl. Acad. Sci. USA 82:488-492; Kunkel y otros (1987) Methods in Enzymol. 154:367-382; Patente de Estados Unidos núm. 4,873,192; Walker y Gaastra, eds. (1983) Techniques in Molecular Biology (MacMillan Publishing Company, Nueva York) y las referencias citadas en estos. Puede encontrarse orientación en cuanto a las sustituciones de aminoácidos apropiadas que no afectan la actividad biológica de la proteína de interés en el modelo de Dayhoff y otros (1978) Atlas of Protein Sequence and Structure (Natl. Biomed. Res. Found., Washington, D.C.), incorporada en la presente descripción como referencia. Pueden ser preferentes las sustituciones conservadoras, tales como el intercambio de un aminoácido por otro que tenga propiedades similares.

Las deleciones o adiciones de sustituciones individuales que alteran, añaden o eliminan un solo aminoácido o un pequeño porcentaje de aminoácidos (típicamente menos del 5 %, más típicamente menos del 1 %) en una secuencia codificada son "variaciones conservadoramente modificadas", donde las alteraciones dan lugar a la sustitución de un aminoácido por un aminoácido químicamente similar. Las tablas de sustitución conservadoras que proporcionan aminoácidos funcionalmente similares son bien conocidas en la técnica. Los siguientes cinco grupos contienen cada uno aminoácidos

que son sustituciones conservadoras entre sí, Alifáticos: Glicina (G), Alanina (A), Valina (V), Leucina (L), Isoleucina (I); Aromático: Fenilalanina (F), Tirosina (Y), Triptófano (W); Con contenido de azufre: Metionina (M), Cisteína (C); Básico: Arginina (R), Lisina (K), Histidina (H); Ácido: Ácido aspártico (D), Ácido glutámico (E), Asparagina (N), Glutamina (Q). Ver también, Creighton, 1984. Además, las sustituciones, deleciones o adiciones individuales que alteran, agregan o eliminan un solo aminoácido o un pequeño porcentaje de aminoácidos en una secuencia codificada también son "variaciones modificadas de forma conservadora".

"Casete de expresión", como se usa en la presente descripción, significa una secuencia de ADN capaz de dirigir la expresión de una secuencia de nucleótidos particular en una célula huésped apropiada, que comprende un promotor operativamente unido a la secuencia de nucleótidos de interés que está operativamente unida a señales de terminación. También comprende típicamente secuencias necesarias para la traducción adecuada de la secuencia de nucleótidos. La región codificante normalmente codifica una proteína de interés pero también puede codificar un ARN funcional de interés, por ejemplo, ARN antisentido o un ARN sin traducir, en la dirección sentido o antisentido. El casete de expresión que comprende la secuencia de nucleótidos de interés puede ser quimérico, lo que significa que al menos uno de sus componentes es heterólogo con respecto a al menos uno de sus otros componentes. El casete de expresión también puede ser uno de origen natural pero que se haya obtenido en una forma recombinante útil para la expresión heteróloga. La expresión de la secuencia de nucleótidos en el casete de expresión puede estar bajo el control de un promotor constitutivo o de un promotor inducible que inicia la transcripción solo cuando la célula huésped está expuesta a algún estímulo externo particular. En el caso de un organismo multicelular, el promotor también puede ser específico para un tejido u órgano en particular o una etapa de desarrollo en animales y/o plantas, lo que incluye las especies de plátano.

Como se usa en la presente descripción, el término "vector", "plásmido" o "constructo" se refiere en términos generales a cualquier plásmido o virus que codifique un ácido nucleico exógeno. El término también debe interpretarse para incluir compuestos no plásmidos y no virales que facilitan la transferencia de ácido nucleico a viriones o células, tales como, por ejemplo, compuestos de polilisina y similares. El vector puede ser un vector viral que sea adecuado como vehículo de administración para la administración del ácido nucleico, o mutante del mismo, a una célula, o el vector puede ser un vector no viral que sea adecuado para el mismo propósito. Los ejemplos de vectores virales y no virales para el suministro de ADN a células y tejidos se conocen

bien en la técnica y se describen, por ejemplo, en Ma y otros (1997, Proc. Natl. Acad. Sci. USA 94: 12744-12746). Los ejemplos de vectores virales incluyen, pero no se limitan a, virus vegetales recombinantes. Los ejemplos no limitantes de virus de plantas incluyen, transfección (transitoria) mediada por TMV en tabaco (Tuipe, T-H. y otros (1993), J. Virology Meth, 42:227-239), virus de genomas de ADNss (por ejemplo, familia *Geminiviridae*), virus de transcripción inversa (por ejemplo, familias *Caulimoviridae*, *Pseudoviridae* y *Metaviridae*), virus con ARNs (por ejemplo, familias *Reoviridae* y *Partitiviridae*), virus con ARNs (-) (por ejemplo, familias *Rhabdoviridae* y *Bunyaviridae*), virus con ARNs (+) (por ejemplo, familias *Bromoviridae*, *Closteroviridae*, *Comoviridae*, *Luteoviridae*, *Potyviridae*, *Sequiviridae* y *Tombusviridae*) y viroides (por ejemplo, familias *Pospiviroidae* y *Avsunviroidae*). Puede encontrarse información detallada sobre la clasificación de virus de plantas en Fauquet y otros (2008, " Geminivirus strain demarcation and nomenclature"). *Archives of Virology* 153:783-821, incorporado en la presente descripción como referencia en su totalidad), y Khan y otros (Plant viruses as molecular pathogens; Publisher Routledge, 2002, ISBN 1560228954, 9781560228950). Los ejemplos de vectores no virales incluyen, pero no se limitan a, liposomas, derivados de poliamina de ADN y similares.

Además, "vector" se define para incluir, entre otros, cualquier plásmido, cósmido, fago o vector binario de *Agrobacterium* en forma lineal o circular, bicatenaria o monocatenaria, que puede o no ser autotransmisibile o movilizable, y que puede transformar procariotas o eucariotas hospederos, ya sea por integración en el genoma celular o exista extracromosómicamente (por ejemplo, plásmido de replicación autónoma con un origen de replicación).

Se incluyen específicamente los vectores lanzadera, por lo que se entiende un vehículo de ADN capaz, de forma natural o por diseño, de replicarse en dos organismos hospederos diferentes, que pueden seleccionarse entre actinomicetos y especies relacionadas, bacterias y eucariotas (por ejemplo, células de plantas superiores, mamíferos, levaduras u hongos).

Preferentemente, el ácido nucleico en el vector está bajo el control de un promotor apropiado u otros elementos reguladores para la transcripción en una célula huésped tal como una célula microbiana, por ejemplo bacteriana o vegetal, y está operativamente unido a él. El vector puede ser un vector de expresión bifuncional que funciona en múltiples huéspedes. En el caso del ADN genómico, este puede contener su propio promotor u otros elementos reguladores y en el caso del ADNc esto puede estar bajo el

control de un promotor apropiado u otros elementos reguladores para la expresión en la célula huésped.

Los "vectores de clonación" contienen típicamente uno o un pequeño número de sitios de reconocimiento de endonucleasas de restricción en los que pueden insertarse secuencias de ADN extrañas de una manera determinable sin pérdida de la función biológica esencial del vector, así como también un gen marcador que es adecuado para su uso en la identificación y selección de células transformadas con el vector de clonación. Los genes marcadores incluyen típicamente genes que proporcionan resistencia a tetraciclina, resistencia a higromicina o resistencia a ampicilina.

- 5
- 10 Como se usa en la presente descripción, el término "resistente", o "resistencia", describe una planta, línea o cultivo que muestra menos síntomas o síntomas reducidos de una plaga o patógeno biótico que una planta, línea o variedad susceptible (o más susceptible) a esa plaga o patógeno biótico. Estos términos se aplican de diversas formas para describir las plantas que no muestran síntomas, así como también las
- 15 plantas que presentan algunos síntomas, pero que aún pueden producir un producto comercializable con un rendimiento aceptable. Algunas líneas a las que se hace referencia como resistentes lo son solo en el sentido de que aún pueden producir un cultivo, aunque las plantas pueden parecer atrofiadas visualmente y el rendimiento se reduce en comparación con las plantas no infectadas. Según la definición de la
- 20 Federación Internacional de Semillas (ISF), una organización no gubernamental sin fines de lucro que representa a la industria de las semillas (ver " Definition of the Terms Describing the Reaction of Plants to Pests or Pathogens and to Abiotic Stresses for the Vegetable Seed Industry", Mayo de 2005), el reconocimiento de si una planta está afectada o sujeta a una plaga o patógeno puede depender del método analítico
- 25 empleado. La ISF define la resistencia como la capacidad de los tipos de plantas para restringir el crecimiento y desarrollo de una plaga o patógeno específico y/o el daño que causan en comparación con variedades de plantas susceptibles en condiciones ambientales y presión de plagas o patógenos similares. Los tipos de plantas resistentes aún pueden presentar algunos síntomas de enfermedades o daños. Se definen dos
- 30 niveles de resistencia. El término "resistencia alta/estándar" se usa para variedades de plantas que restringen en gran medida el crecimiento y desarrollo de la plaga o patógeno especificado bajo presión normal de plaga o patógeno en comparación con variedades susceptibles. La "resistencia moderada/intermedia" se aplica a los tipos de plantas que restringen el crecimiento y desarrollo de la plaga o patógeno especificado, pero exhiben
- 35 una mayor variedad de síntomas o daños en comparación con los tipos de plantas con

alta resistencia. Los tipos de plantas con resistencia intermedia mostrarán síntomas menos graves que las variedades de plantas susceptibles, cuando se cultivan en condiciones de campo y presión de patógenos similares. Los métodos para evaluar la resistencia son bien conocidos por los expertos en la técnica. Dicha evaluación puede realizarse mediante la observación visual de una planta o una parte de la planta (por ejemplo, hojas, raíces, flores, frutos, etc.) para determinar la gravedad de los síntomas. Por ejemplo, cuando a cada planta se le asigna una puntuación de resistencia en una escala de 1 a 5 en función de la gravedad de la reacción o los síntomas, siendo 1 la puntuación de resistencia aplicada a las plantas más resistentes (por ejemplo, sin síntomas o con los menores síntomas), y 5 el puntaje aplicado a las plantas con los síntomas más severos, entonces una línea se califica como resistente cuando al menos el 75 % de las plantas tienen un puntaje de resistencia en un nivel 1, 2 o 3, mientras que las líneas susceptibles son los que tienen más del 25 % de las plantas con una puntuación de 4 o 5. Si es posible una evaluación visual más detallada, entonces se puede usar una escala del 1 al 10 para ampliar el intervalo de puntuaciones y, con suerte, proporcionar una mayor distribución de puntuación entre las plantas que se evalúan.

Otro sistema de puntuación es una prueba de inoculación de raíces basada en el desarrollo de la necrosis después de la inoculación y su posición hacia el cotiledón (tal como una derivada de Bosland y otros, 1991), en donde 0 representa ningún síntoma después de la infección; 1 representa una pequeña necrosis en el hipocótilo después de la infección; 2 se encuentra una necrosis debajo de los cotiledones después de la infección; 3 significa necrosis por encima de los cotiledones después de la infección; 4 representa una necrosis por encima de los cotiledones junto con una marchitez de la planta después de la infección, mientras que, finalmente, 5 representa una planta muerta.

Además de dichas evaluaciones visuales, las evaluaciones de enfermedades pueden realizarse mediante la determinación de la biodensidad del patógeno en una planta o parte de la planta mediante microscopía electrónica y/o mediante métodos biológicos moleculares, tales como la hibridación de proteínas (por ejemplo, ELISA, medición de la densidad de la proteína del patógeno) y/o hibridación de ácidos nucleicos (por ejemplo, PCR-RT, medición de la densidad de ARN de patógenos). En dependencia de la combinación particular de patógeno/planta, puede determinarse que una planta es resistente al patógeno, por ejemplo, si tiene un ARN/ADN del patógeno y/o una densidad de proteína de aproximadamente el 50 %, aproximadamente el 40 % o

aproximadamente el 30 %, o aproximadamente el 20 %, o aproximadamente el 10 %, o aproximadamente el 5 %, o aproximadamente el 2 %, o aproximadamente el 1 %, o alrededor del 0,1 %, o aproximadamente el 0,01 %, o aproximadamente el 0,001 %, o aproximadamente el 0,0001 % del ARN/ADN y/o densidad de proteínas en una planta susceptible.

Los métodos usados en el mejoramiento de plantas para la resistencia a enfermedades son similares a los usados en el mejoramiento de otros caracteres. Es necesario conocer tanto como sea posible sobre la naturaleza de la herencia de los caracteres resistentes en la planta huésped y la existencia de razas fisiológicas o cepas del patógeno.

Como se usa en la presente descripción, el término "resistencia total" se refiere a la falta total de desarrollo del patógeno después de la infección, y puede ser el resultado de la falta de entrada del patógeno en la célula (sin infección inicial) o puede ser el resultado de falla del patógeno para multiplicarse en la célula e infectar las células subsecuentes (sin infección subliminal, sin diseminación). La presencia de resistencia total puede determinarse al establecer la ausencia de proteína patógena o ARN patógeno en las células de la planta, así como también la ausencia de cualquier síntoma de enfermedad en dicha planta, tras la exposición de dicha planta a una dosis infecciosa de patógeno (es decir, después de la 'infección'). Entre los criadores, este fenotipo a menudo se denomina "inmune". "Inmunidad", como se usa en la presente descripción, se refiere por tanto a una forma de resistencia caracterizada por la ausencia de replicación del patógeno incluso cuando el patógeno se transfiere activamente a las células, por ejemplo, mediante electroporación.

Como se usa en la presente descripción, el término "resistencia parcial" se denomina multiplicación reducida del patógeno en la célula, movimiento reducido (sistémico) del patógeno y/o desarrollo reducido de síntomas después de la infección. La presencia de resistencia parcial puede determinarse al establecer la presencia sistémica de baja concentración de proteína patógena o ARN patógeno en la planta y la presencia de síntomas de enfermedad disminuidos o retardados en dicha planta tras la exposición de dicha planta a una dosis infecciosa de patógeno. La concentración de proteína puede determinarse mediante el uso de un método de detección cuantitativo (por ejemplo, un método ELISA o una reacción cuantitativa en cadena de la polimerasa con transcriptasa inversa (PCR-RT)). Entre los criadores, este fenotipo a menudo se denomina "resistencia intermedia".

Como se usa en la presente descripción, el término "tolerante" se usa en la presente descripción para indicar un fenotipo de una planta en donde los síntomas de la

enfermedad permanecen ausentes tras la exposición de dicha planta a una dosis infecciosa de patógeno, por lo que la presencia de una infección patógena sistémica o local, multiplicación de patógenos, puede establecerse al menos la presencia de secuencias genómicas de patógenos en células de dicha planta y/o integración genómica de las mismas. Por lo tanto, las plantas tolerantes son resistentes a la expresión de síntomas pero portadoras asintomáticas del patógeno. A veces, las secuencias de patógenos pueden estar presentes o incluso multiplicarse en las plantas sin provocar síntomas de enfermedad. Este fenómeno también se conoce como "infección latente". En las infecciones latentes, el patógeno puede existir en una forma oculta no infecciosa verdaderamente latente, posiblemente como un genoma integrado o un agente episomal (de modo que la proteína del patógeno no pueda encontrarse en el citoplasma, mientras que los protocolos de PCR pueden indicar la presencia de secuencias de ácido nucleico del patógeno) o como un agente infeccioso y de replicación continua. Un patógeno reactivado puede propagarse e iniciar una epidemia entre los contactos susceptibles. La presencia de una "infección latente" es indistinguible de la presencia de un fenotipo "tolerante" en una planta.

Como se usa en la presente descripción, el término "susceptible" se usa en la presente descripción para referirse a una planta con ausencia o que prácticamente no tiene resistencia al patógeno, lo que da como resultado la entrada del patógeno en la planta y la multiplicación y diseminación sistémica del patógeno, lo que da como resultado síntomas de enfermedad. Por lo tanto, el término "susceptible" equivale a "no resistente".

Como se usa en la presente descripción, el término "descendencia" se refiere a cualquier planta resultante de una reproducción vegetativa o sexual de una o más plantas parentales o descendientes de las mismas. Por ejemplo, puede obtenerse una planta descendiente mediante la clonación o autofecundación de una planta parental o mediante el cruzamiento de dos plantas parentales e incluir autofecundaciones así como también F1 o F2 o incluso generaciones posteriores. Una F1 es una descendencia de primera generación producida a partir de padres, al menos uno de los cuales se usa por primera vez como donante de un rasgo, mientras que la descendencia de la segunda generación (F2) o generaciones posteriores (F3, F4, etc.) son especímenes producidos a partir de autofecundaciones de F1, F2, etc. Una F1 puede ser (y generalmente es) un híbrido resultante de un cruce entre dos progenitores de mejoramiento verdaderos (el mejoramiento verdadero es homocigoto para un rasgo), mientras que una F2 puede ser

(y generalmente lo es) una descendencia resultante de la autopolinización de dichos híbridos F1.

Como se usa en la presente descripción, los términos "dicotiledón", "dicotiledónea" y "dicotiledóneas" se refieren a una planta con flores que tiene un embrión que contiene dos mitades de semillas o cotiledones. Los ejemplos incluyen tabaco; tomate; las legumbres, que incluyen guisantes, alfalfa, trébol y soja; robles; arces; rosas; mentas; calabazas; margaritas; nueces; cactus; violetas y ranúnculos.

Como se usa en la presente descripción, los términos "monocotiledón", "monocotiledónea" o "monocotiledóneas" se refieren a cualquier subclase (Monocotyledoneae) de plantas con flores que tienen un embrión que contiene solo una hoja de semilla y que generalmente tiene hojas de nervaduras paralelas, partes de flores en múltiplos de tres, y no tienen crecimiento secundario en tallos y raíces. Los ejemplos incluyen plátano, narcisos, caña de azúcar, jengibre, lirio, orquídea, arroz, maíz, pastos, tales como festuca alta, pasto de cabra y pasto azul de Kentucky; cereales, tales como trigo, avena y cebada; lirios; cebolla y palma.

Como se usa en la presente descripción, el término "gen" se refiere a cualquier segmento de ADN asociado con una función biológica. Por tanto, los genes incluyen, pero no se limitan a, secuencias codificantes y/o las secuencias reguladoras necesarias para su expresión. Los genes también pueden incluir segmentos de ADN no expresados que, por ejemplo, forman secuencias de reconocimiento para otras proteínas. Los genes pueden obtenerse de una variedad de fuentes, incluida la clonación de una fuente de interés o la síntesis a partir de información de secuencia conocida o predicha, y pueden incluir secuencias diseñadas para tener los parámetros deseados.

Como se usa en la presente descripción, el término "genotipo" se refiere a la composición genética de una célula individual, cultivo celular, tejido, organismo (por ejemplo, una planta) o grupo de organismos.

Como se usa en la presente descripción, el término "alelo" o "alelos" significa cualquiera de una o más formas alternativas de un gen, que se relacionan con al menos un rasgo o característica. En una célula diploide, los dos alelos de un gen determinado ocupan loci correspondientes en un par de cromosomas homólogos. Dado que la presente descripción se refiere a QTL, es decir, regiones genómicas que pueden comprender uno o más genes o secuencias reguladoras, en algunos casos es más preciso referirse a "haplotipo" (es decir, un alelo de un segmento cromosómico) en lugar de "alelo", sin embargo, en esos casos, debe entenderse que el término "alelo" comprende el término "haplotipo". Los alelos se consideran idénticos cuando expresan un fenotipo similar. Las

diferencias en la secuencia son posibles pero no importantes siempre que no influyan en el fenotipo.

Como se usa en la presente descripción, el término "locus" (plural: "loci") se refiere a cualquier sitio que se haya definido genéticamente. Un locus puede ser un gen, o parte de un gen, o una secuencia de ADN que tiene alguna función reguladora y puede estar ocupada por diferentes secuencias.

Como se usa en la presente descripción, el término "marcador molecular" o "marcador genético" se refiere a un indicador que se usa en métodos para visualizar diferencias en las características de las secuencias de ácido nucleico. Ejemplos de estos indicadores son marcadores de polimorfismo de longitud de fragmentos de restricción (RFLP), marcadores de polimorfismo de longitud de fragmentos amplificados (AFLP), polimorfismos de nucleótido único (SNP), mutaciones de inserción, marcadores de microsatélites (SSR), regiones amplificadas caracterizadas por secuencia (SCAR), marcadores de secuencia polimórfica amplificada escindida (CAPS) o marcadores de isoenzimas o combinaciones de los marcadores descritos en la presente descripción, que definen una ubicación genética y cromosómica específica. El mapeo de marcadores moleculares en la vecindad de un alelo es un procedimiento que puede ser realizado con bastante facilidad por una persona promedio experta en técnicas de biología molecular, técnicas que se describen, por ejemplo, en Lefebvre y Chevre, 1995; Lorez y Wenzel, 2007, Srivastava y Narula, 2004, Meksem y Kahl, 2005, Phillips y Vasil, 2001. Puede encontrarse información general sobre la tecnología AFLP en Vos y otros (1995, AFLP: a new technique for DNA fingerprinting, Nucleic Acids Res. 1995 November 11; 23(21): 4407–4414).

Como se usa en la presente descripción, el término "hemicigoto" se refiere a una célula, tejido u organismo en el que un gen está presente solo una vez en un genotipo, como un gen en una célula u organismo haploide, un gen ligado al sexo en el sexo heterogamético, o un gen en un segmento de cromosoma en una célula u organismo diploide donde se ha eliminado su segmento asociado.

Como se usa en la presente descripción, el término "heterocigoto" se refiere a una célula o planta individual diploide o poliploide que tiene diferentes alelos (formas de un gen dado) presentes al menos en un locus.

Como se usa en la presente descripción, el término "heterocigoto" se refiere a la presencia de diferentes alelos (formas de un gen dado) en un locus génico particular.

Como se usa en la presente descripción, el término "homocigoto" se refiere a una célula o planta individual que tiene los mismos alelos en uno o más loci.

Como se usa en la presente descripción, el término "homocigoto" se refiere a la presencia de alelos idénticos en uno o más loci en segmentos cromosómicos homólogos.

5 Como se usa en la presente descripción, el término "homólogo" u "homólogos" se conocen en la técnica y se refiere a secuencias relacionadas que comparten un ancestro común o un miembro de la familia y se determina en base al grado de identidad de secuencia. Los términos "homología", "homólogo", "sustancialmente similar" y "correspondiente sustancialmente" se usan indistintamente en la presente descripción. Los homólogos generalmente controlan, median o influyen en las mismas vías bioquímicas o en vías similares, pero los homólogos particulares pueden dar lugar a fenotipos diferentes. Por lo tanto, se entiende, como apreciarán los expertos en la técnica, que la descripción abarca más que las secuencias ejemplares específicas. Estos términos describen la relación entre un gen que se encuentra en una especie, subespecie, variedad, cultivo o cepa y el gen correspondiente o equivalente en otra especie, subespecie, variedad, cultivo o cepa. Para los propósitos de esta descripción, se comparan secuencias homólogas.

El término "homólogo" se usa a veces para aplicar a la relación entre genes separados por el evento de especiación (ver "ortólogo") o la relación entre genes separados por el evento de duplicación genética (ver "parálogo").

20 El término "homeólogo" se refiere a un gen o cromosoma homeólogo, que resulta de eventos de poliploidía o duplicación cromosómica. Esto contrasta con el término más común "homólogo", que se definió arriba.

El término "ortólogo" se refiere a genes de diferentes especies que evolucionaron a partir de un gen ancestral común por especiación. Normalmente, los ortólogos conservan la misma función en el curso de la evolución. La identificación de ortólogos es fundamental para la predicción fiable de la función genética en genomas recién secuenciados.

25 El término "parálogo" se refiere a genes relacionados por duplicación dentro de un genoma. Si bien los ortólogos generalmente conservan la misma función en el curso de la evolución, los parálogos pueden desarrollar nuevas funciones, incluso si están relacionadas con la original.

30 Se cree o se sabe que las "secuencias homólogas" u "homólogos" u "ortólogos" están relacionadas funcionalmente. Una relación funcional puede indicarse de varias formas, que incluyen, pero no se limitan a: (a) grado de identidad de secuencia y/o (b) función biológica igual o similar. Preferentemente, se indican tanto (a) como (b). El grado de identidad de secuencia puede variar, pero en una modalidad, es al menos 50 % (cuando

- se usan programas de alineamiento de secuencia estándar conocidos en la técnica), al menos 60 %, al menos 65 %, al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos aproximadamente 91 %, al menos aproximadamente 92 %, al menos aproximadamente 93 %, al menos aproximadamente 94 %, al menos
- 5 aproximadamente 95 %, al menos aproximadamente 96 %, al menos aproximadamente el 97 %, al menos aproximadamente el 98 %, o al menos el 98,5 %, o al menos aproximadamente el 99 %, o al menos el 99,5 %, o al menos el 99,8 %, o al menos el 99,9 %. La homología puede determinarse mediante el uso de programas de software fácilmente disponibles en la técnica, tales como los discutidos en *Current Protocols in*
- 10 *Molecular Biology* (FM Ausubel y otros, Eds., 1987) Suplemento 30, sección 7.718, Tabla 7.71. Algunos programas de alineamiento son MacVector (Oxford Molecular Ltd, Oxford, Reino Unido) y ALIGN Plus (Scientific and Educational Software, Pensilvania). Otros programas de alineamiento no limitantes incluyen Sequencher (Gene Codes, Ann Arbor, Michigan), AlignX y Vector NTI (Invitrogen, Carlsbad, CA).
- 15 Como se usa en la presente descripción, el término "híbrido" se refiere a cualquier célula, tejido o planta individual resultante de un cruce entre padres que difieren en uno o más genes.
- Como se usa en la presente descripción, el término "consanguíneo" o "línea consanguínea" se refiere a una cepa de mejoramiento relativamente auténtica.
- 20 El término "planta convertida de un solo alelo", como se usa en la presente descripción, se refiere a aquellas plantas que se desarrollan mediante una técnica de fitomejoramiento llamada retrocruzamiento, en donde esencialmente todas las características morfológicas y fisiológicas deseadas de una endogamia se recuperan además del alelo simple transferido a la endogamia mediante la técnica de
- 25 retrocruzamiento.
- Como se usa en la presente descripción, el término "línea" se usa ampliamente para incluir, pero no se limita a, un grupo de plantas propagadas vegetativamente a partir de una sola planta parental, mediante técnicas de cultivo de tejidos o un grupo de plantas endogámicas que son genéticamente muy similares debido a la descendencia de
- 30 progenitores(s) común(es). Se dice que una planta "pertenece" a una línea particular si (a) es una planta transformante primaria (T0) regenerada a partir de material de esa línea; (b) tiene un pedigrí compuesto por una planta T0 de esa línea; o (c) es genéticamente muy similar debido a un ancestro común (por ejemplo, por endogamia o autofecundación). En este contexto, el término "pedigrí" denota el linaje de una planta,
- 35 por ejemplo, en términos de cruces sexuales afectados de manera que un gen o una

combinación de genes, en condición heterocigótica (hemicigótica) u homocigótica, imparte un rasgo deseado a la planta.

Como se usa en la presente descripción, los términos "introgresión" o "introgresado" se refieren al proceso mediante el cual los genes de una especie, variedad o cultivo se mueven al genoma de otra especie, variedad o cultivo, al cruzar esas especies. El cruce puede ser natural o artificial. El proceso se puede completar opcionalmente retrocruzando con el padre recurrente, en cuyo caso la introgresión se refiere a la infiltración de los genes de una especie en el acervo genético de otra a través del retrocruzamiento repetido de un híbrido interespecífico con uno de sus padres. Una introgresión también puede describirse como un material genético heterólogo integrado de forma estable en el genoma de una planta receptora.

Como se usa en la presente descripción, el término "población" significa una colección genéticamente homogénea o heterogénea de plantas que comparten una derivación genética común.

Como se usa en la presente descripción, el término "variedad" o "cultivo" significa un grupo de plantas similares que por características estructurales y rendimiento se pueden identificar de otras variedades dentro de la misma especie. El término "variedad", como se usa en la presente descripción, tiene un significado idéntico a la definición correspondiente en el Convenio Internacional para la Protección de Nuevas Variedades de Plantas (tratado UPOV), del 2 de diciembre de 1961, revisado en Ginebra el 10 de noviembre de 1972, el 23 de octubre de 1978 y el 19 de marzo de 1991. Por tanto, "variedad" significa una agrupación de plantas dentro de un único taxón botánico del rango más bajo conocido, cuya agrupación, independientemente de que se cumplan plenamente las condiciones para la concesión de un derecho de obtentor, puede i) definirse mediante la expresión de las características resultante de un determinado genotipo o combinación de genotipos, ii) distinguirse de cualquier otro grupo de plantas por la expresión de al menos una de dichas características y iii) considerarse como una unidad en cuanto a su idoneidad para propagarse inalterada.

Como se usa en la presente descripción, el término "selección en masa" se refiere a una forma de selección en la que se seleccionan plantas individuales y la siguiente generación se propaga a partir del agregado de sus semillas. En las especificaciones de la presente descripción se describen más detalles de la selección de masas.

Como se usa en la presente descripción, el término "polinización abierta" se refiere a una población de plantas que está libremente expuesta a algún flujo de genes, en contraposición a una cerrada en la que existe una barrera eficaz para el flujo de genes.

- Como se usa en la presente descripción, los términos "población de polinización abierta" o "variedad de polinización abierta" se refieren a plantas normalmente capaces de al menos alguna fertilización cruzada, seleccionadas según un estándar, que pueden mostrar variación pero que también tienen uno o más características genotípicas o fenotípicas por las que puede diferenciarse la población o la variedad de otras. Un híbrido, que no tiene barreras para la polinización cruzada, es una población de polinización abierta o una variedad de polinización abierta.
- Como se usa en la presente descripción, el término "autocruzamiento" o "autopolinización" significa que el polen de una flor en una planta se aplica (artificial o naturalmente) al óvulo (estigma) de la misma o diferente flor en la misma planta.
- Como se usa en la presente descripción, los términos "cruce", "cruzamiento" o "polinización cruzada" o "mejoramiento cruzado" se refieren al proceso mediante el cual el polen de una flor en una planta se aplica (artificial o naturalmente) al óvulo (estigma) de una flor en otra planta.
- Como se usa en la presente descripción, el término "derivado de" se refiere al origen o fuente, y puede incluir moléculas de origen natural, recombinantes, no purificadas o purificadas. Un ácido nucleico o un aminoácido derivado de un origen o fuente puede tener todo tipo de cambios de nucleótidos o modificación de proteínas como se define en otra parte de la presente descripción.
- El término "cebador", como se usa en la presente descripción, se refiere a un oligonucleótido que es capaz de hibridar con el objetivo de amplificación lo que permite que una ADN polimerasa se una, y sirve de esta manera como punto de inicio de la síntesis de ADN cuando se coloca en condiciones en las que la síntesis del producto de la extensión del cebador es inducida, es decir, en presencia de nucleótidos y un agente de polimerización tal como la ADN polimerasa y a una temperatura y pH adecuados. El cebador (de amplificación) es preferentemente monocatenario para una máxima eficacia en la amplificación. Preferentemente, el cebador es un oligodesoxirribonucleótido. El cebador debe ser lo suficientemente largo para cebar la síntesis de productos de extensión en presencia del agente de polimerización. Las longitudes exactas de los cebadores dependerán de muchos factores, incluida la temperatura y la composición (el contenido de A/T y G/C) del cebador. Un par de cebadores bidireccionales consta de un cebador directo y uno inverso como se usa comúnmente en la técnica de la amplificación de ADN, tal como en la amplificación por PCR.
- Una sonda comprende un ácido nucleico aislado identificable que reconoce una secuencia de ácido nucleico objetivo. Una sonda incluye un ácido nucleico que está

- unido a una ubicación direccionable, un marcador detectable u otra molécula informadora y que se hibrida con una secuencia objetivo. Los marcadores típicos incluyen isótopos radiactivos, sustratos enzimáticos, cofactores, ligandos, agentes quimioluminiscentes o fluorescentes, haptenos y enzimas. Los métodos para el
- 5 etiquetado y la orientación en la elección de etiquetas apropiadas para diversos propósitos se discuten, por ejemplo, en Sambrook y otros. (ed.), *Molecular Cloning: A Laboratory Manual*, 2da ed., vol. 1-3, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989 y Ausubel y otros. *Short Protocols in Molecular Biology*, 4th ed., John Wiley & Sons, Inc., 1999.
- 10 Los métodos para preparar y usar sondas y cebadores de ácidos nucleicos se describen, por ejemplo, en Sambrook y otros (ed.), *Molecular Cloning: A Laboratory Manual*, 2nd ed., vol. 1-3, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989; Ausubel y otros, *Short Protocols in Molecular Biology*, 4ta ed., John Wiley & Sons, Inc., 1999; y Innis y otros, *PCR Protocols, A Guide to Methods and Applications*, Academic
- 15 Press, Inc., San Diego, CA, 1990. Los pares de cebadores de amplificación pueden derivarse de una secuencia conocida, por ejemplo, mediante el uso de programas informáticos destinados para ese fin, tales como PRIMER (Versión 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge, MA). Un experto en la técnica puede entender que la especificidad de una sonda o cebador particular aumenta con su
- 20 longitud. De esta forma, para obtener una mayor especificidad, se pueden seleccionar sondas y cebadores que comprendan al menos 20, 25, 30, 35, 40, 45, 50 o más nucleótidos consecutivos de unas secuencias de nucleótidos objetivo.
- Para las amplificaciones por PCR de los polinucleótidos descritos en la presente descripción, se pueden diseñar cebadores de oligonucleótidos para su uso en
- 25 reacciones de PCR para amplificar las secuencias de ADN correspondientes de ADNc o ADN genómico extraído de cualquier organismo de interés. Los métodos para diseñar cebadores de PCR y clonación de PCR se conocen generalmente en la técnica y se describen en Sambrook y otros, (2001) *Molecular Cloning: A Laboratory Manual* (3ra ed., Cold Spring Harbor Laboratory Press, Plainview, Nueva York). Ver también Innis y
- 30 otros, eds. (1990) *PCR Protocols: A Guide to Methods and Applications* (Academic Press, Nueva York); Innis y Gelfand, eds. (1995) *PCR Strategies* (Academic Press, New York); e Innis y Gelfand, eds. (1999) *PCR Methods Manual* (Academic Press, Nueva York). Los métodos conocidos de PCR incluyen, pero no se limitan a, métodos que usan cebadores emparejados, cebadores anidados, cebadores específicos únicos, cebadores

degenerados, cebadores específicos de genes, cebadores específicos de vector, cebadores parcialmente mal emparejados y similares.

La presente descripción proporciona una secuencia de ácido nucleico aislada que comprende una secuencia seleccionada del grupo que consiste en *FusR* 1, homólogos de *FusR*1, ortólogos de *FusR*1, parálogos de *FusR*1 y fragmentos y variaciones de los mismos. En una modalidad, la presente descripción proporciona un polinucleótido aislado que codifica una proteína producida por la secuencia de ácido nucleico para *FusR*1, que comprende una secuencia de ácido nucleico que comparte al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos el 90 %, al menos el 91 %, al menos el 92 %, al menos el 93 %, al menos el 94 %, al menos el 95 %, al menos el 96 %, al menos el 97 %, al menos el 98 %, al menos el 99 %, al menos 99,1 %, al menos 99,2 %, al menos 99,3 %, al menos 99,4 %, al menos 99,5 %, al menos 99,6 %, al menos 99,7 %, al menos 99,8 % o al menos 99,9 % de identidad con *FusR*1.

Los métodos de alineamiento de secuencias para comparación se conocen bien en la técnica. Diversos programas y algoritmos de alineamiento se describen en: Smith y Waterman (*Adv. Appl. Math.*, 2:482, 1981); Needleman y Wunsch (*J. Mol. Biol.*, 48:443, 1970); Pearson y Lipman (*Proc. Natl. Acad. Sci.*, 85:2444, 1988); Higgins y Sharp (*Gene*, 73:237-44, 1988); Higgins y Sharp (*CABIOS*, 5:151-53, 1989); Corpet y otros, (*Nuc. Acids Res.*, 16:10881-90, 1988); Huang y otros, (*Comp. Appls Biosci.*, 8:155-65, 1992); y Pearson y otros, (*Meth. Mol. Biol.*, 24:307-31, 1994). Altschul y otros, (*Nature Genet.*, 6:119-29, 1994) presenta una consideración detallada de los métodos de alineamiento de secuencias y los cálculos de homología.

La presente descripción también proporciona un gen quimérico que comprende la secuencia de ácido nucleico aislada de cualquiera de los polinucleótidos descritos anteriormente unida operativamente a secuencias reguladoras adecuadas.

La presente descripción también proporciona un constructo recombinante que comprende el gen quimérico como se describió anteriormente. En una modalidad, dicho constructo recombinante es un constructo de silenciamiento de genes, tal como se usa en el silenciamiento de genes de ARNi. En otra modalidad, dicho constructo recombinante es un constructo de edición de genes, tal como se usa en el sistema de edición de genes CRISPR-Cas.

Los vectores de expresión de la presente descripción pueden incluir al menos un marcador seleccionable. Dichos marcadores incluyen genes de dihidrofolato reductasa, G418 o resistencia a neomicina para cultivo de células eucariotas y de resistencia a tetraciclina, kanamicina o ampicilina para cultivo en *E. coli* y otras bacterias.

La presente descripción también proporciona una célula hospedadora transformada que comprende el gen quimérico como se describió anteriormente. En una modalidad, dicha célula huésped se selecciona del grupo que consiste en bacterias, levaduras, hongos filamentosos, algas, animales y plantas que incluyen, pero no se limitan a, el género *Musa*.

Estas secuencias permiten el diseño de cebadores y sondas de genes específicos para *FusR1*, homólogos de *FusR1*, ortólogos de *FusR1*, homeólogos de *FusR1*, parálogos de *FusR1* y fragmentos y variaciones de los mismos.

II. Modulación de la resistencia a enfermedades

La presente descripción se refiere a polinucleótidos y/o polipéptidos de *FusR1* recientemente identificado (*Fusarium* Resistencia 1) y métodos para modular, estimular o mejorar la resistencia a enfermedades en plantas, provocadas por patógenos. Los patógenos de la descripción incluyen, pero no se limitan a, bacterias, hongos, virus o viroides, nematodos, insectos y similares.

Los patógenos bacterianos incluyen pero no se limitan a *Pseudomonas avenae* subsp. *avenae*, *Xanthomonas campestris* pv. *holcicola*, *Enterobacter dissolvens*, *Erwinia dissolvens*, *Erwinia carotovora* subsp. *carotovora*, *Erwinia chrysanthemi* pv. *zeae*, *Pseudomonas andropogonis*, *Pseudomonas syringae* pv. *coronafaciens*, *Clavibacter michiganensis* subsp., *Corynebacterium michiganense* pv. *nebraskense*, *Pseudomonas syringae* pv. *syringae*, bacterias herniparasitarias (ver en hongos), *Bacillus subtilis*, *Erwinia stewartii* y *Spiroplasma kunkelii*.

Los patógenos incluyen pero no se limitan a: *Colleotrichum graminicola*, *Glomerella graminicola* Politis, *Glomerella lucumanensis*, *Aspergillus flavus*, *Rhizoctonia solani* Kuhn, *Thanatephorus cucumeris*, *Acremonium strictum* W. Gams, *Cephalosporium acremonium* Auct. non Corda Black *Lasioidiplodia theobromae*=*BoIr odiplodia* y *theobromae* Borde blanco *Marasmiellus* sp., *Physoderma maydis*, *Cephalosporium Corticium sasakii*, *Curvularia clavata*, *C. maculans*, *Cochhobolus eragrostidis*, *Curvularia inaequahs*, *C. intermedia* (teleomorfo: *Cochhobolus intermedius*), *Curvularia lunata* (teleomorfo: *Cochliobolus lunatus*), *Curvularia pallescens* (teleomorfo: *ochliobolus pallescens*), *Curvularia senegalensis*, *C. luberculata* (teleomorfo: *Cochliobolus tuberculatus*), *Didymella exitalis* *Diplodiaftumenti* (teleomorfo: *Botryosphaeriafestucae*), *Diplodia maydis*=*Stenocarpella maydis*, *Stenocarpella macrospora*=*Diplodia macrospora*, *Sclerophthora rayssiae* var. *zeae*, *Sclerophthora macrospora*=*Sclerospora macrospora*, *Sclerospora graminicola*, *Peronosclerospora maydis*=*Sclerospora maydis*, *Peronosclerospora philippinensis*, *Sclerospora philippinensis*, *Peronosclerospora*

- sorghi=*Sclerospora sorghi*, *Peronosclerospora spontanea*=*Sclerospora spontanea*, *Peronosclerospora sacchari*=*Sclerospora sacchari*, *Nigrospora oryzae* (teleomorfo: *Khuskia oryzae*) *A. lternaria alternata*=*A. tenuis*, *Aspergillus glaucus*, *A. niger*, *Aspergillus* spp., *Botrytis cinerea*, *Cunninghamella* sp., *Curvularia pallescens*,
5 *Doratomyces slemonitis*=*Cephalotrichum slemonitis*, *Fusarium culmorum*, *Gonatobotrys simplex*, *Pithomyces maydis*, *Rhizopus microsporus* Tiegh., *R. stolonifer*=*R. nigricans*, *Scopulariopsis brumptii*, *Claviceps gigantea* (anamorfo: *Sphacelia* sp.) *Aureobasidium zeae*=*Kabatiella zeae*, *Fusarium subglutinans*=*F. moniliforme* var. *subglutinans*, *Fusarium moniliforme*, *Fusarium avenaceum* (teleomorfo: *Gibberella avenacea*),
10 *Botryosphaeria zeae*=*Physalospora zeae* (anamorfo: *Allacrophoma zeae*), *Cercospora sorghi*=*C. sorghi* var. *maydis*, *Helminthosporium pedicellatum* (teleomorfo: *Selosphaeriapedicellata*), *Cladosporium cladosporioides*=*Hormodendrum cladosporioides*, *C. herbarum* (teleomorfo: *Mycosphaerella tassiana*), *Cephalosporium maydis*, *A. lternaria alternata*, *A. scochyta maydis*, *A. tritici*, *A. zeicola*, *Bipolaris victoriae*,
15 *Helminthosporium victoriae* (teleomorfo: *Cochhoholus victoriae*), *C. sativus* (anamorfo: *Bipolaris sorokiniana*=*H. sorokinianum*=*H. sativum*), *Epicoccum nigrum*, *Exserohilum prolatum*=*Drechslera prolata* (teleomorfo: *Setosphaeria prolata*), *Graphium penicillioides*, *Leptosphaeria maydis*, *Leptothyrium zeae*, *Ophiosphaerella herpotricha* (anamorfo: *Scolecosporella* sp.), *Pataphaeosphaeria michotii*, *Phoma* sp., *Septoria zeae*, *S. zeicola*, *S. zeina* *Setosphaeria turcica*, *Exserohilum turcicum*=*Helminthosporium turcicum*, *Cochhoholus carbonum*, *Bipolaris zeicola*=*Helminthosporium carbonum*,
20 *Penicillium* spp., *P. chrysogenum*, *P. expansum*, *P. oxalicum*, *Phaeocytostroma ambiguum*, *Phaeocytoporella zeae*, *Phaeosphaeria maydis*=*Sphaerulina maydis*, *Botryosphaeria festucae*=*Physalospora zeicola* (anamorfo: *Diplodia fumenfi*),
25 *Herniparasitic bacteria and fungi* *Pyrenochaeta Phoma terrestris*=*Pyrenochaeta terrestris*, *Pythium* spp., *P. arrhenomanes*, *P. graminicola*, *Pythium aphanidermatum*=*P. hutleri* L., *Rhizoctonia zeae* (teleomorfo: *Waitea circinata*), *Rhizoctonia solani*, minor *A. lternaria alternata*, *Cercospora sorghi*, *Dictochaeta frutis*, *Fusarium acuminatum* (teleomorfo: *Gibberella acuminata*), *E. equiseti* (teleomorph: *G. intricans*), *E. oxysporum*, *E. pallidroseum*, *E. poae*, *E. roseum*, *G. cyanogena* (anamorfo: *E. sulphureum*), *Microdochium holleyi*, *Mucor* sp., *Periconia circinata*, *Phytophthora cactorum*, *P. drechsleri*, *P. nicotianae* var. *parasitica*, *Phytophthora* spp., *Rhizopus arrhizus*, *Setosphaeria rostrata*, *Exserohilum rostratum*=*Helminthosporium rostratum*, *Puccinia sorghi*, *Physopella pallescens*, *P. zeae*, *Sclerotium rofsii* Sacc.
30 (teleomorfo: *Athelia rofsii*), *Bipolaris sorokiniana*, *B. zeicola*=*Helminthosporium*
35

carbonum, *Diplodia maydis*, *Exserohilum pedicellatum*, *Exserohilum furcicum*=*Helminthosporium turcicum*, *Fusarium avenaceum*, *E. culmorum*, *E. moniliforme*, *Gibberella zeae* (anamorph—*E. graminearum*), *Macrophomina phaseolina*, *Penicillium* spp., *Phomopsis* sp., *Pythium* spp., *Rhizoctonia solani*, *R. zeae*, *Sclerotium rolfsii*, *Spicaria* sp., *Selenophoma* sp., *Gaeumannomyces graminis*, *Myrothecium gramineum*, *Monascus purpureus*, *M. ruber* Smut, *Ustilago zeae*=*U. maydis* Smut, *Ustilaginoidea virens* Smut, *Sphacelotheca reiliana*=*Sporisorium holci*, *Cochliobolus heterostrophus* (anamorfo: *Bipolaris maydis*=*Helminthosporium maydis*), *Stenocarpella macrospora*=*Diplodia macrospora*, *Cercospora sorghi*, *Fusarium episphaeria*, *E. merismoides*, *F. oxysporum* Schlechtend, *Fusarium oxysporum* f. sp. *cubense* (Foc), *Fusarium* spp., *E. poae*, *E. roseum*, *E. solani* (teleomorfo: *Nectria haematococca*), *F. tricinctum*, *Mariannaea elegans*, *Mucor* sp., *Rhopoglyphus zeae*, *Spicaria* sp., *Aspergillus* spp., *Penicillium* spp., *Trichoderma viride*=*T. lignorum* teleomorfo: *Hypocrea* sp., *Stenocarpella maydis*=*Diplodia zeae*, *Ascochyta ischaemi*, *Phyllosticta maydis* (teleomorfo: *Mycosphaerella zeae-maydis*), *Mycosphaerella fijiensis*, *Pseudocercospora* (*Paracercospora*) *fijiensis* y *Gloeocercospora sorghi*.

Los virus o viroides incluyen, pero no se limitan a: el mosaico del virus del mosaico estriado del trigo americano (AWSMV), el virus del mosaico estriado de la cebada (BSMV), el virus del enanismo amarillo de la cebada (BYDV), el virus del racimo de plátano, el virus del mosaico de Brome (BMV), el virus moteado clorótico de los cereales (CCMV), virus de bandas de venas cloróticas del maíz (CCVBV), virus del moteado clorótico del maíz (MCMV), virus del mosaico enanizante del maíz (MDMV), A o B, virus del mosaico del rayado del trigo (WSMV), virus del mosaico del pepino (CMV), virus de la estría clorótica del cinodon (CCSV), virus del mosaico de Johnsongrass (JGMV), acrobacia tupida del maíz u organismo similar al micoplasma (NILO), virus del enanismo clorótico del maíz (MCDV), virus del moteado clorótico del maíz (MCMV), cepas del virus del mosaico enanizante del maíz (MDMV) A, D, E y F, virus de la mancha de la hoja del maíz (MLFV), virus de la línea del maíz (NELV), virus del mosaico del maíz (MMV), virus del moteado del maíz y del enanismo clorótico, virus de la mancha anular pelúcida del maíz (MPRV), virus de la raya gruesa del maíz (MRGV), virus del rayado fino del maíz (MRFV), virus de la hoja roja y de la raya roja del maíz (MRSV), virus del moteado del anillo del maíz (MRMV), virus río cuarto del maíz (MRCV), virus del enanismo rugoso del maíz (MRDV), virus del enanismo estéril del maíz (cepas del virus del estriado amarillo de la cebada), virus del rayado del maíz (MSV), rayado clorótico del maíz, hoja del maíz virus del rayado blanco del maíz, virus del enanismo del maíz, virus del aborto

- de la borla del maíz (MTAV), virus de la vena del maíz (MVEV), virus de la oreja del wallaby del maíz (MAVEV), virus de la hoja blanca del maíz, virus del mosaico de la línea blanca del maíz (NTVVLMMV), virus de la hoja roja del mijo (NMV), virus del familia Nanoviridae, Virus del mosaico del cereal del norte (NCMV), virus de la pseudoroseta de la avena, virus del enanismo estéril de la avena (OSDV), virus del enanismo rayado negro del arroz (RBSDV), virus de la franja del arroz (RSV), virus del mosaico del sorgo (SrMV), anteriormente virus del mosaico de la caña de azúcar (SCMV) cepas H, I y M, virus de la enfermedad de Fiji de la caña de azúcar (FDV), virus del mosaico de la caña de azúcar (SCMV) cepas A, B, D, E, SC, BC, virus de la vena Sabi y NM y virus del
- 5 de la mancha del trigo (WSMV).
- Los nematodos parásitos incluyen, pero no se limitan a: *Punzón Dolichodorus spp.*, *D. heterocephalus* Bulbo y tallo (Europa), *Ditylenchus dipsaci* Excavador *Radopholus similis* Quiste *Heterodera avenae*, *H. zaeae*, *Punctodera chalconensis* Daga *Xiphinema spp.*, *X. americanum*, *X. mediterraneum* Falso nudo de raíz *Nacobbus dorsalis* Lanza,
- 15 *Columbia* *Hoplolaimus columbus* Lanza *Hoplolaimus spp.*, *H. galeatus* Lesión *Pratylenchus spp.*, *P. brachyurus*, *P. crenatus*, *P. hexincisus*, *P. neglectus*, *P. penetrans*, *P. scribneri*, *P. thornei*, *P. zaeae* Aguja *Longidorus spp.*, *L. breviannulatus* Anillo *Criconemella spp.*, *C. ornata* Nudo de la raíz *Meloidogyne spp.*, *M. chitwoodi*, *M. incognita*, *M. javanica* Espiral *Helicotylenchus spp.*, *Belonolaimus spp.*, *B. longicaudatus*
- 20 Raíz rechoncha *Paratrichodorus spp.*, *P. christiei*, *P. minor*, *Quinisulcius aculus* y *Trichodorus spp.*
- Las plagas de insectos incluyen insectos seleccionados de los órdenes Coleoptera, Diptera, Hymenoptera, Lepidoptera, Mallophaga, Homoptera, Hemiptera, Orthoptera, Thysanoptera, Dermaptera, Isoptera, Anoplura, Siphonaptera, Trichoptera, etc.,
- 25 particularmente Coleoptera y Lepidoptera.
- En algunas modalidades, el patógeno vegetal se selecciona de hongos, especialmente hongos del suelo tales como *Fusarium oxysporum*, virus del agua y del aire como *Mycosphaerella fijiensis* (Morelet), *Mycosphaerella musicola* (Leach ex Mulder), *Pseudocercospora* (Paracercospora) *fijiensi*, *Verticillium dahliae*, *Cladosporium* y
- 30 *Ralstonia Solanaceum*.
- En algunas modalidades, dicha enfermedad es la marchitez por *Fusarium*, también conocida como mal de Panamá, que es una enfermedad fúngica letal provocada por el hongo *Fusarium oxysporum f. sp. cubense* (Foc) que se transmite por el suelo. Dicha enfermedad también puede ser conocida como mal de Panamá TR4, Foc, mal de
- 35 Panamá Raza Tropical 4 o TR4. En algunas modalidades, la resistencia a TR4 se

combina dentro de un único cultivar con resistencias o tolerancias genéticas a una o más enfermedades adicionales, tales como resistencia a enfermedades provocadas por bacterias, otros hongos, virus, nematodos, insectos y similares.

La marchitez por *Fusarium* es una de las enfermedades más destructivas y notorias del

- 5 plátano. También se le conoce como mal de Panamá, en reconocimiento a los extensos daños que ocasionó en las plantaciones de exportación de este país centroamericano. Para 1960, la marchitez por *Fusarium* había destruido aproximadamente 40 000 ha de 'Gros Michel' (AAA), lo que provocó que la industria exportadora se convirtiera en cultivares del subgrupo Cavendish (AAA) (Ploetz y Pegg, 2000). La marchitez por
- 10 *Fusarium* es provocada por el hifomiceto del suelo, *Fusarium oxysporum* Schlecht. f sp. *cubense*. Es una de las más de 120 *formae speciales* (formas especiales) de *F. oxysporum* que provocan marchiteces vasculares de las plantas con flores. Este patógeno afecta especies de *Musa* y *Heliconia*, y las cepas se han clasificado en cuatro razas fisiológicas basadas en la patogenicidad para hospedar cultivares en el campo
- 15 (raza 1, 'Gros Michel'; raza 2, 'Bluggoe'; raza 3, *Heliconia* spp.; y raza 4, cultivares Cavendish y todos los cultivares susceptibles a las razas 1 y 2). Se han nombrado cuatro razas de *Fusarium oxysporum*, desde la raza 1 hasta la raza 4. La raza 1 es un patógeno crítico de muchos cultivares del plátano. La raza 2 ataca plátanos de cocina. La raza 3 afecta a los parientes del plátano en las Américas, pero no parece afectar al plátano. La
- 20 amenaza actual proviene de la expansión de *Fusarium oxysporum* raza 4, también conocida como TR4 (Raza Tropical 4), que se designa como 'Foc-TR4'. La raza 4 tiene dos subgrupos, TR4 y SR4 (raza subtropical 4). Hasta hace poco, solo se había registrado que la raza 4 provocaba pérdidas graves en las regiones subtropicales de Australia, Sudáfrica, las Islas Canarias y Taiwán. Los productores de plátano y las
- 25 empresas bananeras han declarado repetidamente que si esta raza se estableciera en las Américas, las industrias de exportación mundiales se verían gravemente afectadas, ya que no existe un reemplazo ampliamente aceptado para los cultivares Cavendish (Bentley y otros, 1998).

Muy recientemente (Stokstad, 2019), se detectó la raza 4 del mal de Panamá (marchitez por *Fusarium*) en el hemisferio occidental. La enfermedad se encontró en cuatro plantaciones en Colombia. Estas cuatro plantaciones fueron inmediatamente puestas en cuarentena. Sin embargo, una parte sustancial del mercado del plátano consiste en exportaciones de América Central y del Sur a los Estados Unidos. Este mercado se encuentra ahora en grave peligro, lo que hace aún más urgente una solución rápida a

la crisis. La reciente aparición del mal de Panamá TR4 en el hemisferio occidental hace que una solución rápida a la crisis sea aún más urgente.

En algunas modalidades, "Marchitez por Fusarium" o "FW" pueden usarse indistintamente, lo que designa la enfermedad como se muestra en plantas de plátano infectadas.

En las décadas de 1950 y 1960, una sola variedad, Gros Michel, se cultivó ampliamente. Fue muy sensible al hongo de fácil propagación *Fusarium oxysporum f. sp. cubense*. En particular, fue Fusarium Tropical Raza 1 (Foc-TR1) la que provocó una enfermedad de marchitez fatal, y la industria bananera mundial casi fue destruida. Se descubrió que la variedad Cavendish era altamente resistente a Foc-TR1 y reemplazó a Gros Michel en la producción mundial de plátano. En la década de 1990, los productores comenzaron a encontrar plantas de plátano infectadas con Foc-TR4, una raza emergente. Foc-TR4 también se propaga fácilmente y se encuentra en plantaciones de plátano en Asia, Medio Oriente y África, lo que amenaza una vez más el cultivo mundial de plátano. La reciente identificación de Foc-TR4 en el Caribe ha provocado una gran preocupación, lo que significa que el hongo ahora tiene una cabeza de playa en el hemisferio occidental, y por tanto amenaza la producción bananera de América Latina. En algunas modalidades, la presente descripción proporciona una solución a los problemas graves en los plátanos provocados por Foc-TR4. En algunas modalidades, la solución se dirige a la identificación de materiales genéticos y/o arquitecturas resistentes a enfermedades y la importación de dichos materiales genéticos y arquitecturas a variedades de plátanos que son susceptibles a hongos patógenos (por ejemplo, Foc-TR4).

Los plátanos también son susceptibles a otros hongos patógenos, particularmente *Mycosphaerella fijiensis* (Morelet), que provoca la enfermedad de la estría negra de la hoja (también conocida como Sigatoka Negra) y *M. musicola*, que provoca la enfermedad de la mancha foliar de la Sigatoka Amarilla. Se sabe que estos hongos (*M. fijiensis* y *M. musicola*) se controlan con fungicidas, pero los fungicidas son ineficaces contra Foc-TR4.

La presente descripción enseña un método para modular, estimular o mejorar la resistencia a enfermedades en plantas, provocadas por patógenos tales como Foc-TR4 mediante el uso de técnicas de fitomejoramiento de próxima generación, también conocidas como nuevas técnicas de mejoramiento.

Las nuevas técnicas de mejoramiento (NBT) se refieren a diversas tecnologías nuevas desarrolladas y/o usadas para crear nuevas características en plantas a través de la variación genética, cuyo objetivo es la mutagénesis dirigida, la introducción dirigida de

nuevos genes o el silenciamiento de genes (RdDM). Las siguientes técnicas de mejoramiento están dentro del alcance de las NBT: cambios de secuencia específicos facilitados mediante el uso de la tecnología de nucleasa de dedos de zinc (ZFN) (ZFN-1, ZFN-2 y ZFN-3, ver la patente de los Estados Unidos número 9,145,565, incorporada como referencia en su totalidad), mutagénesis dirigida por oligonucleótidos (ODM, también conocida como mutagénesis dirigida al sitio), cisgénesis e intragénesis, enfoques epigenéticos tales como la metilación del ADN dependiente de ARN (RdDM, que no cambia necesariamente la secuencia de nucleótidos pero puede cambiar la actividad biológica del secuencia), injerto (en rizoma GM), mejoramiento inverso, agroinfiltración para expresión génica transitoria (agroinfiltración "sensu stricto", agroinoculación, inmersión floral), nucleasas efectoras de tipo activador de la transcripción (TALEN, ver las patentes de los Estados Unidos 8,586,363 y 9,181,535, incorporadas como referencia en su totalidad), el sistema CRISPR/Cas (ver las patentes de los Estados Unidos núms. 8,697,359; 8,771,945; 8,795,965; 8,865,406; 8,871,445; 8,889,356; 8,895,308; 8,906,616; 8,932,814; 8,945,839; 8,993,233; y 8,999,641, que se incorporan aquí como referencia), meganucleasa manipuladas genéticamente, endonucleasas autodirigidas remanipuladas genéticamente, edición del genoma guiada por ADN (Gao y otros, Nature Biotechnology (2016), doi: 10.1038/nbt.3547, incorporado como referencia en su totalidad) y genómica sintética. Una parte importante de la edición dirigida del genoma actual, otra designación para las nuevas técnicas de mejoramiento, son las aplicaciones para inducir una ruptura de la doble hebra del ADN (DSB) en una ubicación seleccionada del genoma donde se pretende realizar la modificación. La reparación dirigida del DSB permite la edición específica del genoma. Dichas aplicaciones pueden utilizarse para generar mutaciones (por ejemplo, mutaciones dirigidas o edición precisa de genes nativos) así como también la inserción precisa de genes (por ejemplo, cisgenes, intragenes o transgenes). Las aplicaciones que conducen a mutaciones a menudo se identifican como tecnología de nucleasa dirigida al sitio (SDN), tal como SDN1, SDN2 y SDN3. Para SDN1, el resultado es una mutación por delección genética no específica dirigida: la posición del ADN DSB se selecciona con precisión, pero la reparación del ADN por parte de la célula huésped es aleatoria y da como resultado pequeñas delecciones, adiciones o sustituciones de nucleótidos. Para SDN2, una SDN se usa para generar una DSB dirigida y una plantilla de reparación de ADN (una secuencia de ADN corta idéntica a la secuencia de ADN de la DSB dirigida excepto por uno o algunos cambios de nucleótidos) se usa para reparar la DSB: esto da como resultado una mutación puntual predeterminada en el gen de interés deseado. En

cuanto a SDN3, la SDN se usa junto con una plantilla de reparación de ADN que contiene una nueva secuencia de ADN (por ejemplo, un gen). El resultado de la tecnología sería la integración de esa secuencia de ADN en el genoma de la planta. La aplicación más probable que ilustra el uso de SDN3 sería la inserción de casetes de expresión cisgénicos, intragénicos o transgénicos en una ubicación genómica seleccionada. Una descripción completa de cada una de estas técnicas puede encontrarse en el informe elaborado por el Joint Research Center (JRC) del Institute for Prospective Technological Studies of the European Commission en 2011 y titulado "New plant breeding techniques - State-of-the-art and prospects for commercial development", que se incorpora como referencia en su totalidad.

En algunas modalidades, se han adoptado diversos enfoques para prevenir o tratar la infección por Foc-TR4. La presente descripción enseña que un enfoque clave para prevenir o tratar Foc-TR4 es (1) encontrar cultivares de plátano resistentes, (2) identificar genes de resistencia y/o rasgos de los cultivares de plátano seleccionados, y (3) criar/introducir los genes de resistencia y/o rasgos en cultivares sensibles de plátano. Zuo y otros (2018) evaluaron 129 accesiones de plátano y encontró 10 que son altamente resistentes a Foc-TR4, lo que proporciona cultivares resistentes naturalmente existentes para su estudio.

Li y otros (2012) analizaron los transcriptomas y los perfiles de expresión de las raíces de un mutante resistente y los compararon con los plátanos Cavendish brasileños de tipo silvestre sensibles en dos puntos de tiempo después del desafío con Foc-TR4. Encontraron unos 88 000 unigenes, con 5000 relacionados con vías de defensa en otras plantas. Llegaron a la conclusión de que unos 2600 genes se expresaron de forma diferencial en el mutante resistente, incluidos algunos genes de lignificación de células vegetales que se expresaron en niveles iguales o inferiores en el mutante resistente.

De manera similar, Bai y otros (2013) compararon los transcriptomas de raíces del cultivar brasileño sensible a Foc-TR4 con el cultivar Yueyoukang1, que se sabe que tiene una gravedad mucho menor a la enfermedad. Bai y otros encontraron expresión diferencial en 500 a 2000 unigenes diferentes en diferentes puntos de tiempo, y estos podrían agruparse en 11 tipos diferentes de vías metabólicas. Bai y otros encontraron genes conectados a la lignificación de la pared celular que estaban regulados diferencialmente entre los cultivares sensibles y resistentes, específicamente 4-coumarato: CoA ligasa (4CL), glutatión S-transferasa (GST), celulosa sintasa, Caffeoil-CoA O-metiltransferasa (CCoAM), y cinamilalcohol deshidrogenasa (CAD) se expresaron en niveles más altos en el cultivar resistente y concluyeron que la

- lignificación de la pared celular podría ser uno de los mecanismos implicados en la resistencia a Foc-TR4. Bai y otros señalaron que esto era incompatible con el resultado encontrado en Li y otros (2012) y concluyeron que diferentes plantas podrían tener diferentes mecanismos de resistencia y que se requiere más trabajo para descifrar cómo los cultivares de plátano pueden resistir Foc-TR4.
- Wang y otros (2017) también analizaron la expresión genética diferencial de las raíces en el momento de la diferenciación de los botones florales y encontraron 107 genes expresados diferencialmente en las raíces entre un cultivar de plátano susceptible y uno sensible.
- Zhang y otros (2018) mostraron que la infección por Foc-TR4 procede de manera similar en las raíces de un cultivar resistente (Pahang) y un cultivar susceptible (Brasileño) hasta llegar al cormo, donde la biomasa fúngica y el grado de necrosis fueron significativamente menores en Pahang. (El 'cormo' de plátano es un tallo subterráneo, o rizoma, del que crecen las raíces).
- Van der Berg y otros (2007) usaron RT-PCR cuantitativa para identificar genes que estaban regulados positivamente en el cultivar de plátano GCTCV-218 tolerante a FW después de la infección con Foc-4. Su control fue el cultivar Williams sensible a FW. Descubrieron que varios genes estaban regulados positivamente en GCTCV-218 tolerante a FW en comparación con Williams sensible a FW. Como era de esperar, muchos de los genes regulados positivamente eran homólogos a genes conocidos asociados a la defensa, incluidos los genes que fortalecen la pared celular. Informaron 13 genes que estaban regulados positivamente en las raíces. Si bien afirman que "Los resultados arrojan luz sobre los genes involucrados en la defensa y proporcionan un paso hacia la comprensión de la marchitez por *Fusarium* del plátano y, de esta manera, el desarrollo de una estrategia de manejo de enfermedades eficaz", el documento no sugiere que ninguno de los 13 genes que depositaron en GenBank pueda usarse para controlar la marchitez por *Fusarium*. No se proporciona una estrategia particular para el uso de estos genes para controlar la marchitez por *Fusarium*.
- Vishnevetsky y otros (2009) (Patente de los Estados Unidos núm. 7,534,930) describen un método para manipular genéticamente plantas de plátano para conferir rasgos de resistencia a enfermedades exógenas, lo que incluye la resistencia a Sigatoka Negra y Amarilla y *Botrytis cinerea*. Vishnevetsky y otros manipularon tres polinucleótidos en plantas de plátano, incluidos genes que codifican endoquitinasa, estilbeno sintasa y superóxido dismutasa.

Paul y otros (2011) aislaron un gen del nematodo *C. elegans* que, cuando se transformó de manera estable en el cultivar de plátano 'Lady finger', confirió en ensayos de invernadero resistencia a la raza 1 del mal de Panamá.

Aunque es poco probable que los consumidores acepten la transformación de plátanos con un gen derivado de un nemátodo, el trabajo de seguimiento realizado por el grupo de Dale con un gen derivado de plátanos es prometedor para lograr la resistencia a *Fusarium* en plátanos transformados con OGM. Por ejemplo, Peraza-Echeverría y otros (2009) aislaron un gen análogo del gen de resistencia (RGA2) de un plátano silvestre, *Musa acuminata malaccensis*. Este gen es miembro de la gran familia de genes de resistencia de tipo NB-LRR. Cuando se transforma en plantas Cavendish sensibles a FW (Dale y otros, 2017), el gen parece conferir resistencia a *Fusarium*. Dale y otros (2017) realizaron pruebas de campo en plantas de plátano transgénicas durante 3 años. Al concluir el ensayo, entre el 67 % y el 100 % de las plantas de control sensibles a FW estaban muertas o infectadas. Sin embargo, en cuatro líneas de plátanos transfectados con su gen candidato, menos del 30 % de los plátanos transformados mostraron signos de infección grave (es decir, ≥ 70 % mostró alguna tolerancia o resistencia). Una línea transformada con RGA2 pareció ser inmune a TR4. Si bien esta es una buena evidencia de que el gen puede conferir cierta resistencia a FW, el gen se aisló por primera vez hace más de una década y no está claro si la industria bananera alguna vez adoptará el gen RGA2.

Es importante señalar que se cree (comunicaciones no publicadas con los mejoradores y científicos de la industria bananera) que puede haber hasta cuatro genes en el genoma de *Musa* que contribuyan con algún grado de resistencia a *Fusarium*, por lo que es poco probable que RGA2 por sí solo resuelva la crisis actual, incluso si es aceptado por los productores. Incluso si RGA2 encuentra aceptación, la industria tiene una gran necesidad de múltiples genes para controlar TR4.

El inventor aclara que FusR1 de la presente descripción no está relacionado en absoluto con RGA2. Los dos genes tienen secuencias de nucleótidos completamente diferentes (es decir, no tienen identidad de secuencia), se encuentran en diferentes cromosomas, tienen diferentes bioquímicas y tienen diferentes mecanismos de acción en la planta.

Wu y otros (2016) secuenciaron un pariente del plátano silvestre resistente a enfermedades, *Musa itinerans*, que se encuentra en la China subtropical. Los valores de Ks se calcularon para estimar los eventos de especiación y paleoploidización en el género *Musa*. También se calcularon los valores de Ka/Ks para mostrar que, como se esperaba, la mayoría de los genes en el genoma de *Musa itinerans* se han sometido a

una selección purificadora. Se sugirió que es conocido que *M. itinerans* es resistente a enfermedades, por lo que su genoma podría extraerse en busca de genes de resistencia a enfermedades.

En algunas modalidades, la presente descripción proporciona métodos para encontrar, identificar y seleccionar genes resistentes a enfermedades, tales como la marchitez por *Fusarium* de cultivares de plátano resistentes a FW. En otras modalidades, la presente descripción proporciona secuencias de nucleótidos y polipéptidos de genes resistentes a *Fusarium* (por ejemplo, *gen FusR1*) identificadas a partir de los métodos de la presente descripción. En modalidades adicionales, la presente descripción enseña métodos para generar y/o producir variedades de plátano que tienen genes y/o rasgos de resistencia mediante el uso de tecnología de fitomejoramiento de nueva generación, que incluye pero no se limita a la tecnología CRISPR descrita en la presente descripción.

III. Identificación del gen *FusR1* del género *Musa*

Los plátanos cultivados son generalmente triploides (aunque algunos son diploides) como resultado de su compleja historia evolutiva y de domesticación, que involucró una serie de eventos de hibridación interespecíficos e intraespecíficos, tanto naturales como hechos por humanos. Los plátanos comestibles cultivados son en gran parte el resultado de la hibridación entre dos especies diploides silvestres, *Musa acuminata* y *Musa balbisiana* (Christelová y otros, 2017). La domesticación humana de los plátanos comenzó hace unos 7000 años en el sudeste asiático (D'Hont y otros, 2012). Los genomas de plátano derivados de *M. acuminata* se conocen como genomas "A", mientras que los plátanos derivados de *M. balbisiana* son genomas "B" (D'Hont y otros, 2012). Por tanto, la estructura del genoma del diploide *M. acuminata* se denomina AA y la estructura genómica del diploide *M. balbisiana* es BB. Por lo tanto, los cultivares de plátano comestible pueden tener genomas triploides AAA (como Cavendish o Gros Michel), genomas AAB (como en muchos plátanos) o genomas ABB (como la variedad local Cachaco). *M. acuminata* probablemente surgió en Malasia o Indonesia (Christelová y otros, 2017). Por el contrario, se cree que *M. balbisiana* se originó en India, Tailandia o Filipinas (Christelová y otros, 2017). Por lo tanto, estas dos especies eran originalmente alopátricas y el aislamiento geográfico proporcionó una oportunidad para que cada especie desarrollara rasgos únicos. Cuando, posteriormente, los humanos trasladaron los cultivares de *M. acuminata* a áreas pobladas por *M. balbisiana*, tuvo lugar la hibridación interespecífica.

El cultivar Cavendish, económicamente crítico, que representa al menos el 99 % de la producción comercial de exportación de plátano, exhibe esterilidad inducida por

triploides. Esto, combinado con la partenocarpia, da lugar a frutos comestibles sin semillas, pero dificulta gravemente el mejoramiento, por lo que los plátanos Cavendish se propagan vegetativamente (clonalmente). El genotipo Cavendish tiene tres genomas "A" derivados de *M. acuminata*.

- 5 En algunas modalidades, el inventor identificó genes que controlan eficazmente la marchitez por Fusarium en el plátano. Por ejemplo, la presente descripción enseña que un gen, que se nombra *FusR1* (Fusarium Resistencia 1) se identificó mediante el uso de enfoque de análisis de la evolución molecular del inventor. La presente descripción enseña que el *gen FusR1* es un gen nativo en las especies de *Musa*, incluidos los
- 10 plátanos cultivados, *M. itinerans*, *M. acuminata*, *M. balbisiana*, *M. basjoo*, así como también *Musella lasiocarpa*, el único miembro de un género estrechamente relacionado. El ortólogo (dos alelos) del pariente del plátano silvestre, *Musa itinerans*, se proporciona aquí como SEQ ID NO: 1 y SEQ ID NO: 4. Las secuencias de *M. itinerans FusR1* se obtuvieron de múltiples accesiones (que incluyen, pero no se limitan a, ITC1526,
- 15 ITC1571 y PT-BA-00223). Todas las accesiones de *M. itinerans* son extremadamente resistentes a FW (Li y otros, 2015; Wu y otros, 2016).

La presente descripción enseña que el inventor identificó dos alelos de *FusR1* en *M. itinerans*. SEQ ID NO: 1 da el alelo #1 de la secuencia de ARNm de *FusR1*. SEQ ID NO: 2 da la secuencia codificante del alelo #1. SEQ ID NO: 4 da el alelo #2 de la secuencia de ARNm de *FusR1*. SEQ ID NO: 5 da la secuencia codificante del alelo #2. Los alelos 1 y 2 son muy similares en secuencia: codifican solo cuatro diferencias de aminoácidos. Se identificó una segunda transcripción de *FusR1* (SEQ ID NO: 7) de *M. itinerans*; esta transcripción tiene un intrón expresado (es decir, sin empalmar) que da como resultado la interrupción del marco de lectura adecuado. Esto se expresa en niveles muy bajos.

- 25 *M. itinerans* es de forma natural extremadamente resistente a los efectos de la marchitez por Fusarium (Li y otros, 2015; Wu y otros, 2016). En algunas modalidades, el gen *FusR1* de *M. itinerans* es responsable de la resistencia a la marchitez por Fusarium.

La presente descripción enseña además que el inventor identificó tres alelos de *FusR1* en *M. acuminata*. Dos de estos alelos se aislaron de accesiones resistentes a FW de *M.*

- 30 *acuminata*. El tercer alelo se aisló de una accesión de *M. acuminata* sensible a FW. Las secuencias resistentes a FW de *M. acuminata FusR1* se obtuvieron de múltiples accesiones resistentes a FW, incluidas ITC0896 (*M. a.* subespecie *banksii*) y PT_BA-00281 (Pisang Bangkahulu). La secuencia sensible a FW de *M. acuminata* procede de las accesiones sensibles a FW ITC0507, ITC0685, PT-BA-00304, PT-BA-00310 y PT-
- 35 BA-00315.

SEQ ID NO: 8 da la secuencia de ARNm del alelo 1 del gen *FusR1* resistente a FW de *M. acuminata*. SEQ ID NO: 10 da la secuencia de ARNm del alelo 2 del gen *FusR1* resistente a FW de *M. acuminata*. La secuencia codificante del alelo 1 resistente a FW de *M. acuminata* se proporciona en SEQ ID NO: 9. SEQ ID NO: 11 da la secuencia

5 codificante del alelo 2 resistente a FW de *M. acuminata*.

SEQ ID NO: 13 da la secuencia de ARNm del alelo *FusR1* sensible a FW de *M. acuminata*. (La secuencia sensible a FW de *M. acuminata* se identificó a partir de las accesiones ITC0507, ITC0685, PT-BA-00304, PT-BA-00310 y PT-BA-00315. Estas

10 accesiones incluyen múltiples muestras de cultivares de plátano tales como Pisang Madu, Pisang Pipit y Pisang Rojo Uter, todos los cuales se han caracterizado bien como sensibles a FW (Chen y otros, 2019).

El inventor identificó un promotor central putativo para *FusR1* de *M. acuminata*. El inventor usó dos aplicaciones de predicción de promotores diferentes en un intento de encontrar predicciones congruentes de diferentes algoritmos/software.

15 Como primera etapa, el inventor amplificó y secuenció un fragmento de secuencia de 753 pb (SEQ ID NO: 31), que comienza aguas arriba de la región codificante del alelo resistente a FW del gen *FusR1* derivado de *M. acuminata*. Este fragmento es 100 % idéntico a pb7868911 - pb7869210 y bp7869341 - bp7869743 del número de accesión de GenBank NC_025206 (*Musa acuminata* subsp. *Malaccensis* cromosoma 5,

20 ASM31385v2, secuencia de genoma completa), que se encuentra en el cromosoma 5 de *M. acuminata*.

El inventor analizó primero la región aguas arriba de *FusR1* mediante el uso de la "Neural Network Promoter Prediction" (NNPP), que está disponible en el Berkeley *Drosophila* Genome Project (BDGP). BDGP es un consorcio del Centro del Genoma de

25 *Drosophila*, financiado por el Instituto Nacional de Investigación del Genoma Humano, el Instituto Nacional del Cáncer y el Instituto Médico Howard Hughes. El software NNPP se "entrenó" en secuencias promotoras humanas y de *Drosophila melanogaster*, pero ha demostrado ser generalmente eficaz para identificar secuencias promotoras, incluso en plantas (Reese, 2001).

30 El análisis de NNPP identificó con éxito un promotor central para *FusR1*. A continuación se muestran los resultados del análisis. Las primeras 189 bases de la SEQ ID NO: 31 (mostradas en minúsculas) son secuencias aguas arriba no codificantes, que incluyen la secuencia UTR 5' de *FusR1*; las siguientes 423 bases son secuencias codificantes (mostradas en MAYÚSCULAS). Esta secuencia de codificación es idéntica a la SEQ ID

35 NO: 9. Las últimas 141 bases son UTR 3' (mostradas en minúsculas). Las bases 92-141

de SEQ ID NO: 31 (**atcgtggcactataaataggacaagaggagggatgaggtaaacgcactc**) son la secuencia promotora predicha por NNPP, mostrada en **negrita y minúscula**. El sitio de inicio de la transcripción (TSS) en el par de bases 132 se muestra en **minúsculas y subrayado en negrita**. NNPP asigna una puntuación de .88 (es decir, un nivel de confianza del 88 %) a este promotor.

SEQ ID NO: 31:

gtagagacacttgagttgaattctgaatcattattcttctcatgaacgcatacgtcccaccatacacaccaaatcttaattg
ctcaagc**catcgtggcactataaataggacaagaggaggatgaggtaaacgcactc**cctcatacttgcacaggt
acgtgttgtagaagaattcagaggtaaagcgATGGCTGGAGGAGGCAAAAGAGGTGAAGCGTCG
TCTCTTCTACTTGTGACGCTGCTCGTGACGTTGTTGGCTTTCTTCGCCACCAACTC
CTCGGCAGCCCGTGTACACCCCGTCCGCAATCCCTCGCCAGAGCGGCACTGAG
TGCGGTGGGGGCAAGGCAAGATGAGCCGTGCTGCAGATGCGCGTGTCTCTCAT
TTACCCACCTACTTGGTGCAATTGCGCGGCATATGGCAAGGCTCCTGCCCTTCC
GCCTGCAACAACCTGCCAGTGTGTCTCAACGAGTGCACCTTGCCCTCGATCTTATGG
ACCCCAAGGTCTGCGAGGCCAACTCCTGTCCCTGGCCTGTTGCAGCCCCCAAAG
TAGAGCCGGCGCAGCAGTGGGCTATCGAAGAAACCGGTGGGAAATTAGCGATGA
TGGTGTGATcccaattgtgtttgtagcgtcgtcgtcttctctcgtccgtccatccatctatccatccatctactataatct
atgtcgtgtaccgctcgtgtgggtgtgcttctcctcagtaataaaaaataaaatgcttctcgtttt

El inventor después analizó la región aguas arriba de *FusR1* de *M. acuminata* mediante el uso del software "Prediction of PLANT Promoters" (TSSP), que está dirigido específicamente a la identificación de secuencias promotoras de plantas (Solovveyev y Shahmuradov, 2003). Esto es parte de un conjunto de software de análisis de secuencias producido por Softberry, Inc. TSSP identificó el sitio de inicio de la transcripción (TSS) como la posición 132 en la SEQ ID NO: 31, que es idéntico a los resultados del software NNPP (ver arriba). TSSP ubicó la caja TATA de *FusR1* (mostrada arriba en ***cursiva minúscula***) en las bases 102-107 de la SEQ ID NO: 31. Por tanto, la caja TATA de *FusR1* se encuentra, como se esperaba, 25 pares de bases aguas arriba del TSS.

Como estas 2 aplicaciones de predicción de promotores diferentes dan resultados congruentes, el inventor identificó la secuencia de promotor correcta para *M. acuminata*. La presente descripción enseña métodos para introducir el gen *FusR1* recién identificado y sus variantes en plátanos cultivados, particularmente el cultivo Cavendish sensible a *Fusarium* con el fin de hacer que estos cultivos sean resistentes a la marchitez por *Fusarium*. En algunas modalidades, la presente descripción enseña que pueden usarse métodos tradicionales de fitomejoramiento para introducir el gen/rasgo *FusR1* de

M. itinerans en Cavendish y otros plátanos cultivados. En otras modalidades, la presente descripción enseña que pueden usarse métodos de fitomejoramiento de próxima generación para introducir el gen/rasgo *FusR1* de *M. itinerans* en Cavendish y otros plátanos cultivados. En modalidades adicionales, la presente descripción enseña

5 métodos para introducir el gen/rasgo *FusR1* de *M. itinerans* en Cavendish y otros plátanos cultivados mediante el uso de técnicas de edición del genoma como el sistema de edición del genoma dirigido mediante el uso de nucleasas de dedos de zinc (ZFN), activador de la transcripción como nucleasas efectoras (TALEN) o tecnología del sistema CRISPR/Cas9 que explota la actividad endonucleasa de proteínas asociadas a

10 CRISPR (Cas) con especificidad de secuencia dirigida por ARN CRISPR (ARNcr). Dada la amenaza de extinción probable para Cavendish, la presente descripción proporciona un enfoque de edición del genoma rápido, eficiente y preciso mediante el uso del sistema CRISPR/Cas9 adaptado para la producción de plátanos mínimamente editados genéticamente que tienen un gen/rasgo resistente a *Fusarium*, que será

15 aceptado especialmente en los países en desarrollo donde el plátano proporciona una seguridad económica y alimentaria fundamental. La presente descripción enseña que la transferencia del gen *FusR1* nativo de *M. itinerans* a plátanos cultivados puede lograrse mejor con la tecnología CRISPR, que permite una transferencia dirigida, limpia y eficiente y que, en comparación con las técnicas de edición genética más tradicionales,

20 minimiza posibles efectos secundarios. En algunas realizaciones, los alelos útiles de *FusR1* (SEQ ID NO: 8 y SEQ ID NO: 10) se identifican a partir de poblaciones de *M. acuminata* naturalmente resistentes a FW. Estos alelos confieren resistencia a FW. La presente descripción enseña que el alelo *FusR1* derivado de *M. acuminata* puede usarse, en combinación con alelos *FusR1*

25 derivados de *M. itinerans* (SEQ ID NO: 2 y SEQ ID NO: 5), para mejorar la resistencia a FW en plátanos cultivados, particularmente Cavendish. La presente descripción enseña el apilamiento de genes con al menos dos genes *FusR1* identificados por el inventor descrito en la presente descripción.

30 Tanto el ortólogo de *M. itinerans* *FusR1* (SEQ ID NO: 2 y SEQ ID NO: 5) como los alelos resistentes a FW de *M. acuminata* (SEQ ID NO: 8 y SEQ ID NO: 10) pueden usarse en el fitomejoramiento tradicional y/o enfoques de fitomejoramiento de nueva generación. Los enfoques de fitomejoramiento de nueva generación incluyen, entre otros, la selección asistida por marcadores (MAS) y/o técnicas de edición del genoma en plátanos cultivados.

Algunas accesiones de *M. balbisiana* se han caracterizado rigurosamente como muy resistentes a FW, mientras que otras son extremadamente sensibles a FW. Si bien podría esperarse que las accesiones silvestres de *M. balbisiana* fueran resistentes a un patógeno como *Fusarium*, ha sido difícil para los investigadores comprender por qué las accesiones estrechamente relacionadas difieren de manera tan significativa en términos de resistencia a FW.

El inventor descubrió una diferencia estructural de las secuencias de nucleótidos del gen *FusR1* en accesiones de *M. balbisiana* sensibles a FW como se muestra en la Figura 5. Todas las accesiones de *M. balbisiana* sensibles a FW analizadas por el inventor contienen una transcripción de *FusR1* "rota". Este análisis se limita a los genes *FusR1* "rotos" que se encuentran en todas las accesiones sensibles a FW que se examinaron. Los ARNm de *FusR1* en todas las accesiones de *M. balbisiana* que el inventor examinó tenían un intrón expresado sin empalmar que interrumpe el marco de lectura adecuado. Además, el inventor encontró (i) una delección larga de 82 u 84 pb en varios ARNm de *FusR1* (2) en todas las accesiones, una delección más pequeña de 1 pb, o (ii), en algunas accesiones, una inserción de 4 pb, cada una de las cuales también interrumpe el marco de lectura abierto, lo que codifica para una proteína FUSR1 mutada, no funcional. Todas las accesiones de *M. balbisiana* sensibles a FW tienen uno o más de estos disruptores del marco de lectura descritos anteriormente, lo que da como resultado una proteína no funcional. En algunas modalidades, la presente descripción enseña que algunas accesiones de *M. balbisiana* tienen los cuatro disruptores del marco de lectura. Ver la Figura 5.

En otras modalidades, el inventor también descubrió otra diferencia significativa al estudiar accesiones de *M. acuminata* resistentes a FW frente a sensibles a FW. En algunas modalidades, *FusR1* en *M. acuminata* da resistencia frente a la sensibilidad en dependencia de los alelos *FusR1*. La presente descripción enseña que dos alelos, que resultaron ser "alelos resistentes", confieren resistencia a FW; SEQ ID NO: 8 y SEQ ID NO: 10. Estos dos alelos son muy similares en secuencia al ortólogo *FusR1* derivado de los espacios de plátano silvestre resistentes a FW, *M. basjoo* (SEQ ID NO: 17 y SEQ ID NO: 20). El tercer alelo, el alelo sensible a FW, se encuentra solo en accesiones de *M. acuminata* sensibles a FW (SEQ ID NO: 13).

La secuencia de *FusR1* de *M. balbisiana* (SEQ ID NO: 26 y SEQ ID NO: 27) no confiere resistencia a FW, porque este gen está dañado (como lo está en todas las accesiones de *M. balbisiana* sensibles a FW examinadas) por indeles que interrumpen el marco de

lectura y/o intrones expresados sin empalmar que provocan pérdida de resistencia a FW.

En modalidades adicionales, las secuencias de *FusR1* derivadas de accesiones de *M. acuminata* resistentes a FW (SEQ ID NO: 8 y SEQ ID NO: 10) tienen una similitud de secuencia muy alta con el ortólogo *FusR1* derivado de *M. basjoo* (SEQ ID NO: 17). *M. basjoo* es una especie de plátano silvestre muy resistente a FW (Li y otros, 2015). En otras modalidades, la secuencia *FusR1* (SEQ ID NO: 13) de accesiones de *M. acuminata* sensibles a FW difiere de los alelos de *M. acuminata* resistentes a FW (SEQ ID NO: 8 y SEQ ID NO: 10).

La presente descripción enseña que la resistencia a FW en *M. acuminata* depende de que el alelo se encuentre sólo en accesiones resistentes a FW. Aunque *M. acuminata* y *M. balbisiana* están más estrechamente relacionados entre sí que con *M. itinerans* o *M. basjoo*, las secuencias de *FusR1* que controlan la resistencia a FW se agrupan en contraste directo con la forma en que las especies están realmente relacionadas. En otras modalidades, el gen *FusR1* se ha adaptado (es decir, se ha seleccionado positivamente) de modo que *FusR1* no refleja las relaciones reales dentro de la especie *Musa*. La presente descripción enseña dos eventos adaptativos independientes (evolución convergente) o quizás la versión de *FusR1* resistente a FW se ha comercializado entre diversas especies de *Musa* (transferencia de genes).

El inventor confirmó las verdaderas relaciones filogenéticas entre estas especies de *Musa* mediante la secuenciación de dos genes de copia única, conservados y diferentes, *C2H2* y *TOPO6*, de varias especies de *Musa*. Las proteínas de dedos de zinc de tipo C2H2 desempeñan un papel importante en el desarrollo y el crecimiento de las plantas, así como también en la resistencia al estrés abiótico, incluida la maduración de la fruta en el plátano (Han y otros, Front. Plant Sci., Vol. 11, artículo 115:1-13, 20 de febrero de 2020; Han y otros, Postharvest Biology and Technology, 116:8-15, junio de 2016). *TOPO6*, una región marcadora de genes nucleares de la subunidad B del homólogo vegetal de la topoisomerasa VI arquea, se produce como un locus de copia única en el genoma haploide de la mayoría de los grupos de plantas (Frank R. Blattner, Plant Systematics and Evolution, Vol. 302:239 -244, 2016). Estos dos genes (cuyas funciones bioquímicas son bien conocidas) no tienen ningún papel en el control de patógenos, lo que los hace ideales como 'controles' para comprender los cambios adaptativos impuestos al plátano *FusR1* como resultado de la exposición a *Fusarium*. Por tanto, la descripción enseña que el consenso en la literatura de que *M. acuminata* y *M. balbisiana* son especies hermanas es correcto, lo que significa que se han producido cambios

significativos en nuestro gen recientemente identificado, *FusR1*, en estas especies de plátano, lo que proporciona aún más evidencia de que *FusR1* confiere resistencia a FW. Ver los árboles filogenéticos proporcionados en las Figuras 3 y 4.

La presente descripción enseña las diferencias de secuencia críticas entre los alelos *FusR1* fuertemente resistentes a FW de *M. itinerans*, lo que permite al inventor determinar los pocos nucleótidos exactos que hacen que *FusR1* sea capaz de controlar FW. Basado en los hallazgos del inventor, la presente descripción enseña un método de uso del sistema CRISPR/Cas para conferir resistencia a FW en Cavendish sensible a FW (así como también en todos los demás plátanos cultivados), mediante el cambio precisamente de solo unos pocos nucleótidos críticos en *FusR1*. Además, la presente descripción también enseña un método para usar estos nucleótidos críticos para crear una nueva secuencia de *FusR1* con mayor resistencia a FW que el gen nativo.

IV. Gen *FusR1* y variantes del mismo

La presente descripción se basa, en parte, en el aislamiento del nuevo gen *FusR1* de variedades y especies de plátano. Las secuencias de nucleótidos de este gen *FusR1* y sus secuencias de ortólogos se presentan en SEQ ID NO: 1-2, 4-5, 7-11, 13-14, 16-18, 20-21, 23-24 y 26-31, respectivamente.

En algunas modalidades, la SEQ ID NO: 1 es una secuencia de ARNm parcial para el alelo 1 de *FusR1* de *Musa itinerans*, la especie de plátano silvestre más resistente a *Fusarium*. La SEQ ID NO: 4 es una secuencia de ARNm parcial para el alelo 2 de *FusR1* de *Musa itinerans*.

Los alelos *FusR1* antes mencionados de *M. itinerans* (SEQ ID NO: 1 y SEQ ID NO: 4) codifican para proteínas ligeramente diferentes, que son las SEQ ID NO: 3 y SEQ ID NO: 6, respectivamente. El polipéptido traducido de SEQ ID NO: 1 se presenta como SEQ ID NO: 3. El polipéptido traducido de SEQ ID NO: 4 se presenta como SEQ ID NO: 6. Estos son solo ligeramente diferentes, y los pocos residuos de aminoácidos diferentes son todos bioquímicamente conservadores. En algunas modalidades, se secuenciaron 5 accesiones diferentes de *M. itinerans* y todas las accesiones tenían estos mismos dos alelos *FusR1*.

En algunas modalidades, las SEQ ID NO: 8 y SEQ ID NO: 10 son ARNm parciales (lo que incluye las secuencias codificantes completas). Estos son los alelos resistentes a FW de *FusR1* de *Musa acuminata* ssp. *banksia* (número de accesión ITC0896) y PT_BA-00281 (Pisang Bankahulu). Estos dos alelos difieren en un solo sitio silencioso.

En otras modalidades, SEQ ID NO: 13 representa el alelo sensible a FW de *M. acuminata*. En modalidades adicionales, SEQ ID NO: 9 y SEQ ID NO: 11 representan la

secuencia codificante para los alelos resistentes a FW de *M. acuminata*. Además, SEQ ID NO: 12 representa la secuencia de proteína resistente a FW de *M. acuminata*, que es una secuencia polipeptídica traducida de SEQ ID NO: 8 y SEQ ID NO: 10.

En algunas modalidades, las SEQ ID NO: 17 y SEQ ID NO: 20 son secuencias de alelos parciales de ARNm *FusR1* de *M. basjoo*, una especie de plátano silvestre que es resistente a *Fusarium*. En otras modalidades, SEQ ID NO: 23 es la secuencia *FusR1* de otro pariente de plátano silvestre, *Musella lasiocarpa*.

Se observa que todas las secuencias de ARNm que los inventores informan en la presente descripción son técnicamente parciales, ya que carecen de un poco de UTR 5' y normalmente algunas bases del extremo de UTR 3'. La gran mayoría de los ARNm descritos en la presente descripción están muy cerca de ser de secuencia completa.

En algunas modalidades, las SEQ ID NO: 26 y SEQ ID NO: 28-30 son las secuencias de ARNm parcial *FusR1* de varias accesiones diferentes de *M. balbisiana*. La SEQ ID NO: 27 es la secuencia codificante de *FusR1* de *M. balbisiana*. En algunas modalidades,

se examinó un gran número de accesiones de *M. balbisiana* sensibles a FW. En todas las secuencias de *FusR1* de accesiones de *M. balbisiana* sensibles a FW, la estructura de la secuencia de *FusR1* está rota y/o dañada. Todas las accesiones de *M. balbisiana* sensibles a FW tenían una secuencia codificante de *FusR1* con una delección de 1 pb en la posición 340 de la secuencia codificante. Todas las accesiones de *M. balbisiana* sensibles a FW también tenían un intrón largo expresado sin empalmar en la secuencia codificante. Varios también tenían una delección larga (82-84 pb), algunos tenían otra delección de 4 pb y, en todos los casos, una delección de un par de bases (en relación con *FusR1* de otras especies de plantas, incluidas todas las demás accesiones de plátano. Si bien es cierto que 84 pb, como múltiplo de tres, no interrumpe el marco de lectura, elimina 28 residuos de aminoácidos de la estructura primaria de la proteína, lo que potencialmente interrumpe la estructura terciaria de la proteína plegada y, por lo tanto, afecta negativamente la función. En cualquier caso, basado en nuestros hallazgos, la ubicua delección de 1 pb siempre da como resultado una interrupción del marco de lectura.

El inventor incluyó secuencias de ARNm de accesiones de *Musa balbisiana* a partir de las cuales el inventor secuenció *FusR1*. Estos ilustran las diversas formas en las que *FusR1* se 'rompe' en *M. balbisiana*. El inventor señala en la presente descripción que TODAS las accesiones de *M. balbisiana* que el inventor analizó tienen una transcripción de ARNm de *FusR1* rota. La Figura 5 muestra estas secuencias de *M. balbisiana FusR1* alineadas.

- La accesión ITC1016 de *M. balbisiana* (SEQ ID NO: 26) contiene un intrón expresado sin empalmar de 82 pares de bases. Este intrón interrumpe el marco de lectura, lo que da como resultado un codón de terminación prematuro ubicado a 8 pb en el intrón, lo que provoca una secuencia codificante truncada de 141 pb (en oposición a la secuencia codificante de 423 pb adecuada). Además, esta accesión (y, de hecho, todas las accesiones de *M. balbisiana*) también tiene una delección de un par de bases, ubicada aproximadamente a 90 pb en 5' del codón de terminación verdadero, que (incluso si el intrón se hubiera empalmado correctamente) da como resultado un codón de parada prematuro, dando una secuencia codificante truncada.
- 5 La accesión ITC0545 de *M. balbisiana* (SEQ ID NO: 28) contiene el mismo intrón expresado, sin empalmar, de 82 pares de bases. Este intrón interrumpe el marco de lectura, lo que da como resultado un codón de parada prematuro ubicado a 8 pb en el intrón, lo que provoca una secuencia codificante de 141 pb truncada (en oposición a la secuencia codificante de 423 pb adecuada). Otros 27 pb aguas abajo del intrón
- 15 expresado se encuentra una delección de 85 pb. Si bien esto en combinación con el intrón expresado de 84 pb restauraría matemáticamente el marco de lectura correcto (85 pb -82 pb = 3 pb), como se explicó anteriormente, provoca la pérdida de 28 residuos de aminoácidos que se encuentran en una región funcionalmente crítica de la proteína FusR1 plegada. Además, esta accesión también tiene la delección de un par de bases,
- 20 ubicada aproximadamente a 90 pb en 5' del codón de terminación verdadero, que (incluso si el intrón se hubiera cortado y empalmado correctamente) da como resultado un codón de terminación prematuro, lo que da una secuencia codificante truncada. Finalmente, el ARNm de *FusR1* de esta accesión también tiene una inserción de 4 pb que interrumpe el marco aguas abajo.
- 25 La accesión ITC0080 de *M. balbisiana* (SEQ ID NO: 29) contiene el mismo intrón expresado sin empalmar que las accesiones anteriores, excepto que esta versión del intrón sin empalmar tiene una longitud de 84 pb. Si bien este intrón expresado no interrumpe el marco de lectura, introduce 28 residuos de aminoácidos adicionales que se encuentran en una región funcionalmente crítica de la proteína plegada y, por tanto,
- 30 es muy probable que evite el plegamiento adecuado de la proteína FusR1. Además, esta accesión también tiene la delección de un par de bases, ubicada aproximadamente a 90 pb en 5' del codón de terminación verdadero, que (incluso si el intrón se hubiera cortado y empalmado correctamente) da como resultado un codón de terminación prematuro, lo que da una secuencia codificante truncada.

La accesión ITC1527 de *M. balbisiana* (SEQ ID NO: 30) contiene el mismo intrón expresado sin empalmar que las accesiones anteriores, esta vez de 82 pb de longitud. De nuevo, este intrón interrumpe el marco de lectura, lo que da como resultado un codón de parada prematuro ubicado a 8 pb en el intrón, lo que provoca una secuencia codificante truncada de 141 pb (en oposición a la secuencia codificante adecuada de 423 pb). Además, el ARNm de *FusR1* de esta accesión tiene una inserción de 4 pb más aguas abajo. Además, esta accesión también tiene la delección de un par de bases, ubicada aproximadamente a 90 pb en 5' del codón de terminación verdadero, que (incluso si el intrón se hubiera cortado y empalmado correctamente) da como resultado un codón de terminación prematuro, lo que da una secuencia codificante truncada. Todas las accesiones de *M. balbisiana* analizadas por el inventor tienen alguna combinación de uno o más de estos diversos defectos en su ARNm de *FusR1*. La Tabla 1 resume la información de secuencia de la presente descripción.

Tabla 1. Resumen de la información de la secuencia

SEQ ID NO.	Tipo de secuencia	Origen	Descripción breve
1	Nucleótido	<i>Musa itinerans</i>	Secuencia de ARNm parcial para el transcrito 1 de <i>FusR1</i> resistente a FW* , alelo 1 de <i>Musa itinerans</i>
2	Nucleótido	<i>Musa itinerans</i>	Secuencia codificante resistente a FW del alelo 1 de <i>FusR1</i> de <i>M. itinerans</i>
3	Proteína	<i>Musa itinerans</i>	Secuencia de proteínas de <i>FUSR1</i> resistente a FW alelo 1 de <i>M. itinerans</i>
4	Nucleótido	<i>Musa itinerans</i>	Secuencia de ARNm parcial para el transcrito 1 de <i>FusR1</i> resistente a FW alelo 2 de <i>Musa itinerans</i>
5	Nucleótido	<i>Musa itinerans</i>	Secuencia codificante del alelo 2 resistente a FW de <i>FusR1</i> de <i>M. itinerans</i>
6	Proteína	<i>Musa itinerans</i>	Secuencia de proteínas de <i>FUSR1</i> resistente a FW alelo 2 de <i>M. itinerans</i>
7	Nucleótido	<i>Musa⁶⁷ itinerans</i>	Secuencia de ARNm parcial para el transcrito 2 de <i>FusR1</i> de <i>Musa itinerans</i>
	Nucleótido	<i>Musa</i>	Secuencia de ARNm parcial

(Tabla 1. Continuación)

10	SEQ ID NO.	Tipo de secuencia	Origen	Descripción breve
	SEQ ID NO: 9	Nucleótido	<i>Musa acuminata ssp. banksii</i>	Secuencia de codificación del alelo 1 de <i>FusR1</i> resistente a FW de <i>M. acuminata</i>
15	SEQ ID NO: 10	Nucleótido	<i>Musa acuminata ssp. banksii</i>	Secuencia de ARNm parcial para el alelo 2 de <i>FusR1</i> resistente a FW de <i>M. acuminata</i>
20	SEQ ID NO: 11	Nucleótido	<i>Musa acuminata ssp. banksii</i>	Secuencia de codificación del alelo 2 de <i>FusR1</i> resistente a FW de <i>M. acuminata</i>
25	SEQ ID NO: 12	Proteína	<i>Musa acuminata ssp. banksii</i>	Secuencia de proteínas de <i>FUSR1</i> resistente a FW de <i>M. acuminata</i>
	SEQ ID NO: 13	Nucleótido	<i>Musa acuminata</i>	Secuencia de ARNm parcial para el alelo de <i>FusR1</i> sensible a FW de <i>M. acuminata</i>
30	SEQ ID NO: 14	Nucleótido	<i>Musa acuminata</i>	Secuencia de codificación del alelo <i>FusR1</i> sensible a FW de <i>M. acuminata</i>
	SEQ ID NO: 15	Proteína	<i>Musa acuminata</i>	Secuencia de proteínas de FusR1 sensible a FW de <i>M. acuminata</i>
35	SEQ ID NO: 16	Nucleótido	<i>Musa⁶⁸ acuminata</i>	Secuencia de ARNm parcial del transcrito 2 de <i>FusR1</i> sensible a FW de <i>M. acuminata</i>

(Tabla 1. Continuación)

10	SEQ ID NO.	Tipo de secuencia	Origen	Descripción breve
	SEQ ID NO: 18	Nucleótido	<i>Musa basjoo</i>	Secuencia de codificación de <i>FusR1</i> resistente a FW alelo 1 de <i>M. basjoo</i>
15	SEQ ID NO: 19	Proteína	<i>Musa basjoo</i>	Secuencia de proteínas de <i>FusR1</i> resistente a FW alelo 1 de <i>Musa basjoo</i>
	SEQ ID NO: 20	Nucleótido	<i>Musa basjoo</i>	Secuencia de ARNm parcial de resistente a FW alelo 2 de <i>FusR1</i> de <i>M. basjoo</i>
20	SEQ ID NO: 21	Nucleótido	<i>M. basjoo</i>	Secuencia de codificación parcial de <i>FusR1</i> resistente a FW alelo 2 de <i>M. basjoo</i>
25	SEQ ID NO: 22	Proteína	<i>M. basjoo</i>	Secuencia de proteínas parcial del alelo 2 resistente a FW de <i>FusR1</i> de <i>M. basjoo</i>
	SEQ ID NO: 23	Nucleótido	<i>Musella lasiocarpa</i>	Secuencia de ARNm parcial de <i>FusR1</i> de <i>Musella lasiocarpa</i>
30	SEQ ID NO: 24	Nucleótido	<i>Musella lasiocarpa</i>	Secuencia de codificación de <i>FusR1</i> de <i>M. lasiocarpa</i>
	SEQ ID NO: 25	Proteína	<i>Musella lasiocarpa</i>	Secuencia de proteínas de <i>FUSR1</i> de <i>M. lasiocarpa</i>
35	SEQ ID NO: 26	Nucleótido	<i>M. balbisiana</i>	Secuencia de ARNm parcial de <i>FusR1</i> de la accesión ITC1016 de <i>M. balbisiana</i>
	SEQ ID NO: 27	Nucleótido	⁶⁹ <i>M. balbisiana</i>	Secuencia de codificación "hipotética" de la accesión ITC1016 de <i>M. balbisiana</i>

Tabla 1. Continuación)

10	SEQ ID NO.	Tipo de secuencia	Origen	Descripción breve
	SEQ ID NO: 29	Nucleótido	<i>M. balbisiana</i>	Secuencia de ARNm parcial de <i>FusR1</i> de la accesión ITC0080 de <i>M. balbisiana</i>
15	SEQ ID NO: 30	Nucleótido	<i>M. balbisiana</i>	Secuencia de ARNm parcial de <i>FusR1</i> de la accesión ITC1527 de <i>M. balbisiana</i>
20	SEQ ID NO: 31	Nucleótido	<i>M. acuminata</i> <i>ssp. banksii</i>	Secuencia aguas arriba, que incluye la secuencia promotora , del alelo 1 resistente a FW de <i>FusR1</i> de <i>M. acuminata</i>
	SEQ ID NO: 32	Proteína	<i>M. balbisiana</i>	Secuencia de proteínas de <i>FUSR1</i> de <i>M. balbisiana</i>

*FW - Marchitez por Fusarium

- 25 De acuerdo con la presente descripción, el nuevo gen *FusR1* y sus ortólogos serán útiles para facilitar la construcción de plantas de cultivo que sean resistentes a enfermedades patógenas, especialmente enfermedades provocadas por patógenos fúngicos, virus, nemátodos, insectos y similares. Los genes *FusR1* de la presente descripción también pueden usarse como marcadores en el mapeo genético así como también para evaluar
- 30 la resistencia a enfermedades en una planta de interés. Por tanto, las secuencias pueden usarse en programas de mejoramiento. Ver, por ejemplo, Gentzbittel y otros (1998, Theor. Apl. Genet. 96:519-523). Los usos adicionales de las secuencias de la descripción incluyen el uso de las secuencias como cebo para aislar otros componentes de señalización en las rutas de defensa/resistencia y para aislar las secuencias
- 35 promotoras correspondientes. Las secuencias también pueden usarse para modular los

procesos de desarrollo de las plantas, tales como el desarrollo del polen, la regulación de la forma de los órganos, la diferenciación de aleurona y la epidermis de los brotes, la competencia embriónica y las interacciones célula/célula. Ver, en general, Sambrook y otros, (1989) *Molecular Cloning: A Laboratory Manual* (2da ed., Cold Spring Harbor Laboratory Press, Plainview, N.Y.). Las secuencias de la presente descripción también pueden usarse para generar variantes (por ejemplo, mediante "intercambio de dominios") para la generación de nuevas especificidades de resistencia.

La descripción abarca composiciones de proteínas o ácidos nucleicos aislados o sustancialmente purificados. Una molécula de ácido nucleico o proteína "aislada" o "purificada", o una porción biológicamente activa de la misma, está sustancial o esencialmente libre de componentes que normalmente acompañan o interactúan con la molécula de ácido nucleico o proteína que se encuentra en su entorno natural. Por tanto, un polinucleótido o polipéptido aislado o purificado está sustancialmente libre de otro material celular o medio de cultivo cuando se produce mediante técnicas recombinantes, o sustancialmente libre de precursores químicos u otras sustancias químicas cuando se sintetiza químicamente. De manera adecuada, un polinucleótido "aislado" está libre de secuencias (especialmente secuencias que codifican proteínas) que flanquean naturalmente el polinucleótido (es decir, secuencias ubicadas en los extremos 5' y 3' del polinucleótido) en el ADN genómico del organismo del cual el polinucleótido se derivó.

Por ejemplo, en diversas modalidades, el polinucleótido aislado puede contener menos de aproximadamente 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0,5 kb o 0,1 kb de secuencias de nucleótidos que flanquean naturalmente el polinucleótido en el ADN genómico de la célula de la cual se derivó el polinucleótido. Un polipéptido que está sustancialmente libre de material celular incluye preparaciones de proteína que tienen menos de aproximadamente 30 %, 20 %, 10 %, 5 % (en peso seco) de proteína contaminante. Cuando la proteína de la descripción o la porción biológicamente activa de la misma se produce de manera recombinante, el medio de cultivo representa adecuadamente menos de aproximadamente 30 %, 20 %, 10 % o 5 % (en peso seco) de precursores químicos o químicos de proteínas que no son de interés.

Una porción de una secuencia de nucleótidos *FusR1* que codifica una porción biológicamente activa de un polipéptido *FusR1* de la descripción codificará al menos alrededor de 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 40, 50, 60, 70, 80, 90, 100, 120, 150, 300, 400, 500, 600, 700, 800, 900 o 1000 residuos de aminoácidos contiguos, o casi hasta el número total de aminoácidos presentes en un polipéptido *FUSR1* de longitud completa de la descripción (por ejemplo,

140 residuos de aminoácidos para las SEQ ID NO: 3, 6, 12, 19, o 22, respectivamente). Las porciones de una secuencia de nucleótidos de *FusR1* que son útiles como sondas de hibridación o cebadores de PCR generalmente no necesitan codificar una porción biológicamente activa de un polipéptido *FUSR1*.

- 5 Por tanto, una porción de una secuencia de nucleótidos *FusR1* puede codificar una porción biológicamente activa de un polipéptido *FUSR1*, o puede ser un fragmento que puede usarse como sonda de hibridación o cebador de PCR mediante el uso de métodos estándar conocidos en la técnica. Una porción biológicamente activa de un polipéptido *FUSR1* puede prepararse por aislamiento de una porción de una de las
- 10 secuencias de nucleótidos *FusR1* de la descripción, al expresar la porción codificada del polipéptido *FUSR 1* (por ejemplo, por expresión recombinante in vitro), y mediante la evaluación de la actividad de la porción codificada del polipéptido *FUSR1*. Las moléculas de ácido nucleico que son porciones de una secuencia de nucleótidos *FusR1* comprenden al menos aproximadamente 15, 16, 17, 18, 19, 20, 25, 30, 50, 75, 100, 150,
- 15 200, 250, 300, 350, 400, 450, 500, 550, 600 o 650 nucleótidos, o casi hasta el número de nucleótidos presentes en una secuencia de nucleótidos *FusR1* de longitud completa descrita en la presente descripción (por ejemplo, aproximadamente de 350 a 650 nucleótidos para SEQ ID NO: 1-2, 4-5, 8-10, 17-18 o 20-21, respectivamente). La descripción también contempla variantes de las secuencias de nucleótidos descritas.
- 20 Las variantes de ácido nucleico pueden ser de origen natural, como variantes alélicas (mismo locus), homólogos (locus diferente) y ortólogos (organismo diferente) o pueden ser de origen no natural. Las variantes de origen natural como estas pueden identificarse con el uso de técnicas de biología molecular bien conocidas, como, por ejemplo, con la reacción en cadena de la polimerasa (PCR) y las técnicas de hibridación como se
- 25 conocen en la técnica. Pueden prepararse variantes no naturales mediante técnicas de mutagénesis, incluidas las aplicadas a polinucleótidos, células u organismos. Las variantes pueden contener sustituciones, deleciones, inversiones e inserciones de nucleótidos. La variación puede ocurrir en una o ambas regiones codificantes y no codificantes. Las variaciones pueden producir sustituciones de aminoácidos tanto
- 30 conservadoras como no conservadoras (en comparación con el producto codificado). Para las secuencias de nucleótidos, las variantes conservadoras incluyen aquellas secuencias que, debido a la degeneración del código genético, codifican la secuencia de aminoácidos de uno de los polipéptidos *FUSR1* de la descripción. Las secuencias de nucleótidos variantes también incluyen secuencias de nucleótidos derivadas
- 35 sintéticamente, tales como las generadas, por ejemplo, mediante el uso de mutagénesis

dirigida al sitio pero que todavía codifican un polipéptido FUSR1 de la descripción. Generalmente, las variantes de una secuencia de nucleótidos particular de la descripción tendrán al menos aproximadamente 30 %, 40 %, 50 %, 55 %, 60 %, 65 %, 70 %, generalmente al menos aproximadamente 75 %, 80 %, 85 %, convenientemente

5 aproximadamente del 90 % al 95 % o más, y más adecuadamente aproximadamente 98 % o más de identidad de secuencia con esa secuencia de nucleótidos particular determinada por los programas de alineamiento de secuencia descritos en otra parte de la presente descripción mediante el uso de parámetros predeterminados.

Las variantes de secuencias de nucleótidos también abarcan secuencias derivadas de

10 procedimientos mutagénicos o recombinantes tales como "barajado de ADN" que pueden usarse para intercambiar dominios en un polipéptido de interés con dominios de otros polipéptidos. Con el barajado de ADN, pueden manipularse una o más secuencias codificantes de *FusR1* diferentes para crear una nueva secuencia de *FusR1* que posea las propiedades deseadas. En este procedimiento, se generan bibliotecas de

15 polinucleótidos recombinantes a partir de una población de polinucleótidos relacionados que comprenden regiones de secuencia que tienen una identidad de secuencia sustancial y pueden recombinarse de forma homóloga in vitro o in vivo. Por ejemplo, mediante el uso de este enfoque, los motivos de secuencia que codifican un dominio de interés pueden barajarse entre el gen *FusR1* de la descripción y otros genes *FusR1*

20 conocidos para obtener un nuevo gen que codifica una proteína con una propiedad de interés mejorada, tal como un espectro amplio de resistencia a una enfermedad. Se conocen en la técnica estrategias para el barajado de ADN. Ver, por ejemplo: Stemmer (1994, Proc. Natl. Acad. Sci. USA 91:10747-10751; 1994, Nature 370:389-391); Crameri y otros, (1997, Nature Biotech. 15:436-438); Moore y otros, (1997, J. Mol. Biol. 272:336-

25 347); Zlang y otros, (1997 Proc. Natl. Acad. Sci. USA 94:450-44509); Crameri y otros, (1998, Nature 391:288-291); y las Patentes de los Estados Unidos núms. 5,605,793 y 5,837,458.

La presente descripción proporciona secuencias de nucleótidos que comprenden al menos una porción de las proteínas aisladas codificadas por secuencias de nucleótidos

30 para *FusR1*, homólogos de *FusR1*, ortólogos de *FusR1*, parálogos de *FusR1* y fragmentos y variaciones de los mismos.

En algunas modalidades, la presente descripción proporciona una secuencia de nucleótidos que codifica FUSR1 y/o fragmentos funcionales y variaciones de la misma que comprenden una secuencia de nucleótidos que comparte al menos

35 aproximadamente el 70 %, aproximadamente el 75 %, aproximadamente el 80 %,

aproximadamente el 81 %, aproximadamente el 82 %, aproximadamente el 83 %, aproximadamente el 84 %, aproximadamente el 85 %, aproximadamente el 86 %, aproximadamente el 87 %, aproximadamente el 88 %, aproximadamente el 89 %, aproximadamente el 90 %, aproximadamente el 91 %, aproximadamente el 92 %, aproximadamente el 93 %, aproximadamente el 94 %, aproximadamente el 95 %, aproximadamente el 96 %, aproximadamente el 97 %, aproximadamente el 98 % o aproximadamente el 99 %, aproximadamente el 99,1 %, aproximadamente el 99,2 %, aproximadamente el 99,3 %, aproximadamente el 99,4 %, aproximadamente el 99,5 %, aproximadamente el 99,6 %, aproximadamente el 99,7 %, aproximadamente el 99,8 %, o aproximadamente el 99,9 % de identidad de secuencia con respecto a la SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 17 o SEQ ID NO: 18. En algunas modalidades, una secuencia de nucleótidos que codifica FUSR1 tiene la secuencia de ácido nucleico de SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 17 o SEQ ID NO: 18.

En algunas modalidades, la presente descripción proporciona secuencias de nucleótidos para *FusR1*, homólogos de *FusR1*, ortólogos de *FusR1*, parálogos de *FusR1* y fragmentos y variaciones de los mismos que comprenden secuencias de nucleótidos que comparten al menos aproximadamente el 70 %, aproximadamente el 75 %, aproximadamente el 80 %, aproximadamente el 81 %, aproximadamente el 82 %, aproximadamente el 83 %, aproximadamente el 84 %, aproximadamente el 85 %, aproximadamente el 86 %, aproximadamente el 87 %, aproximadamente el 88 %, aproximadamente el 89 %, aproximadamente el 90 %, aproximadamente el 91 %, aproximadamente el 92 %, aproximadamente el 93 %, aproximadamente el 94 %, aproximadamente el 95 %, aproximadamente el 96 %, aproximadamente el 97 %, aproximadamente el 98 % o aproximadamente el 99 %, aproximadamente el 99,1 %, aproximadamente el 99,2 %, aproximadamente el 99,3 %, aproximadamente el 99,4 %, aproximadamente el 99,5 %, aproximadamente el 99,6 %, aproximadamente el 99,7 %, aproximadamente el 99,8 % o aproximadamente el 99,9 % de identidad de secuencia con respecto a la SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 17 o SEQ ID NO: 18. En algunas modalidades, las secuencias de nucleótidos para *FusR1*, homólogos de *FusR1*, ortólogos de *FusR1*, parálogos de *FusR1* y fragmentos y variaciones de los mismos tienen las secuencias de ácido nucleico de SEQ ID NO: 1, SEQ ID NO: 2, SEQ

ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 17 o SEQ ID NO: 18.

En algunas modalidades, las secuencias de nucleótidos para *FusR1*, homólogos de *FusR1*, ortólogos de *FusR1*, parálogos de *FusR1* y fragmentos y variaciones de los

5 mismos pueden usarse para expresarse en plantas. En algunas modalidades, dichas secuencias de nucleótidos pueden usarse para incorporarse en un casete de expresión, que es capaz de dirigir la expresión de una secuencia de nucleótidos para *FusR1*, homólogos de *FusR1*, ortólogos de *FusR1*, parálogos de *FusR1* y fragmentos y variaciones de los mismos en una célula vegetal, por ejemplo, variedades de plátano
10 descritas en la presente descripción. Este casete de expresión comprende un promotor unido operativamente a la secuencia de nucleótidos de interés (es decir, *FusR1*, ortólogos de *FusR1* y fragmentos y variaciones de los mismos) que está unido operativamente a señales de terminación. También comprende típicamente secuencias necesarias para la traducción adecuada de la secuencia de nucleótidos. La región
15 codificante codifica generalmente una proteína de interés (es decir, FUSR1). En algunas modalidades, el casete de expresión que comprende la secuencia de nucleótidos de *FusR1*, homólogos de *FusR1*, ortólogos de *FusR1*, parálogos de *FusR1* y fragmentos y variaciones de los mismos es quimérico, de modo que al menos uno de sus componentes es heterólogo con respecto a al menos uno de sus otros componentes.

20 En otras modalidades, el casete de expresión es uno que es de origen natural pero que se ha obtenido en una forma recombinante útil para la expresión heteróloga. La expresión de la secuencia de nucleótidos en el casete de expresión puede estar bajo el control de un promotor constitutivo o de un promotor inducible que inicia la transcripción solo cuando la célula huésped se expone a algún estímulo externo particular. Además,
25 la expresión de la secuencia de nucleótidos en el casete de expresión puede estar bajo el control de un promotor específico de tejido. En el caso de un organismo multicelular, el promotor también puede ser específico para un tejido u órgano en particular o una etapa de desarrollo en animales y/o plantas, lo que incluye las especies de plátano.

La presente descripción proporciona polipéptidos y secuencias de aminoácidos que
30 comprenden al menos una porción de las proteínas codificadas por secuencias de nucleótidos para *FusR1*, homólogos de *FusR1*, ortólogos de *FusR1*, parálogos de *FusR1* y fragmentos y variaciones de los mismos.

La presente descripción también proporciona una secuencia de aminoácidos codificada por las secuencias de ácido nucleico de *FusR1*, homólogos de *FusR1*, ortólogos de
35 *FusR1*, parálogos de *FusR1* y/o fragmentos y variaciones de los mismos. En algunas

modalidades, la presente descripción proporciona un polipéptido aislado que comprende una secuencia de aminoácidos que comparte al menos aproximadamente el 70 %, aproximadamente el 75 %, aproximadamente el 80 %, aproximadamente el 85 %, al menos aproximadamente el 90 %, aproximadamente el 91 %, aproximadamente el 92 %, aproximadamente el 93 %, aproximadamente el 94 %, aproximadamente el 95 %, aproximadamente el 96 %, aproximadamente el 97 %, aproximadamente el 98 %, aproximadamente el 99 %, aproximadamente el 99,1 %, aproximadamente el 99,2 %, aproximadamente el 99,3 %, aproximadamente el 99,4 %, aproximadamente el 99,5 %, aproximadamente el 99,6 %, aproximadamente el 99,7 %, aproximadamente el 99,8 % o aproximadamente el 99,9 % de identidad con respecto a una secuencia de aminoácidos codificada por las secuencias de ácido nucleico de *FusR1*, homólogos de *FusR1*, ortólogos de *FusR1*, parálogos de *FusR1* y/o fragmentos y variaciones de los mismos. En una modalidad, la presente descripción proporciona un polipéptido aislado que comprende una secuencia de aminoácidos que codifica una secuencia de aminoácidos que comparte al menos aproximadamente el 85 %, aproximadamente el 86 %, aproximadamente el 87 %, aproximadamente el 88 %, aproximadamente el 89 %, aproximadamente el 90 %, aproximadamente el 91 %, aproximadamente el 92 %, aproximadamente el 93 %, aproximadamente el 94 %, aproximadamente el 95 %, aproximadamente el 96 %, aproximadamente el 97 %, aproximadamente el 98 %, aproximadamente el 99 %, aproximadamente el 99,1 %, aproximadamente el 99,2 %, aproximadamente el 99,3 %, aproximadamente el 99,4 %, aproximadamente el 99,5 %, aproximadamente el 99,6 %, aproximadamente el 99,7 %, aproximadamente el 99,8 % o aproximadamente el 99,9 % de identidad con respecto a una secuencia de aminoácidos codificada por las secuencias de ácido nucleico de *FusR1*, homólogos de *FusR1*, ortólogos de *FusR1*, parálogos de *FusR1* y/o fragmentos y variaciones de los mismos.

La descripción también abarca variantes y fragmentos de proteínas de una secuencia de aminoácidos codificada por las secuencias de ácido nucleico de *FusR1*, homólogos de *FusR1*, ortólogos de *FusR1* y/o parálogos de *FusR1*. Las variantes pueden contener alteraciones en las secuencias de aminoácidos de las proteínas constituyentes. El término "variante" con respecto a un polipéptido se refiere a una secuencia de aminoácidos que está alterada por uno o más aminoácidos con respecto a una secuencia de referencia. La variante puede tener cambios "conservadores" o cambios "no conservadores", por ejemplo, las variaciones menores análogas también pueden incluir deleciones o inserciones de aminoácidos, o ambas.

Los fragmentos funcionales y variantes de un polipéptido incluyen aquellos fragmentos y variantes que mantienen una o más funciones del polipéptido original. Se reconoce que el gen o ADNc que codifica un polipéptido puede mutar considerablemente sin alterar materialmente una o más de las funciones del polipéptido. Primero, se reconoce

5 que el código genético está degenerado y, por tanto, diferentes codones codifican los mismos aminoácidos. En segundo lugar, incluso cuando se introduce una sustitución de aminoácidos, la mutación puede ser conservadora y no tener un impacto material en la función o funciones esenciales de una proteína. Ver, por ejemplo, Stryer Biochemistry 3ra Ed., 1988. En tercer lugar, parte de una cadena polipeptídica puede eliminarse sin

10 alterar o eliminar todas sus funciones. En cuarto lugar, pueden realizarse inserciones o adiciones en la cadena polipeptídica, por ejemplo, añadir etiquetas de epítipo, sin alterar o eliminar sus funciones (Ausubel y otros, J. Immunol. 159(5): 2502-12, 1997). Otras modificaciones que pueden realizarse sin perjudicar materialmente una o más funciones de un polipéptido pueden incluir, por ejemplo, modificaciones químicas y

15 bioquímicas in vivo o in vitro o la incorporación de aminoácidos inusuales. Dichas modificaciones incluyen, pero no se limitan a, por ejemplo, acetilación, carboxilación, fosforilación, glicosilación, ubiquinación, marcaje, por ejemplo, con radionucleótidos, y diversas modificaciones enzimáticas, como apreciarán fácilmente los expertos en la técnica. Una variedad de métodos para marcar polipéptidos y marcadores útiles para

20 tales propósitos se conocen bien en la técnica e incluyen isótopos radiactivos tales como ³²P, ligandos que se unen o están unidos por parejas de unión específicos marcados (por ejemplo, anticuerpos), fluoróforos, agentes quimioluminiscentes, enzimas y anti-ligandos. Los fragmentos y variantes funcionales pueden ser de longitud variable. Por ejemplo, algunos fragmentos tienen al menos 10, 25, 50, 75, 100, 200 o incluso más

25 residuos de aminoácidos. Estas mutaciones pueden ser naturales o cambiadas intencionalmente. En algunas modalidades, las mutaciones que contienen alteraciones que producen sustituciones, adiciones o deleciones silenciosas, pero que no alteran las propiedades o actividades de las proteínas o cómo se fabrican las proteínas son una modalidad de la descripción.

30 Las sustituciones conservadoras de aminoácidos son aquellas sustituciones que, cuando se realizan, interfieren menos con las propiedades de la proteína original, es decir, la estructura y especialmente la función de la proteína se conserva y no cambia significativamente por dichas sustituciones. Las sustituciones conservadoras mantienen generalmente (a) la estructura de la cadena polipeptídica en el área de la sustitución,

35 por ejemplo, como una conformación de lámina o helicoidal, (b) la carga o hidrofobicidad

de la molécula en el sitio objetivo, o (c) la mayor parte de la cadena lateral. Puede encontrarse más información sobre sustituciones conservadoras, por ejemplo, en Ben Bassat y otros, (J. Bacteriol., 169:751 757, 1987), O'Regan y otros, (Gene, 77:237 251, 1989), Sahin Toth y otros, (Protein Sci., 3:240 247, 1994), Hochuli y otros, (Bio/Technology, 6:1321 1325, 1988) y en libros de texto de genética y biología molecular ampliamente usados. Las matrices de Blosum se usan comúnmente para determinar el parentesco de secuencias de polipéptidos. Las matrices de Blosum se crearon mediante el uso de una gran base de datos de alineamientos confiables (la base de datos BLOCKS), en la que se contaron los alineamientos de secuencia por pares relacionados por menos de algún porcentaje de identidad umbral (Henikoff y otros, Proc. Natl. Acad. Sci. USA, 89:10915-10919, 1992). Se usó un umbral de identidad del 90 % para las frecuencias objetivo altamente conservadas de la matriz BLOSUM90. Se usó un umbral de identidad del 65 % para la matriz BLOSUM65. Las puntuaciones de cero y superiores en las matrices de Blosum se consideran "sustituciones conservadoras" en el porcentaje de identidad seleccionado. La siguiente tabla 2 muestra sustituciones de aminoácidos conservativas ilustrativas.

20

25

30

35

Tabla 2. Sustituciones conservadoras de aminoácidos ilustrativas enumeradas

	Residuo original	Sustituciones muy altamente conservadas	Sustituciones altamente conservadas (a partir de la matriz Blosum90)	Sustituciones conservadas (a partir de la matriz Blosum65)
5	Ala	Ser	Gly, Ser, Thr	Cys, Gly, Ser, Thr, Val
	Arg	Lys	Gln, His, Lys	Asn, Gln, Glu, His, Lys
10	Asn	Gln; His	Asp, Gln, His, Lys, Ser, Thr	Arg, Asp, Gln, Glu, His, Lys, Ser, Thr
	Asp	Glu	Asn, Glu	Asn, Gln, Glu, Ser
	Cys	Ser	None	Ala
	Gln	Asn	Arg, Asn, Glu, His, Lys, Met	Arg, Asn, Asp, Glu, His, Lys, Met, Ser
15	Glu	Asp	Asp, Gln, Lys	Arg, Asn, Asp, Gln, His, Lys, Ser
	Gly	Pro	Ala	Ala, Ser
	His	Asn; Gln	Arg, Asn, Gln, Tyr	Arg, Asn, Gln, Glu, Tyr
	Ile	Leu; Val	Leu, Met, Val	Leu, Met, Phe, Val
20	Leu	Ile; Val	Ile, Met, Phe, Val	Ile, Met, Phe, Val
	Lys	Arg; Gln; Glu	Arg, Asn, Gln, Glu	Arg, Asn, Gln, Glu, Ser,
	Met	Leu; Ile	Gln, Ile, Leu, Val	Gln, Ile, Leu, Phe, Val
	Phe	Met; Leu; Tyr	Leu, Trp, Tyr	Ile, Leu, Met, Trp, Tyr
	Ser	Thr	Ala, Asn, Thr	Ala, Asn, Asp, Gln, Glu, Gly, Lys, Thr
25	Thr	Ser	Ala, Asn, Ser	Ala, Asn, Ser, Val
	Trp	Tyr	Phe, Tyr	Phe, Tyr
	Tyr	Trp; Phe	His, Phe, Trp	His, Phe, Trp
	Val	Ile; Leu	Ile, Leu, Met	Ala, Ile, Leu, Met, Thr

30 En algunos ejemplos, las variantes no pueden tener más de 3, 5, 10, 15, 20, 25, 30, 40, 50 o 100 cambios conservadores de aminoácidos (tal como sustituciones de aminoácidos muy altamente conservadas o altamente conservadas). En otros ejemplos, uno o varios residuos hidrófobos (tales como Leu, Ile, Val, Met, Phe o Trp) en una secuencia variante pueden reemplazarse con un residuo hidrófobo diferente (tal como

Leu, Ile, Val, Met, Phe, o Trp) para crear una variante funcionalmente similar a las secuencias de aminoácidos descritas codificadas por las secuencias de ácido nucleico de *FusR1*, homólogos de *FusR1*, ortólogos de *FusR1* y/o parálogos de *FusR1*, y/o fragmentos y variaciones de los mismos.

- 5 En algunas modalidades, las variantes pueden diferir de las secuencias descritas mediante la alteración de la región codificante para ajustarse al sesgo de uso del codón del organismo particular en el que se va a introducir la molécula. En otras modalidades, la región codificante puede alterarse al aprovechar la degeneración del código genético para alterar la secuencia codificante de manera que, aunque la secuencia de nucleótidos
- 10 está sustancialmente alterada, codifica no obstante una proteína que tiene una secuencia de aminoácidos sustancialmente similar a la descrita, secuencias de aminoácidos codificadas por las secuencias de ácidos nucleicos de *FusR1*, homólogos de *FusR1*, ortólogos de *FusR1* y/o parálogos de *FusR1*, y/o fragmentos y variaciones de los mismos.
- 15 En algunas modalidades, se proporcionan fragmentos funcionales derivados de los ortólogos de *FusR1* de la presente descripción. Los fragmentos funcionales aún pueden conferir resistencia a patógenos cuando se expresan en una planta. En algunas modalidades, los fragmentos funcionales contienen al menos la región conservada o el dominio inhibidor de Bowman-Birk de un ortólogo de *FusR1* de tipo silvestre, o variantes
- 20 funcionales de los mismos. En algunas modalidades, los fragmentos funcionales contienen una o más regiones conservadas compartidas por dos o más ortólogos de *FusR1*, compartidas por dos o más ortólogos de *FusR1* en el mismo género de plantas, compartidas por dos o más ortólogos de *FUSR1* de dicotiledóneas y/o compartidas por dos o más ortólogos de *FusR1* de monocotiledóneas. Las regiones conservadas o los
- 25 dominios del inhibidor de Bowman-Birk pueden determinarse mediante cualquier programa informático adecuado, tal como el programa BLAST de proteínas de NCBI y el programa Alignment de NCBI, o programas equivalentes. En algunas modalidades, los fragmentos funcionales son 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41,
- 30 42, 43, 44, 45, 46, 47, 48, 49, 50 o más aminoácidos más cortos en comparación con los ortólogos de *FusR1* de la presente descripción. En algunas modalidades, los fragmentos funcionales se obtienen al eliminar uno o más aminoácidos de los ortólogos de *FusR1* de la presente descripción. En algunas modalidades, los fragmentos funcionales comparten al menos 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % o
- 35 más de identidad con respecto a los ortólogos de *FusR1* de la presente descripción.

En algunas modalidades, se proporcionan polipéptidos funcionales quiméricos o sintéticos derivados de los ortólogos de *FusR1* de la presente descripción. Los polipéptidos funcionales quiméricos o sintéticos aún pueden conferir resistencia a patógenos cuando se expresan en una planta. En algunas modalidades, los polipéptidos sintéticos o quiméricos funcionales contienen al menos la región conservada o el dominio inhibidor de Bowman-Birk de un FUSR1 de tipo silvestre ortólogos, o variantes funcionales de los mismos. En algunas modalidades, los polipéptidos sintéticos o quiméricos funcionales contienen una o más regiones conservadas compartidas por dos o más ortólogos de FUSR1, compartidas por dos o más ortólogos de *FusR1* en el mismo género de plantas, compartidas por dos o más ortólogos de *FusR1* de monocotiledóneas y/o compartidas por dos o más ortólogos de FUSR1 de dicotiledóneas. Las regiones conservadas ilustrativas no limitantes se muestran en la Figura 2. Las regiones conservadas o los dominios del inhibidor de Bowman-Birk pueden determinarse mediante cualquier programa informático adecuado, tal como el programa BLAST de proteínas de NCBI y el programa Alignment de NCBI, o programas equivalentes. En algunas modalidades, los polipéptidos sintéticos o quiméricos funcionales comparten al menos 80 %, 85 %, 90 %, 95 %, 96 %, 97 %, 98 %, 99 % o más de identidad con respecto a los ortólogos de *FusR1* de la presente descripción.

Las secuencias de regiones conservadas exclusivas de los alelos sensibles a FW también pueden usarse para reducir el nivel de uno o más ortólogos de *FusR1*. En algunas modalidades, pueden usarse secuencias de regiones conservadas para generar moléculas de silenciamiento génico para dirigirse a uno o más ortólogos de *FusR1*. En algunas modalidades, las moléculas de silenciamiento génico se seleccionan del grupo que consiste en polinucleótidos bicatenarios, polinucleótidos monocatenarios u oligonucleótidos dúplex mixtos. En algunas modalidades, las moléculas de silenciamiento génico comprenden un fragmento de ADN/ARN de aproximadamente 10 pb, 15 pb, 19 pb, 20 pb, 21 pb, 25 pb, 30 pb, 40 pb, 50 pb, 60 pb, 70 pb, 80 pb, 90 pb, 100 pb, 150 pb, 200 pb, 250 pb, 300 pb, 350 pb, 400 pb, 500 pb, 600 pb, 700 pb, 800 pb, 900 pb, 1000 pb o más polinucleótidos, en donde el fragmento de ADN/ARN comparte al menos el 90 %, 95 %, 99 % o más de identidad con respecto a una región conservada de las secuencias de ortólogos de *FusR1* de la presente descripción, o secuencias complementarias de las mismas.

V. Transformación de plantas

Los presentes polinucleótidos que codifican FUSR1, homólogos de *FusR1*, ortólogos de *FusR1* y/o parálogos de *FusR1*, y/o fragmentos y variaciones de los mismos de la presente descripción pueden transformarse en plátano u otros géneros de plantas.

Los expertos en la técnica conocen bien métodos para producir plantas transgénicas.

- 5 Ahora pueden producirse plantas transgénicas mediante una variedad de métodos de transformación diferentes, lo que incluye, pero no se limitan a, electroporación; microinyección; bombardeo de microproyectiles, también conocido como aceleración de partículas o bombardeo biolístico; transformación mediada por virus; y transformación mediada por *Agrobacterium*. Ver, por ejemplo, las Patentes de los Estados Unidos
- 10 núms. 5,405,765; 5,472,869; 5,538,877; 5,538,880; 5,550,318; 5,641,664; 5,736,369 y 5,736,369; Publicaciones de Solicitud de Patente Internacional núms. WO2002/038779 y WO/2009/117555; Lu y otros, (Plant Cell Reports, 2008, 27:273-278); Watson y otros, Recombinant DNA, Scientific American Books (1992); Hinchey y otros, Bio/Tech. 6:915-922 (1988); McCabe y otros, Bio/Tech. 6:923-926 (1988); Toriyama y otros, Bio/Tech.
- 15 6:1072-1074 (1988); Fromm y otros, Bio/Tech. 8:833-839 (1990); Mullins y otros, Bio/Tech. 8:833-839 (1990); Hiei y otros, Plant Molecular Biology 35:205-218 (1997); Ishida y otros, Nature Biotechnology 14:745-750 (1996); Zhang y otros, Molecular Biotechnology 8:223-231 (1997); Ku y otros, Nature Biotechnology 17:76-80 (1999); y Raineri y otros, Bio/Tech. 8:33-38 (1990)), cada uno de los cuales se incorpora en la
- 20 presente descripción expresamente como referencia en su totalidad.

Agrobacterium tumefaciens es una bacteria de origen natural que es capaz de insertar su ADN (información genética) en las plantas, lo que da como resultado un tipo de daño a la planta conocido como agalla de la corona. La mayoría de las especies de plantas ahora pueden transformarse mediante el uso de este método, lo que incluye las

25 especies de cucurbitáceas.

El bombardeo de microproyectiles también se conoce como aceleración de partículas, bombardeo biolístico y pistola de genes (Biolistic® Gene Gun). La pistola de genes se usa para disparar gránulos que están recubiertos con genes (por ejemplo, para los rasgos deseados) en semillas de plantas o tejidos de plantas para que las células de las

30 plantas expresen los nuevos genes. La pistola de genes usa un explosivo real (blanco calibre .22) para propulsar el material. También puede usarse aire comprimido o vapor como propulsor. La pistola de genes Biolistic® se inventó en 1983-1984 en la Universidad de Cornell por John Sanford, Edward Wolf y Nelson Allen. Tanto ella como su marca registrada son ahora propiedad de E.I. du Pont de Nemours and Company. La

mayoría de las especies de plantas se han transformado mediante el uso de este método.

El método más común para la introducción de nuevo material genético en el genoma de una planta implica el uso de células vivas del patógeno bacteriano *Agrobacterium tumefaciens* para inyectar literalmente un fragmento de ADN, llamado ADN-T o de transferencia, en células vegetales individuales (generalmente después de una herida del tejido) donde se dirige al núcleo de la planta para la integración cromosómica. Existen numerosas patentes que rigen la transformación mediada por *Agrobacterium* y plásmidos de suministro de ADN particulares diseñados específicamente para su uso con *Agrobacterium* --- por ejemplo, US4536475, EP0265556, EP0270822, WO8504899, WO8603516, US5591616, EP0604662, EP0672752, WO8603776, WO9209696, WO9419930, WO9967357, US4399216, WO8303259, US5731179, EP068730, WO9516031, US5693512, US6051757 y EP904362A1. La transformación de plantas mediada por *Agrobacterium* implica como primera etapa la colocación de fragmentos de ADN clonados en plásmidos en células de *Agrobacterium* vivas, que después se usan para la transformación en células vegetales individuales. La transformación de plantas mediada por *Agrobacterium* es, por tanto, un método indirecto de transformación de plantas. Los métodos de transformación de plantas mediada por *Agrobacterium* que implican el uso de vectores sin ADN-T también se conocen bien por los expertos en la técnica y pueden tener aplicabilidad en la presente descripción. Ver, por ejemplo, la Patente de los Estados Unidos núm. 7,250,554, que usa ADN-P en lugar de T-DNA en el vector de transformación.

Una planta transgénica formada mediante el uso de métodos de transformación por *Agrobacterium* típicamente contiene un solo gen en un cromosoma, aunque son posibles múltiples copias. Puede hacerse referencia a dichas plantas transgénicas como hemigotas para el gen añadido. Un nombre más exacto para dicha planta es un segregante independiente, porque cada planta transformada representa un evento de integración del ADN-T único (Patente de los Estados Unidos núm. 6,156,953). Un locus transgénico se caracteriza generalmente por la presencia y/o ausencia del transgén. Un genotipo heterocigoto en el que un alelo corresponde a la ausencia del transgén también se designa como hemigoto (Patente de los Estados Unidos núm. 6,008,437).

También se informan métodos de transformación directa de plantas mediante el uso de ADN. El primero de estos que se informó históricamente es la electroporación, que utiliza una corriente eléctrica aplicada a una solución que contiene células vegetales (M. E. Fromm y otros, *Nature*, 319, 791 (1986); H. Jones y otros, *Plant Mol. Biol.*, 13, 501 (1989)

y H. Yang y otros, *Plant Cell Reports*, 7, 421 (1988). Otro método directo, llamado "bombardeo biolístico", utiliza partículas ultrafinas, generalmente de tungsteno u oro, que se recubren con ADN y después se rocían sobre la superficie de un tejido vegetal con fuerza suficiente para hacer que las partículas penetren en las células vegetales, lo que incluye la gruesa pared celular, membrana y envoltura nuclear, pero sin matar al menos a algunas de ellas (documentos US 5,204,253, US 5,015,580). Un tercer método directo utiliza formas fibrosas de metal o cerámica que consisten en proyecciones afiladas, porosas o huecas en forma de aguja que literalmente atraviesan las células, y también la envoltura nuclear de las células. Tanto las fibras de carburo de silicio como de borato de aluminio se han usado para la transformación de plantas (Mizuno y otros, 2004; Petolino y otros, 2000; US5302523 Solicitud de Estados Unidos 20040197909) y también para la transformación de bacterias y animales (Kaepler y otros, 1992; Raloff, 1990; Wang, 1995). Existen otros métodos informados y, sin duda, se desarrollarán métodos adicionales. Sin embargo, las eficiencias de cada uno de estos métodos indirectos o directos para introducir ADN extraño en las células vegetales son invariablemente extremadamente bajas, lo que hace necesario usar algún método para la selección de solo aquellas células que se han transformado y, además, permitir el crecimiento y la regeneración en plantas de sólo aquellas células que se han transformado.

Para una transformación de plantas eficiente, debe emplearse un método de selección de manera que las plantas completas se regeneren a partir de una sola célula transformada y cada célula de la planta transformada porte el ADN de interés. Estos métodos pueden emplear selección positiva, mediante la cual se suministra un gen extraño a una célula vegetal que le permite utilizar un sustrato presente en el medio que de cualquier otra manera no podría usar, tal como manosa o xilosa (por ejemplo, ver los documentos US 5767378; US 5994629). Sin embargo, se usa más típicamente la selección negativa porque es más eficiente, ya que utiliza agentes selectivos tales como herbicidas o antibióticos que matan o inhiben el crecimiento de células vegetales no transformadas y reducen la posibilidad de quimeras. Los genes de resistencia que son eficaces contra agentes de selección negativa se proporcionan en el ADN extraño introducido que se utiliza para la transformación de la planta. Por ejemplo, uno de los agentes selectivos más populares que se utiliza es el antibiótico kanamicina, junto con el gen de resistencia neomicina fosfotransferasa (*nptII*), que confiere resistencia a kanamicina y antibióticos relacionados (ver, por ejemplo, Messing & Vierra, *Gene* 19:259- 268 (1982); Bevan y otros, *Nature* 304:184-187 (1983)). Sin embargo, pueden

usarse muchos antibióticos diferentes y genes de resistencia a antibióticos con propósitos de transformación (ver los documentos US 5034322, US 6174724 y US 6255560). Además, pueden usarse varios herbicidas y genes de resistencia a herbicidas para la transformación, lo que incluye el gen bar, que confiere resistencia al herbicida fosfinotricina (White y otros, Nucl Acids Res 18: 1062 (1990), Spencer y otros, Theor Appl Genet 79: 625-631(1990), documentos US 4795855, US 5378824 y US 6107549). Además, el gen dhfr, que confiere resistencia al agente anticanceroso metotrexato, se ha usado para la selección (Bourouis y otros, EMBO J. 2(7): 1099-1104 (1983).

Los elementos de control de expresión que se utilizan para regular la expresión de una proteína dada pueden ser el elemento de control de la expresión que normalmente se encuentra asociado con la secuencia codificante (elemento de expresión homólogo) o puede ser un elemento de control de expresión heterólogo. Se conocen en la técnica una variedad de elementos de control de expresión homólogos y heterólogos y pueden usarse fácilmente para generar unidades de expresión para su uso en la presente descripción. Las regiones de iniciación de la transcripción, por ejemplo, pueden incluir cualquiera de las diversas regiones de iniciación de opina, tales como octopina, manopina, nopalina y similares que se encuentran en los plásmidos Ti de *Agrobacterium tumefaciens*. De manera alternativa, también pueden usarse promotores virales de plantas, tales como los promotores 19S y 35S del virus del mosaico de la coliflor (promotores CaMV 19S y CaMV 35S, respectivamente) para controlar la expresión génica en una planta (Patentes de los Estados Unidos núms. 5,352,605; 5,530,196 y 5,858,742, por ejemplo). También pueden utilizarse secuencias potenciadoras derivadas del CaMV (Patentes de los Estados Unidos núms. 5,164,316; 5,196,525; 5,322,938; 5,530,196; 5,352,605; 5,359,142; y 5,858,742, por ejemplo). Por último, también pueden utilizarse promotores de plantas tales como el promotor prolifera, promotores específicos de la fruta, promotor Ap3, promotores de choque térmico, promotores específicos de la semilla, etc.

Ya sea un promotor específico de gametos, un promotor constitutivo (tal como el promotor CaMV o Nos), un promotor específico de órgano (tal como el promotor E8 de tomate) o un promotor inducible se une típicamente a la proteína o región codificante antisentido mediante el uso de técnicas estándar conocidas en la técnica. La unidad de expresión puede optimizarse aún más mediante el empleo de elementos suplementarios tales como terminadores de la transcripción y/o elementos potenciadores.

Por tanto, para la expresión en plantas, las unidades de expresión contendrán típicamente, además de la secuencia proteica, una región promotora de la planta, un

sitio de iniciación de la transcripción y una secuencia de terminación de la transcripción. Típicamente se incluyen sitios únicos de enzimas de restricción en los extremos 5' y 3' de la unidad de expresión para permitir una fácil inserción en un vector preexistente.

En la construcción de promotor/gen estructural heterólogo o combinaciones antisentido, el promotor se sitúa, preferentemente, a la misma distancia del sitio de inicio de la transcripción heterólogo que del sitio de inicio de la transcripción en su entorno natural. Sin embargo, como se conoce en la técnica, puede acomodarse alguna variación en esta distancia sin pérdida de la función promotora.

Además de una secuencia promotora, el casete de expresión también puede contener una región de terminación de la transcripción a continuación del gen estructural para proporcionar una terminación eficiente. La región de terminación puede obtenerse del mismo gen que la secuencia promotora o puede obtenerse de genes diferentes. Si el ARNm codificado por el gen estructural se va a procesar eficientemente, también se añaden comúnmente al constructo del vector secuencias de ADN que dirigen la poliadenilación del ARN. Las secuencias de poliadenilación incluyen, pero no se limitan a, la señal de octopina sintasa de *Agrobacterium* (Gielen y otros, EMBO J 3:835-846 (1984)) o la señal de nopalina sintasa (Depicker y otros, Mol. and Appl. Genet. 1:561-573 (1982)). La unidad de expresión resultante se une o se construye de cualquier otra manera para incluirla en un vector que es apropiado para la transformación de plantas superiores. Pueden incluirse una o más unidades de expresión en el mismo vector. El vector contendrá típicamente una unidad de expresión del gen marcador seleccionable mediante la cual las células vegetales transformadas pueden identificarse en cultivo. Generalmente, el gen marcador codificará la resistencia a un antibiótico, tal como G418, higromicina, bleomicina, kanamicina o gentamicina, o a un herbicida, tal como glifosato (Round-Up) o glufosinato (BASTA) o atrazina. Generalmente, también se incluyen secuencias de replicación, de origen bacteriano o viral, para permitir que el vector se clone en un huésped bacteriano o fago; preferentemente, se incluye una amplia gama de huéspedes para el origen de replicación procariota. También puede incluirse un marcador seleccionable para bacterias que permita la selección de células bacterianas que portan el constructo deseado. Los marcadores seleccionables procariotas adecuados incluyen la resistencia a antibióticos tales como ampicilina, kanamicina o tetraciclina. También pueden estar presentes en el vector otras secuencias de ADN que codifican funciones adicionales, como se conoce en la técnica. Por ejemplo, en el caso de las transformaciones por *Agrobacterium*, también se incluirán secuencias de ADN-T para su posterior transferencia a los cromosomas de las plantas.

Para introducir un gen deseado o un conjunto de genes mediante métodos convencionales se requiere un cruzamiento sexual entre dos líneas, y, a continuación, realizar repetidos retrocruzamientos entre la descendencia híbrida y uno de los progenitores hasta obtener una planta con las características deseadas. Este proceso, sin embargo, está restringido a plantas que pueden hibridarse sexualmente, y se transferirán otros genes además del gen deseado.

Las técnicas de ADN recombinante permiten a los investigadores de plantas eludir estas limitaciones al permitir a los genetistas de plantas identificar y clonar genes específicos para rasgos convenientes, tales como una composición mejorada de ácidos grasos, e introducir estos genes en variedades de plantas que ya son útiles. Una vez que los genes extraños se introducen en una planta, esta planta puede usarse después en esquemas de reproducción de plantas (por ejemplo, mejoramiento de pedigrí, esquemas de mejoramiento de descendencia de una sola semilla, selección recíproca recurrente) para producir una progenie que también contenga el gen de interés.

Los genes pueden introducirse en el sitio de una manera dirigida mediante el uso de recombinación homóloga. La recombinación homóloga permite modificaciones específicas de sitio en genes endógenos y, por tanto, pueden corregirse mutaciones heredadas o adquiridas y/o pueden introducirse nuevas alteraciones en el genoma. La recombinación homóloga y la integración dirigida al sitio en plantas se analiza, por ejemplo, en las Patentes de los Estados Unidos núms. 5,451,513; 5,501,967 y 5,527,695.

De acuerdo con Ploetz (2015, *Phytopathology* 105:1512-1521), "La transformación genética de los plátanos se ha vuelto un lugar común, y la resistencia a enfermedades es uno de los rasgos más buscados [se omiten las citas]". Las técnicas para transformar y regenerar plantas de plátano se conocen bien en la técnica. Ver, por ejemplo, la Patente de los Estados Unidos núm. 7,534,930; Patente de los Estados Unidos núm. 6,133,035; Sagi y otros, *Bio/Technology* 13, 481-485, 1995; May y otros, *Bio/Technology* 13, 485-492, 1995; Vishnevetsky y otros, *Transgenic Res.* 20(1):61-71, 2011; Paul y otros (2011); Zhong y otros, *Plant Physiol.* 110, 1097-1107, 1996; y Dugdale y otros, *Journal of General Virology* 79:2301-2311, 1998, cada uno de los cuales se incorpora expresamente en la presente descripción como referencia en su totalidad. Para resúmenes e historia, ver, por ejemplo, Mohan y Swennen (editores), 2004, *Banana improvement: cellular, molecular biology, and induced mutations*, Science Publishers, Inc.; y, Remy y otros, 2013, *Genetically modified bananas: Past, present and future*, Acta

Horticulturae 974:71-80, cada uno de los cuales se incorpora expresamente en la presente descripción como referencia en su totalidad.

Al reducir la presente invención a la práctica, el inventor puede construir un constructo de expresión que incluye secuencias de nucleótidos que codifican FUSR1, homólogos de *FusR1*, ortólogos de *FusR1* y/o parálogos de *FusR1*, y/o fragmentos y variaciones de los mismos. El constructo de expresión de la presente invención puede introducirse en el callo embriogénico del plátano comercial y las células transformadas resultantes pueden regenerarse en plantas. Se espera que las plantas de plátano transgénicas tengan expresión de la proteína FUSR1 resistente a FW y resistencia a patógenos.

De acuerdo con un aspecto de la presente invención, se proporciona un método para producir una ~~planta de plátano~~planta de banano resistente a enfermedades. El método se efectúa mediante la transformación de una célula de plátano con al menos un polinucleótido exógeno que codifica un polipéptido (tal como *FusR1* resistente a FW) capaz de conferir resistencia a enfermedades a una ~~planta de plátano~~planta de banano.

De acuerdo con otro aspecto de la presente invención, se proporciona un método para producir una ~~planta de plátano~~planta de banano resistente a enfermedades. El método se efectúa al transformar una célula de plátano con al menos un casete de expresión exógeno que contiene polinucleótidos que codifican una proteína efectora asociada a CRISPR y un ARN guía capaz de dirigirse al menos a un alelo *FusR1* sensible a FW, lo que confiere de esta manera resistencia a la enfermedad a una ~~planta de plátano~~planta de banano.

La célula de plátano de la presente invención puede ser cualquier variedad o cultivo de plátano, lo que incluye, pero no se limita a, *M. acuminata* comercialmente importante (Cavendish, Cavendish enano, Grand Nain, etc.). Preferentemente, la célula de plátano usada para la transformación es una célula embriogénica que es capaz de formar una planta completa. Con mayor preferencia, la célula de plátano es una célula de callo embriogénico.

La frase "célula de callo embriogénico" usada en la presente descripción se refiere a una célula embriogénica contenida en una masa celular producida in vitro.

Las células de callo embriogénico de plátano adecuadas para la transformación pueden generarse mediante el uso de una metodología bien conocida. Por ejemplo, las flores masculinas inmaduras (inflorescencias) pueden disecarse e incubarse en medio M1 (ver contenido en la Tabla 1 de la presente descripción más abajo) bajo una intensidad de luz reducida (50-100 lux) a 25 °C. Después de 3-5 meses de incubación en medio M1, se transfieren callos embriogénicos amarillos a medio M2 (ver contenido en la Tabla 1

más abajo) y se incuban a 27 °C en la oscuridad durante al menos cuatro meses para promover la embriogénesis.

Como se mencionó anteriormente, dichas células de callos embriogénicos de plátano son adecuadas para la transformación con un constructo de ácido nucleico que incluye al menos un polinucleótido que codifica un polipéptido resistente a la enfermedad.

Las frases "polipéptido capaz de conferir resistencia a la enfermedad" y "polipéptido de resistencia a la enfermedad" se usan indistintamente en la presente descripción para referirse a cualquier péptido, polipéptido o proteína que sea capaz de proteger una ~~planta de plátano~~ planta de banano (que expresa el polipéptido) de la infección por patógenos o los efectos dañinos resultantes de la infección por patógenos.

Un polipéptido de resistencia a la enfermedad adecuado también puede ser un polipéptido capaz de inducir o potenciar la resistencia en plantas tal como se describe, por ejemplo, en las Patentes de los Estados Unidos núms. 6,091,004 y 6,316,697.

Como se mencionó anteriormente, el método de la presente invención se efectúa al transformar una célula de plátano con al menos un polinucleótido que codifica un polipéptido capaz de conferir resistencia a la enfermedad a una ~~planta de plátano~~ planta de banano.

En algunas modalidades, la célula de plátano se transforma con una secuencia de polinucleótidos que codifica la proteína FUSR1 de *Musa itinerans*, un ejemplo de la cual se expone en la SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 4 y SEQ ID NO: 5.

En algunas modalidades, la célula de plátano se transforma con una secuencia de polinucleótidos que codifica la proteína FUSR1 de *Musa acuminata*, un ejemplo de la cual se expone en la SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10 y SEQ ID NO: 11.

En algunas modalidades, la célula de plátano se transforma con una secuencia de polinucleótidos que codifica la proteína FUSR1 de *Musa basjoo*, un ejemplo de la cual se expone en la SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 20 y SEQ ID NO: 21.

En algunas modalidades, la célula de plátano se transforma con una secuencia de polinucleótidos que codifica la proteína FUSR1 de *Musella lasiocarpa*, un ejemplo de la cual se expone en la SEQ ID NO: 23.

En algunas modalidades, la célula de plátano se transforma con una secuencia de polinucleótidos que codifica la proteína FUSR1 de *Musa balbisiana*, un ejemplo de la cual se expone en SEQ ID NO: 26.

En algunas modalidades, las plantas transformadas con un solo polipéptido exógeno de resistencia a la enfermedad, tal como FUSR1, pueden exhibir solo una protección parcial y de corta duración (ver, por ejemplo, en Jach y otros, Plant J. 8:97-108, 1995). En otras

modalidades, la célula/~~planta de plátano~~planta de banano de la presente invención expresa, preferentemente, una pluralidad de polipéptidos exógenos de resistencia a la enfermedad y es, por tanto, sustancialmente más resistente a la enfermedad que las plantas no modificadas.

- 5 Pueden utilizarse varios enfoques para transformar y coexpresar estos polinucleótidos en células vegetales.

Aunque se prefiere menos, cada una de las secuencias de polinucleótidos descritas anteriormente puede introducirse por separado en una célula de plátano mediante el uso de tres constructos de ácido nucleico separados. En algunas modalidades, las tres

- 10 secuencias de polinucleótidos pueden introducirse y expresarse conjuntamente en la célula de plátano mediante el uso de un único constructo de ácido nucleico. Dicho constructo puede diseñarse con una sola secuencia promotora que puede transcribir un mensaje policistrónico que incluye las tres secuencias polinucleotídicas. Para permitir la traducción conjunta de los tres polipéptidos codificados por el mensaje policistrónico, las
- 15 secuencias de polinucleótidos pueden estar interconectadas mediante una secuencia del sitio de entrada al ribosoma interno (IRES) que facilita la traducción de secuencias de polinucleótidos situadas a continuación de la secuencia IRES. En este caso, una molécula de ARN policistrónico transcrita que codifica los tres polipéptidos descritos anteriormente se traducirá tanto a partir del extremo 5' tapado como de las dos
- 20 secuencias IRES internas de la molécula de ARN policistrónico para producir de esta manera en la célula los tres polipéptidos.

De manera alternativa, los segmentos de polinucleótidos que codifican la pluralidad de polipéptidos capaces de conferir resistencia a la enfermedad pueden fusionarse traslacionalmente mediante un sitio de reconocimiento de proteasa escindible por una

25 proteasa expresada por la célula que se va a transformar con el constructo de ácido nucleico. En este caso, un polipéptido quimérico traducido será escindido por una proteasa expresada en células para generar de esta manera la pluralidad de polipéptidos.

- En otras modalidades, la presente invención utiliza un constructo de ácido nucleico que
- 30 incluye tres secuencias promotoras, cada una de ellas capaz de dirigir la transcripción de una secuencia de polinucleótidos específica de las secuencias de polinucleótidos descritas anteriormente.

Los promotores adecuados que pueden usarse con el ácido nucleico de la presente invención incluyen promotores constitutivos, inducibles o específicos de tejido.

Los promotores constitutivos adecuados incluyen, por ejemplo, el promotor CaMV 35S (Odell y otros, *Nature* 313:810-812, 1985); maíz Ubi 1 (Christensen y otros, *Plant Sol. Biol.* 18:675-689, 1992); actina de arroz (McElroy y otros, *Plant Cell* 2:163-171, 1990); pEMU (Last y otros, *Theor. Appl. Genet.* 81:581-588, 1991); y Super MAS Sintético (Ni y
5 otros, *The Plant Journal* 7: 661-76, 1995). Otros promotores constitutivos incluyen los de las Patentes de los Estados Unidos núms. 5,659,026, 5,608,149; 5,608,144; 5,604,121; 5,569,597; 5,466,785; 5,399,680; 5,268,463 y 5,608,142.

Los promotores inducibles adecuados pueden ser promotores inducibles por patógenos tales como, por ejemplo, el promotor PR10 de alfalfa (Coutos-Thevenot y otros, *Journal*
10 *of Experimental Botany* 52: 901-910, 2001 y los promotores descritos por Marineau y otros, *Plant Mol. Biol.* 9:335-342, 1987; Matton y otros, *Molecular Plant-Microbe Interactions* 2:325-331, 1989; Somsisch y otros, *Proc. Natl. Acad. Sci. USA* 83:2427-2430, 1986; Somsisch y otros, *Mol. Gen. Genet.* 2:93-98, 1988; y Yang, *Proc. Natl. Acad. Sci. USA* 93:14972-14977, 1996.

15 Los promotores específicos de tejido adecuados incluyen, pero no se limitan a, promotores específicos de hojas tales como los descritos, por ejemplo, por Yamamoto y otros, *Plant J.* 12:255-265, 1997; Kwon y otros, *Plant Physiol.* 105:357-67, 1994; Yamamoto y otros, *Plant Cell Physiol.* 35:773-778, 1994; Gotor y otros, *Plant J.* 3:509-18, 1993; Orozco y otros, *Plant Mol. Biol.* 23:1129-1138, 1993; y Matsuoka y otros, *Proc.*
20 *Natl. Acad. Sci. USA* 90:9586-9590, 1993.

El constructo de ácido nucleico de la presente invención también puede incluir al menos un marcador seleccionable tal como, por ejemplo, nptII. Preferentemente, el constructo de ácido nucleico es un vector lanzadera, que puede propagarse tanto en *E. coli* (en donde el constructo comprende un marcador seleccionable apropiado y un origen de
25 replicación) como ser compatible para su propagación en células. El constructo de acuerdo con la presente invención puede ser, por ejemplo, un plásmido, un bácmido, un fagémido, un cósmido, un fago, un virus o un cromosoma artificial, preferiblemente un plásmido.

El constructo de ácido nucleico de la presente invención puede utilizarse para
30 transformar de forma estable células de plátano. Los métodos principales para provocar la integración estable del ADN exógeno en el genoma del plátano incluyen dos enfoques principales:

(i) Transferencia de genes mediada por *Agrobacterium*: Klee y otros, (1987) *Annu. Rev. Plant Physiol.* 38:467-486; Klee y Rogers en *Cell Culture and Somatic Cell Genetics of*
35 *Plants*, Vol. 6, *Molecular Biology of Plant Nuclear Genes*, eds. Schell, J. y Vasil, L. K.,

- Academic Publishers, San Diego, Calif. (1989) p. 2-25; Gatenby, en *Plant Biotechnology*, eds. Kung, S. y Arntzen, C. J., Butterworth Publishers, Boston, Mass. (1989) p. 93-112.
- (ii) Captación directa de ADN: Paszkowski y otros, en *Cell Culture and Somatic Cell Genetics of Plants*, Vol. 6, *Molecular Biology of Plant Nuclear Genes* eds. Schell, J. y Vasil, L. K., Academic Publishers, San Diego, Calif. (1989) p. 52-68; lo que incluye métodos para la captación directa de ADN en protoplastos, Toriyama, K. y otros, (1988) *Bio/Technology* 6:1072-1074. DNA uptake induced by brief electric shock of plant cells: Zhang y otros, *Plant Cell Rep.* (1988) 7:379-384. Fromm y otros, *Nature* (1986) 319:791-793. DNA injection into plant cells or tissues by particle bombardment, Klein y otros, *Bio/Technology* (1988) 6:559-563; McCabe y otros, *Bio/Technology* (1988) 6:923-926; Sanford, *Physiol. Plant.* (1990) 79:206-209; mediante el uso de sistemas de micropipetas: Neuhaus y otros, *Theor. Apl. Genet.* (1987) 75:30-36; Neuhaus y Spangenberg, *Physiol. Plant.* (1990) 79:213-217; fibras de vidrio o transformación de cultivos celulares por fibras de carburo de silicio, embriones o tejido de callos, Patente de los Estados Unidos 5,464,765 o por incubación directa de ADN con polen en germinación, DeWet y otros, en *Experimental Manipulation of Ovule Tissue*, eds. Chapman, G. P. y Mantell, S. H. y Daniels, W. Longman, London, (1985) p. 197-209; y Ohta, *Proc. Natl. Acad. Sci. USA* (1986) 83:715-719.
- El sistema *Agrobacterium* incluye el uso de vectores plasmídicos que contienen segmentos de ADN definidos que se integran en el ADN genómico de la planta. Los métodos de inoculación del tejido vegetal varían en dependencia de la especie vegetal y del sistema de administración de *Agrobacterium*. Un enfoque que se utiliza ampliamente es el procedimiento del disco de la hoja, que puede realizarse con cualquier explante de tejido que proporcione una buena fuente para el inicio de la diferenciación de la planta completa. Horsch y otros, en *Plant Molecular Biology Manual* A5, Kluwer Academic Publishers, Dordrecht (1988) p. 1-9. Un enfoque complementario emplea el sistema de administración de *Agrobacterium* en combinación con la infiltración al vacío. Los procedimientos adecuados mediados por *Agrobacterium* para introducir ADN exógeno en las células de plátano se describen por Dougale y otros. (*Journal of General Virology*, 79:2301-2311, 1998) y en la Patente de los Estados Unidos núm. 6,395,962.
- Existen diversos métodos de transferencia directa de ADN a las células vegetales. En la electroporación, los protoplastos se exponen brevemente a un fuerte campo eléctrico. En la microinyección, el ADN se inyecta mecánicamente directamente en las células mediante el uso de micropipetas muy pequeñas. En el bombardeo de micropartículas,

el ADN se adsorbe en microprojectiles, tales como cristales de sulfato de magnesio o partículas de tungsteno, y los microprojectiles se aceleran físicamente hacia las células o tejidos vegetales.

De manera alternativa, el constructo de ácido nucleico de la presente invención puede introducirse en células de plátano mediante un bombardeo de microprojectiles. En esta técnica, las partículas de tungsteno u oro recubiertas con ADN exógeno se aceleran hacia las células objetivo. Se describen procedimientos adecuados de transformación de plátano mediante el bombardeo de microprojectiles por Sagi y otros. (Biotechnology 13:481-485, 1995) y por Dougale y otros, (Journal of General Virology, 79:2301-2311, 1998). Preferentemente, el constructo de ácido nucleico de la presente invención se introduce en células de plátano mediante un procedimiento de bombardeo de microprojectiles como se describe en el Ejemplo 4 en la presente descripción más abajo.

Después de la transformación, las células transformadas se micropropagan para proporcionar una reproducción rápida y consistente del material transformado.

La micropropagación es un proceso de cultivo de plantas de nueva generación a partir de un solo pedazo de tejido que se extirpa de una planta o cultivo progenitor seleccionado. Este proceso permite la reproducción masiva de plantas que tienen el tejido preferido que expresa la proteína de fusión. Las plantas de nueva generación que se producen son genéticamente idénticas y tienen todas las características de la planta original. La micropropagación permite la producción masiva de material vegetal de calidad en un corto período de tiempo y ofrece una rápida multiplicación de cultivares seleccionados en la preservación de las características de la planta transgénica o transformada original. Las ventajas de la clonación de plantas son la velocidad de multiplicación de las plantas y la calidad y uniformidad de las plantas producidas.

La micropropagación es un procedimiento de múltiples etapas que requiere la alteración del medio de cultivo o las condiciones de crecimiento entre etapas. Por tanto, el proceso de micropropagación implica cuatro etapas básicas: etapa uno, cultivo de tejidos inicial; etapa dos, multiplicación de cultivos de tejidos; etapa tres, diferenciación y formación de la planta y etapa cuatro, cultivo en invernadero y endurecimiento. Durante la etapa uno, cultivo de tejido inicial, el cultivo de tejido se establece y se certifica como libre de contaminantes. Durante la etapa dos, el cultivo de tejido inicial se multiplica hasta que se produce una cantidad suficiente de muestras de tejido para cumplir los objetivos de producción. Durante la etapa tres, las muestras de tejido cultivadas en la etapa dos se dividen y crecen en plántulas individuales. En la etapa cuatro, las plántulas

transformadas se transfieren a un invernadero para su endurecimiento, donde se aumenta gradualmente la tolerancia de las plantas a la luz para poder crecer en el medio natural.

Por tanto, las células de plátano transformadas pueden micropropagarse y regenerarse en plantas mediante el uso de métodos conocidos en la técnica tales como los descritos, por ejemplo, en la Patente de los Estados Unidos núm. 6,133,035 y de Novak y otros, 1989; Dhed'a y otros, 1991; Cote y otros, 1996; Becker y otros, 2000; Sagi y otros, Plant Cell Reports 13:262-266, 1994; Grapin y otros, Cell Dev. Biol. Plant. 32:66-71, 1996; Marroquin y otros, In Vivo Cell. Div. Biol. 29P:43-46, 1993; y Escalant y otros, In Vivo Cell Dev. Biol. 30:181-186, 1994).

La integración estable de la secuencia de ADN exógeno en el genoma de las plantas transformadas puede determinarse mediante el uso de técnicas de biología molecular estándar bien conocidas en la técnica tales como PCR e hibridación por transferencia de Southern.

Aunque actualmente se prefiere la transformación estable, la presente invención también contempla la transformación transitoria de células cultivadas, células foliares, células meristemáticas o la planta completa.

La transformación transitoria puede efectuarse mediante cualquiera de los métodos de transferencia directa de ADN descritos anteriormente o mediante infección viral mediante el uso de virus vegetales modificados.

Se prefiere la infección viral ya que permite evitar la micropropagación y la regeneración de una planta completa a partir de células cultivadas. Los virus que han demostrado ser útiles para la transformación de plantas huéspedes incluyen CaMV, TMV y BV. La transformación de plantas mediante el uso de virus de plantas se describe en la Patente de los Estados Unidos núm. 4,855,237 (BGV), EP-A 67,553 (TMV), Solicitud Publicada en Japonés núm. 63-14693 (TMV), EPA 194,809 (BV), EPA 278,667 (BV); y Gluzman y otros. (Communications in Molecular Biology: Viral Vectors, Cold Spring Harbor Laboratory, New York, pp. 172-189, 1988). En el documento WO 87/06261 se describen partículas de pseudovirus para su uso en la expresión de ADN extraño en muchos huéspedes, lo que incluye plantas.

La construcción de virus de ARN de plantas para la introducción y expresión de secuencias de ácidos nucleicos exógenos no virales en plantas está demostrada por las referencias anteriores así como también por Dawson y otros. (Virology 172:285-292, 1989; Takamatsu y otros, EMBO J. 6:307-311, 1987; French y otros, (Science 231:1294-1297, 1986); y Takamatsu y otros, (FEBS Letters 269:73-76, 1990).

Cuando el virus es un virus de ADN, pueden realizarse las modificaciones adecuadas en el propio virus. De manera alternativa, el virus puede clonarse primero en un plásmido bacteriano para facilitar la construcción del vector viral deseado con el ADN extraño. A continuación, el virus puede extirparse del plásmido. Si el virus es un virus de ADN,

5 puede unirse un origen de replicación bacteriano al ADN viral, que después se replica por la bacteria. La transcripción y traducción de este ADN producirá la proteína de la cubierta que encapsulará el ADN viral.

Si el virus es un virus de ARN, el virus generalmente se clona como un ADNc y se inserta en un plásmido. A continuación, el plásmido se utiliza para crear todos los constructos.

10 El virus de ARN se produce a continuación mediante la transcripción de la secuencia viral del plásmido y la traducción de los genes virales para producir la(s) proteína(s) de cubierta que encapsulan el ARN viral.

La construcción de los virus de ARN de plantas para la introducción y expresión en plantas de secuencias de ácidos nucleicos exógenos no virales, tales como las incluidas

15 en el constructo de la presente invención, se demuestra mediante las referencias anteriores, así como también en la Patente de los Estados Unidos núm. 5,316,931.

En una modalidad, se proporciona un ácido nucleico viral vegetal en el que se suprimió la secuencia codificante de la proteína de cubierta nativa de un ácido nucleico viral, se insertó una secuencia codificante de la proteína de cubierta viral vegetal no nativa y un

20 promotor no nativo, preferentemente, el promotor subgenómico de la secuencia codificante de la proteína de cubierta no nativa, capaz de expresarse en el huésped vegetal, de empaquetar el ácido nucleico viral vegetal recombinante y de garantizar una infección sistémica del huésped por el ácido nucleico viral vegetal recombinante. De manera alternativa, el gen de la proteína de la cubierta puede inactivarse mediante la

25 inserción de la secuencia de ácido nucleico no nativa dentro de él, de manera que se produzca una proteína. El ácido nucleico viral vegetal recombinante puede contener uno o más promotores subgenómicos no nativos adicionales. Cada promotor subgenómico no nativo es capaz de transcribir o expresar genes o secuencias de ácido nucleico adyacentes en el huésped vegetal y es incapaz de recombinarse entre sí y con

30 promotores subgenómicos nativos. Pueden insertarse secuencias de ácido nucleico no nativas (extrañas) adyacentes al promotor subgenómico viral vegetal nativo o los promotores subgenómicos virales vegetales nativos y no nativos si se incluye más de una secuencia de ácido nucleico. Las secuencias de ácido nucleico no nativas se transcriben o expresan en la planta huésped bajo el control del promotor subgenómico

35 para producir los productos deseados.

En una segunda modalidad, se proporciona un ácido nucleico viral vegetal recombinante como en la primera modalidad, excepto que la secuencia codificante de la proteína de cubierta nativa se coloca adyacente a uno de los promotores subgenómicos de la proteína de cubierta no nativa en lugar de una secuencia codificante de la proteína de cubierta no nativa.

En una tercera modalidad, se proporciona un ácido nucleico viral vegetal recombinante en el que el gen de la proteína de cubierta nativa es adyacente a su promotor subgenómico y se han insertado uno o más promotores subgenómicos no nativos en el ácido nucleico viral. Los promotores subgenómicos no nativos insertados son capaces de transcribir o expresar genes adyacentes en una planta huésped y son incapaces de recombinarse entre sí y con promotores subgenómicos nativos. Las secuencias de ácido nucleico no nativas pueden insertarse junto a los promotores virales subgenómicos no nativos de la planta, de manera que las secuencias se transcriban o expresen en la planta huésped bajo el control de los promotores subgenómicos para producir el producto deseado.

En una cuarta modalidad, se proporciona un ácido nucleico viral vegetal recombinante como en la tercera modalidad, excepto que la secuencia codificante de la proteína de cubierta nativa se reemplaza por una secuencia codificante de la proteína de cubierta no nativa.

Los vectores virales son encapsulados por las proteínas de la cubierta codificadas por el ácido nucleico viral vegetal recombinante para producir un virus vegetal recombinante. El ácido nucleico viral vegetal recombinante o el virus vegetal recombinante se usa para infectar las plantas huésped apropiadas. El ácido nucleico viral vegetal recombinante es capaz de replicarse en el huésped, diseminarse sistémicamente en el mismo y transcribir o expresar los genes extraños (ácido nucleico aislado) en el huésped para producir la proteína deseada.

Además de lo anterior, la molécula de ácido nucleico de la presente invención también puede introducirse en un genoma de un cloroplasto lo que permite de esta manera la expresión del cloroplasto.

Se conoce una técnica para introducir secuencias de ácido nucleico exógenas en el genoma de los cloroplastos. Esta técnica implica los siguientes procedimientos. Primero, las células vegetales se tratan químicamente para reducir el número de cloroplastos por célula a aproximadamente uno. A continuación, el ácido nucleico exógeno se introduce mediante bombardeo de partículas en las células con el objetivo de introducir al menos una molécula de ácido nucleico exógeno en los cloroplastos. El ácido nucleico exógeno

se selecciona de manera que sea integrable en el genoma del cloroplasto mediante recombinación homóloga que se efectúa fácilmente mediante enzimas inherentes al cloroplasto. Para este propósito, el ácido nucleico exógeno incluye, además de un gen de interés, al menos un tramo de ácido nucleico que se deriva del genoma del cloroplasto. Además, el ácido nucleico exógeno incluye un marcador seleccionable, que sirve mediante procedimientos de selección secuencial para verificar que todas o sustancialmente todas las copias de los genomas de los cloroplastos después de dicha selección incluirán el ácido nucleico exógeno. Pueden encontrarse más detalles en relación con esta técnica en las Patentes de los Estados Unidos núms. 4,945,050; y 5,693,507 que se incorporan en la presente descripción como referencia. Por tanto, un polipéptido puede producirse por el sistema de expresión de proteínas del cloroplasto e integrarse en la membrana interna del cloroplasto.

En caso de que el polipéptido exógeno confiera resistencia a la enfermedad a la planta, la expresión puede determinarse basado en el aumento de la resistencia o tolerancia a patógenos, preferentemente, en comparación con una planta similar de tipo salvaje (no transformada). La evaluación comparativa de plantas para determinar su resistencia o tolerancia a patógenos puede realizarse mediante el uso de bioensayos in vitro o in vivo bien conocidos en la técnica de patología vegetal tales como los descritos, por ejemplo, por Agrios, G. N., ed. (Plant Pathology, Tercera Edición, Academic Press, Nueva York, 1988).

La evaluación de la resistencia o tolerancia de las plantas a los patógenos puede realizarse mediante la exposición de un patógeno a un extracto obtenido de tejido vegetal y la determinación del efecto del extracto en el crecimiento del patógeno in vitro. En algunas modalidades, la evaluación de la resistencia o tolerancia de las plantas a los patógenos se efectúa mediante la exposición de un patógeno a un tejido vegetal (por ejemplo, un tejido foliar).

En otras modalidades, la evaluación de la resistencia o tolerancia de las plantas los a patógenos se efectúa mediante la exposición de un patógeno a una planta completa. Por ejemplo, la evaluación de la resistencia o tolerancia de las plantas a *Fusarium oxysporum f. sp. Cubense* (Foc) (el agente causal del mal de Panamá) puede afectarse al plantar plantas de plátano transformadas en un campo abierto cerca de plantas no transformadas que están infectadas con el patógeno (usadas como fuente de inóculo). La gravedad de la enfermedad que se desarrolla posteriormente en plantas transformadas se evalúa comparativamente con las plantas no transformadas. La gravedad de la enfermedad se evalúa, preferentemente, de manera visual (el daño

generalmente aparece en los retoños que tienen al menos 5-12 hojas) y se analiza estadísticamente para determinar diferencias significativas en la resistencia o tolerancia entre líneas de plantas al mal de Panamá.

Por tanto, la presente invención proporciona constructos de ácido nucleico que incluyen uno o más polinucleótidos que codifican polipéptidos de resistencia a la enfermedad, células de plátano transformadas y plantas de plátano transformadas que expresan rasgos de resistencia a la enfermedad exógenas, y métodos para producirlos.

VI. Métodos de mejoramiento

Poblaciones de polinización abierta. La mejora de las poblaciones de polinización abierta de cultivos tales como el centeno, muchos maíces y la remolacha azucarera, pastos herbáceos, legumbres tales como la alfalfa y el trébol, y cultivos de árboles tropicales tales como el cacao, los cocos, la palma aceitera y algo de caucho, depende esencialmente de cambios de las frecuencias génicas hacia la fijación de alelos favorables a la vez que se mantiene un alto (pero lejos del máximo) grado de heterocigosidad. La uniformidad en dichas poblaciones es imposible y la correspondencia con el tipo en una variedad de polinización abierta es una característica estadística de la población en su conjunto, no una característica de plantas individuales. Por tanto, la heterogeneidad de las poblaciones de polinización abierta contrasta con la homogeneidad (o prácticamente) de las líneas, los clones y los híbridos endogámicos.

Los métodos de mejora de la población se dividen naturalmente en dos grupos, los que se basan en la selección puramente fenotípica, normalmente denominada selección en masa, y los que se basan en la selección con pruebas de progenie. La mejora de la interpoblacional utiliza el concepto de poblaciones de mejoramiento abiertas; lo que permite que los genes fluyan de una población a otra. Las plantas de una población (cultivar, cepa, ecotipo o cualquier fuente de germoplasma) se cruzan de forma natural (por ejemplo, por el viento) o a mano o por abejas (comúnmente *Apis mellifera* L. o *Megachile rotundata* F.) con plantas de otras poblaciones. La selección se aplica para mejorar una (o a veces ambas) poblaciones mediante el aislamiento de plantas con rasgos convenientes de ambas fuentes.

Existen básicamente dos métodos principales de mejora de la población de polinización abierta. En primer lugar, está la situación en la que una población cambia en masa mediante un procedimiento de selección elegido. El resultado es una población mejorada que se propaga indefinidamente mediante el apareamiento aleatorio dentro de sí misma de forma aislada. En segundo lugar, la variedad sintética logra el mismo resultado final que es el mejoramiento de la población, pero no es en sí misma

propagable como tal; tiene que reconstruirse a partir de líneas parentales o clones. Estos procedimientos de mejoramiento de plantas para mejorar las poblaciones de polinización abierta se conocen bien por los expertos en la técnica, y numerosos textos y artículos proporcionan revisiones exhaustivas de los procedimientos de mejoramiento

5 que se utilizan habitualmente para mejorar las plantas de polinización cruzada, lo que incluye: Allard, *Principles of Plant Breeding*, John Wiley & Sons, Inc. (1960); Simmonds, *Principles of Crop Improvement*, Longman Group Limited (1979); Hallauer y Miranda, *Genética cuantitativa en el mejoramiento de maíz*, Iowa State University Press (1981); y Jensen, *Plant Breeding Methodology*, John Wiley & Sons, Inc. (1988). Para conocer

10 los métodos de mejora de la población específicos para la soja, ver, por ejemplo, J.R. Wilcox, editor (1987) *SOYBEANS: Improvement, Production, and Uses*, Segunda edición, American Society of Agronomy, Inc., Crop Science Society of America, Inc., y Soil Science Society of America, Inc., editores, 888 páginas.

Selección Masiva. En la selección masiva, se eligen las plantas individuales

15 convenientes, se cosechan y la semilla se compone sin someter a prueba la progenie para producir la siguiente generación. Dado que la selección se basa únicamente en el progenitor materno y no existe control sobre la polinización, la selección masiva equivale a una forma de apareamiento aleatoria con selección. Como se indicó anteriormente, el propósito de la selección masiva es aumentar la proporción de genotipos superiores en

20 la población.

Sintéticos. Una variedad sintética se produce mediante el cruzamiento entre sí de un número de genotipos que se seleccionan por su buena capacidad de combinación en todas las combinaciones híbridas posibles, con el subsiguiente mantenimiento de la variedad mediante polinización abierta. Si los progenitores son líneas (más o menos

25 endogámicas) propagadas por semillas, como en algunas remolachas azucareras y frijoles (*Vicia*) o clones, como en pastos herbáceos, tréboles y alfalfa, no hay diferencia en principio. Los progenitores se seleccionan en función de la capacidad de combinación general, a veces mediante cruzamientos de prueba o cruces superiores, más generalmente mediante cruzamientos múltiples. Las líneas de semillas parentales

30 pueden ser endogámicas deliberadamente (por ejemplo, mediante autofecundación o cruce de hermanos). Sin embargo, incluso si los progenitores no son endogámicos deliberadamente, la selección dentro de las líneas durante el mantenimiento de la línea asegurará que se produzca cierta endogamicidad. Los progenitores clonales, por supuesto, permanecerán inalterados y muy heterocigotos.

- Si una semilla sintética puede pasar directamente de la parcela de producción de semillas parentales al agricultor o deba someterse primero a uno o dos ciclos de multiplicación, depende de la producción de semillas y de la escala de la demanda de semillas. En la práctica, los pastos y los tréboles generalmente se multiplican una o dos veces y, por tanto, se apartan considerablemente del sintético original.
- Si bien a veces se utiliza la selección masiva, generalmente se prefiere someter a prueba la progenie para los cruzamientos múltiples, debido a su simplicidad operativa y su relevancia obvia para el objetivo, es decir, la explotación de la capacidad de combinación general en un sintético.
- El número de líneas parentales o clones que entran en un sintético varía ampliamente. En la práctica, el número de líneas parentales varía de 10 a varios cientos, donde 100-200 es la media. Se espera que los sintéticos de base amplia que se forman a partir de 100 o más clones sean más estables durante la multiplicación de semillas que los sintéticos de base estrecha.
- Híbridos. Como se analizó anteriormente, el híbrido es una planta individual resultante de un cruzamiento entre progenitores de diferentes genotipos. Los híbridos comerciales se usan ahora ampliamente en muchos cultivos, lo que incluye el maíz (maíz), el sorgo, la remolacha azucarera, el girasol y el brócoli. Los híbridos pueden formarse de varias formas diferentes, entre ellas por el cruzamiento de dos progenitores directamente (híbridos de un solo cruzamiento), al cruzar un híbrido de un solo cruzamiento con otro progenitor (híbridos de cruzamiento triple o de tres vías), o al cruzar dos híbridos diferentes (híbridos de cruzamiento doble o de cuatro vías).
- En sentido estricto, la mayoría de los individuos de una población exogámica (es decir, de polinización abierta) son híbridos, pero el término generalmente se reserva para los casos en los que los progenitores son individuos cuyos genomas son lo suficientemente distintos como para ser reconocidos como especies o subespecies diferentes. Los híbridos pueden ser fértiles o estériles en dependencia de las diferencias cualitativas y/o cuantitativas en los genomas de los dos progenitores. La heterosis, o vigor híbrido, generalmente se asocia con una mayor heterocigosidad que da como resultado un mayor vigor de crecimiento, supervivencia y fertilidad de los híbridos en comparación con las líneas parentales que se usaron para formar el híbrido. La heterosis máxima generalmente se logra al cruzar dos líneas genéticamente diferentes y altamente endogámicas.
- La producción de híbridos es una industria bien desarrollada, que involucra la producción aislada tanto de las líneas parentales como de los híbridos resultantes del

cruzamiento de esas líneas. Para un análisis detallado del proceso de producción de híbridos, ver, por ejemplo, Wright, *Commercial Hybrid Seed Production* 8:161-176, en Hybridization of Crop Plants.

Análisis de segregación a granel (BSA). BSA, también conocido como análisis de segregación por agrupamiento, o análisis de agrupamiento segregante, es un método descrito por Michelmore y otros. (Michelmore y otros, 1991, Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proceedings of the National Academy of Sciences, USA* 99:9828-9832) y Quarrie y otros. (Quarrie y otros, Bulk segregant analysis with molecular markers and its use for improving drought resistance in maize, 1999, *Journal of Experimental Botany*, 50(337):1299-1306).

Para BSA de un rasgo de interés, se eligen líneas parentales con determinados fenotipos diferentes y se cruzan para generar poblaciones F2, doble haploides o recombinantes endogámicos con análisis de QTL. A continuación, la población se fenotipa para identificar plantas individuales o líneas que tienen una expresión alta o baja del rasgo. Se preparan dos agrupamientos de ADN, uno de los individuos que tienen un fenotipo (por ejemplo, resistente a patógenos) y el otro de los individuos que tienen el fenotipo inverso (por ejemplo susceptibles a patógenos), y se analizan para determinar la frecuencia alélica con marcadores moleculares. Solo se requieren unos pocos individuos en cada agrupamiento (por ejemplo, 10 plantas de cada uno) si los marcadores son dominantes (por ejemplo, RAPDs). Se necesitan más individuos cuando los marcadores son codominantes (por ejemplo, RFLPs). Los marcadores vinculados al fenotipo pueden identificarse y utilizarse para el mejoramiento o el mapeo de QTL.

Piramidación de genes. El método para combinar en un solo genotipo una serie de genes objetivo identificados en diferentes progenitores se suele denominar piramidación de genes. La primera parte de un mejoramiento por piramidación de genes se denomina un pedigrí y tiene como objetivo acumular una copia de todos los genes objetivo en un solo genotipo (llamado genotipo raíz). La segunda parte se denomina etapas de fijación y tiene como objetivo fijar los genes objetivo en un estado homocigoto, es decir, derivar el genotipo ideal (ideotipo) a partir del genotipo raíz. La piramidación de genes puede combinarse con la selección asistida por marcadores (MAS, ver Hospital y otros, 1992, 1997a y 1997b, y Moreau y otros, 1998) o la selección recurrente basada en marcadores (MBRS, ver Hospital y otros, 2000).

Los programas de mejoramiento de plátanos, especialmente para los plátanos comestibles, se ven obstaculizados por la alta esterilidad, la triploidía y la falta de semillas. Pocos clones de plátano diploides producen polen viable, y el germoplasma de los clones comerciales de plátano es estéril tanto para masculinos como para
5 femeninos. A pesar de estos problemas y desafíos, se lograron importantes avances en el mejoramiento genético de Musa en los últimos años, y los programas de mejoramiento del plátano no dejan de ofrecer nuevas variedades (Escalant y Jain, Chapter 30, Banana improvement with cellular and molecular biology, and induced mutations: future and perspectives, 8 páginas, En Jain y Swennan, editores, Banana Improvement: Cellular,
10 Molecular Biology, and Induced Mutations, 2004, Food and Agriculture Organization of the United Nations, Science Publishers, Inc.).

Para obtener información sobre el mejoramiento de plátanos, ver, por ejemplo, Heslop-Harrison y Schwarzacher, Annals of Botany 100:1073-1084, 2007; Bakry y otros, Capítulo 1, Genetic Improvement in Banana, 50 páginas, En Breeding Plantation Tree
15 Crops: Tropical Species, 2009; Heslop-Harrison y otros, Genomics, Banana Breeding and Superdomestication, Acta Hort. 897:55-62, 2011; Jenny y otros, En Jacome y otros, editores, *Mycosphaerella* leaf spot diseases of banana: present status and outlook, Proceedings of the 2nd International Workshop on *Mycosphaerella* leaf spot diseases
celebrado en San José, Costa Rica, 20-23 de mayo de 2002, sesión 4, páginas 199-
20 208; Ortiz y otros, Banana and Plantain Breeding, Chapter 10, páginas 110-146, En Gowen y otros, editors, Bananas and Plantains, World Crop Series, Springer Link, 1995; Batte y otros, Frontiers in Plant Science, Vol. 10, artículo 81, 9 páginas, Febrero de 2019.

VII. Edición de Genes

Como se usa en la presente descripción, el término "sistema de edición de genes" se
25 refiere a un sistema que comprende uno o más dominios o componentes de unión al ADN y uno o más dominios o componentes modificadores del ADN, o ácidos nucleicos aislados, por ejemplo, uno o más vectores, que codifican dicho dominios o componentes de unión y modificación del ADN. Los sistemas de edición de genes se usan para modificar el ácido nucleico de un gen objetivo y/o para modular la expresión de un gen
30 objetivo. En los sistemas de edición de genes conocidos, por ejemplo, los uno o más dominios o componentes de unión al ADN están asociados con los uno o más dominios o componentes modificadores del ADN, de modo que los uno o más dominios de unión al ADN dirigen los uno o más dominios o componentes modificadores del ADN a un sitio de ácido nucleico específico. Los métodos y composiciones para potenciar la edición de
35 genes se conocen bien en la técnica. Ver, por ejemplo, la Publicación de Solicitud de

Patente de los Estados Unidos núm. 2018/0245065, que se incorpora como referencia en su totalidad.

Determinados sistemas de edición de genes se conocen en la técnica, e incluyen, pero no se limitan a, nucleasas con dedos de zinc, nucleasas efectoras de tipo activador de la transcripción (TALEN); sistemas de repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas (CRISPR)/Cas, sistemas de meganucleasa y edición de genes mediada por vectores virales.

En algunas modalidades, la presente descripción enseña métodos para la edición/clonación de genes mediante el uso de ADN nucleasas. Complejos CRISPR, nucleasas efectoras de tipo activador de la transcripción (TALEN), nucleasas con dedos de zinc (ZFN) y enzimas de restricción FokI, que son algunas de las nucleasas específicas de secuencia que se utilizan como herramientas de edición de genes. Estas enzimas pueden dirigir sus actividades de nucleasa a los loci objetivo deseados a través de interacciones con regiones guía que se manipulan genéticamente para reconocer secuencias de interés. En algunas modalidades, la presente descripción enseña métodos de edición de genes basados en CRISPR para manipular genéticamente el genoma de las especies de plátano de la presente descripción para estimular, mejorar o modular la resistencia a la enfermedad de los patógenos.

(i) Sistemas CRISPR

CRISPR (repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas) y las endonucleasas asociadas a CRISPR (cas) se descubrieron originalmente como sistemas de inmunidad adaptativa desarrollados por bacterias y arqueas para protegerse contra la invasión viral y de plásmidos. Los sistemas CRISPR/Cas naturales en bacterias se componen de uno o más genes Cas y una o más matrices CRISPR que consisten en repeticiones palindrómicas cortas de secuencias de bases separadas por secuencias de direccionamiento del genoma adquiridas de virus y plásmidos previamente encontrados (denominados espaciadores). (Wiedenheft, B. y otros, Nature. 2012; 482:331; Bhaya, D. y otros, Annu. Rev. Genet. 2011; 45:231; y Terms, M.P. y otros, Curr. Opin. Microbiol. 2011; 14:321). Las bacterias y arqueas que poseen uno o más loci CRISPR responden al desafío viral o plasmídico al integrar fragmentos cortos de secuencia extraña (protoespaciadores) en el cromosoma del huésped en el extremo proximal de la matriz CRISPR. La transcripción de los loci CRISPR genera una biblioteca de ARN derivados de CRISPR (ARNcr) que contienen secuencias complementarias a los ácidos nucleicos invasores encontrados previamente (Haurwitz, RE, y otros, Science. 2012:329;1355; Gesner, E.M., y otros, Nat. Struct. Mol. Biol. 2001:18,688;

Jinek, M., y otros, *Science*. 2012:337; 816-21). El reconocimiento del objetivo por los ARNcr se produce a través del emparejamiento de bases complementarias con el ADN objetivo, lo que dirige la escisión de secuencias extrañas por medio de proteínas Cas. (Jinek y otros, 2012 "A Programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity." *Science*. 2012:337; 816-821).

Existen al menos cinco tipos principales de sistemas CRISPR (Tipo I, II, III, IV y V) y al menos 16 subtipos distintos (Makarova, K.S., y otros, *Nat Rev Microbiol*. 2015. *Nat. Rev. Microbiol.* 13, 722–736). Los sistemas CRISPR también se clasifican basado en sus proteínas efectoras. Los sistemas de Clase 1 poseen complejos de ARNcr-efector de múltiples subunidades, mientras que en los sistemas de Clase 2 todas las funciones del complejo efector son llevadas a cabo por una única proteína (por ejemplo, Cas9 o Cpf1). En algunas modalidades, la presente descripción proporciona el uso de sistemas efectores de subunidad única de tipo II y/o tipo V.

Como estos se producen naturalmente en muchos tipos diferentes de bacterias, la disposición exacta de CRISPR y la estructura, función y número de genes Cas y su producto difieren algo de una especie a otra. Haft y otros (2005) *PLoS Comput. Biol.* 1: e60; Kunitz y otros (2007) *Genome Biol.* 8: R61; Mojica y otros (2005) *J. Mol. Evol.* 60: 174-182; Bolotin y otros (2005) *Microbiol.* 151: 2551-2561; Pourcel y otros (2005) *Microbiol.* 151: 653-663; y Stern y otros (2010) *Trends. Genet.* 28: 335-340. Por ejemplo, las proteínas Cse (subtipo Cas, *E. coli*) (por ejemplo, CasA) forman un complejo funcional, Cascade, que procesa los transcritos de ARN de CRISPR en unidades espaciadoras repetidas que Cascade retiene. Brouns y otros (2008) *Science* 321: 960-964. En otros procariotas, Cas6 procesa el transcrito de CRISPR. La inactivación de fagos basada en CRISPR en *E. coli* requiere Cascade y Cas3, pero no Cas1 o Cas2. Las proteínas Cmr (módulo Cas RAMP) en *Pyrococcus furiosus* y otros procariotas forman un complejo funcional con pequeños ARN de CRISPR que reconocen y escinden los ARN objetivo complementarios. Un sistema CRISPR más simple se basa en la proteína Cas9, que es una nucleasa con dos sitios de corte activos, uno para cada cadena de la doble hélice. La combinación de Cas9 y el ARN del locus CRISPR modificado puede usarse en un sistema para la edición de genes. Pennisi (2013) *Science* 341: 833-836.

(ii) CRISPR/Cas9

En algunas modalidades, la presente descripción proporciona métodos de edición de genes mediante el uso de un sistema CRISPR de Tipo II. Los sistemas de tipo II se basan en una i) proteína endonucleasa única, ii) un ARNcr transactivo (ARNtracr), y iii)

un ARNcr donde una porción de ~20 nucleótidos (nt) del extremo 5' del ARNcr es complementaria a un ácido nucleico objetivo. La región de una cadena de ARNcr de CRISPR que es complementaria a su protoespaciador de ADN objetivo se denomina aquí "secuencia guía".

- 5 En algunas modalidades, los componentes ARNtracr y ARNcr de un sistema de Tipo II pueden reemplazarse por un ARN guía único (ARNsg), también conocido como ARN guía (ARNg). El ARNsg puede incluir, por ejemplo, una secuencia de nucleótidos que comprende una secuencia de al menos 12-20 nucleótidos complementaria a la secuencia de ADN objetivo (secuencia guía) y puede incluir una secuencia de ARN de andamio común en su extremo 3'. Como se usa en la presente descripción, "un ARN de andamio común" se refiere a cualquier secuencia de ARN que imita la secuencia de ARNtracr o cualquier secuencia de ARN que funcione como ARNtracr.

- 10 Las endonucleasas Cas9 producen rupturas de ADN de extremos romos, y se reclutan para el ADN objetivo mediante una combinación de oligos de ARNcr y ARNtracr, que unen la endonucleasa a través de la hibridación complementaria del complejo ARN CRISPR.

- 15 En algunas modalidades, el reconocimiento del ADN por el complejo ARNcr/endonucleasa requiere el apareamiento adicional de bases complementarias con un motivo adyacente protoespaciador (PAM) (por ejemplo, 5'-NGG-3') situado en una porción 3' del ADN objetivo, aguas abajo del protoespaciador objetivo. (Jinek, M. y otros, Science. 2012, 337:816-821). En algunas modalidades, el motivo PAM reconocido por un Cas9 varía para diferentes proteínas Cas9.

- 20 En algunas modalidades, el Cas9 descrito en la presente descripción puede ser cualquier variante derivada o aislada de cualquier fuente. En otras modalidades, el péptido Cas9 de la presente descripción puede incluir una o más de las mutaciones descritas en la bibliografía, lo que incluye, pero no se limita a las mutaciones funcionales descritas en: Fonfara y otros, Nucleic Acids Res. Febrero de 2014;42(4):2577-90; Nishimasu H. y otros, Cell. 27 de febrero de 2014, 156(5):935-49; Jinek M. y otros Science. 2012 337:816-21; y Jinek M. y otros, Science. 14 de marzo de 2014, 343 (6176); ver también la Solicitud de Patente de los Estados Unidos núm. 13/842,859, presentada el 15 de marzo de 2013, que se incorpora aquí como referencia; además, ver las Patentes de los Estados Unidos núms. 8,697,359; 8,771,945; 8,795,965; 8,865,406; 8,871,445; 8,889,356; 8,895,308; 8,906,616; 8,932,814; 8,945,839; 8,993,233; y 8,999,641, que se incorporan aquí por la presente como referencia. Por 30 tanto, en algunas modalidades, los sistemas y métodos descritos en la presente

descripción pueden usarse con la proteína Cas9 de tipo silvestre que tiene actividad nucleasa bicatenaria, mutantes Cas9 que actúan como nicasas monocatenarias, u otros mutantes con actividad nucleasa modificada.

De acuerdo con la presente descripción, las moléculas Cas9 de, derivadas, o basadas

- 5 en las proteínas Cas9 de una variedad de especies pueden usarse en los métodos y composiciones descritos en la presente descripción. Por ejemplo, las moléculas Cas9 de, derivadas o basadas en, por ejemplo, las moléculas Cas9 de *S. pyogenes*, *S. thermophilus*, *Staphylococcus aureus* y/o *Neisseria meningitidis*, pueden usarse en los sistemas, métodos y composiciones descritos en la presente descripción. Especies
- 10 adicionales de Cas9 incluyen: *Acidovorax avenae*, *Actinobacillus pleuropneumoniae*, *Actinobacillus succinogenes*, *Actinobacillus suis*, *Actinomyces* sp, *denitrificans cycliphilus*, *Aminomonas paucivorans*, *Bacillus cereus*, *Bacillus smithii*, *Bacillus thuringiensis*, *Bacteroides* sp, *Blastopirellula marina*, *Bradyrhizobium* sp, *Brevibacillus latemsporus*, *Campylobacter coli*, *Campylobacter jejuni*, *Campylobacter lari*, *Candidatus*
- 15 *Puniceispirillum*, *Clostridium cellulolyticum*, *Clostridium perfringens*, *Corynebacterium accolens*, *Corynebacterium difteria*, *Corynebacterium matruchotii*, *Dinoroseobacter sliibae*, *Eubacterium dolichum*, gamma proteobacteria, *Gluconacetobacter diazotrophicus*, *Haemophilus parainfluenzae*, *Haemophilus sputorum*, *Helicobacter canadensis*, *Helicobacter cinaedi*, *Helicobacter mustelae*, *Ilyobacter polytropus*, *Kingella*
- 20 *kingae*, *Lactobacillus crispatus*, *Listeria ivanovii*, *Listeria monocytogenes*, *Listeriaceae* bacteria, *Methylocystis* sp., *Methylosinus trichosporium*, *Mobiluncus mulieris*, *Neisseria bacilliformis*, *Neisseria cinerea*, *Neisseria flavescens*, *Neisseria lactamica*. *Neisseria* sp., *Neisseria wadsworthii*, *Nitrosomonas* sp., *Phascolarctobacterium succinatutens*, *Pasteurella multocida*, *Phascolarctobacterium succinatutens*, *Ralstonia syzygii*,
- 25 *Rhodopseudomonas palustris*, *Rhodovulum* sp., *Simonsiella muelleri*, *Sphingomonas* sp., *Sporolactobacillus manteletes*, *Staphylococcus lugdunensis*, *Streptococcus* sp., *Subdoligranulum* sp., *Tisliella mobilis*, *Treponema* sp. o *Verminephrobacter eiseniae*.

En algunas modalidades, la presente descripción enseña el uso de herramientas para técnicas de edición del genoma en plantas tales como cultivos y métodos de edición de

genes mediante el uso de endonucleasas asociadas a CRISPR (cas), lo que incluye

30 SpyCas9, SaCas9, St1Cas9. Estas poderosas herramientas para la edición del genoma, que pueden aplicarse a la edición del genoma de una planta, se conocen bien en la técnica. Ver ejemplo, Song y otros (2016), CRISPR/Cas9: A powerful tool for crop

genome editing, *The Crop Journal* 4: 75-82, Mali y otros (2013) RNA-guided human

35 genome engineering via cas9, *Science* 339: 823-826; Ran y otros (2015) In vivo genome

editing using staphylococcus aureus cas9, Nature 520: 186-191; Esvelt y otros (2013) Orthogonal cas9 proteins for rna-guided gene regulation and editing, Nature methods 10(11): 1116-1121, cada uno de los cuales se incorpora aquí como referencia en su totalidad para todos los propósitos.

5 (iii) *CRISPR/Cpf1*

En otras modalidades, la presente descripción proporciona métodos de edición de genes mediante el uso de un sistema CRISPR de Tipo V. En algunas modalidades, la presente descripción proporciona métodos de edición de genes mediante el uso de CRISPR de *Prevotella*, *Francisella*, *Acidaminococcus*, *Lachnospiraceae*, y *Moraxella* (Cpf1).

- 10 Los sistemas CRISPR Cpf1 de la presente descripción comprenden **i)** una única proteína endonucleasa, y **ii)** un ARNcr, en donde una porción del extremo 3' del ARNcr contiene la secuencia guía complementaria a un ácido nucleico objetivo. En este sistema, la nucleasa Cpf1 es reclutada directamente por el ARNcr en el ADN objetivo. En algunas modalidades, las secuencias guía para Cpf1 deben ser al menos de 12 nt,
- 15 13 nt, 14 nt, 15 nt o 16 nt para lograr una escisión de ADN detectable, y un mínimo de 14 nt, 15 nt, 16 nt, 17 nt o 18 nt para lograr una escisión de ADN eficiente.

Los sistemas Cpf1 de la presente descripción se diferencian de Cas9 en una variedad de formas. Primero, a diferencia de Cas9, Cpf1 no requiere un ARNtracr separado para la escisión. En algunas modalidades, los ARNcr de Cpf1 pueden ser tan cortos como

20 aproximadamente de 42-44 bases de longitud, de las cuales de 23-25 nt es la secuencia guía y 19 nt es la secuencia constitutiva de repetición directa. Por el contrario, las secuencias sintéticas combinadas de ARNcr y ARNtracr de Cas9 pueden tener una longitud de aproximadamente 100 bases.

En segundo lugar, ciertos sistemas de Cpf1 prefieren un motivo PAM "TTN" que se

25 localiza 5' aguas arriba de su objetivo. Esto contrasta con los motivos PAM "NGG" localizados en el 3' del ADN objetivo para sistemas Cas9 comunes tales como el Cas9 de *Streptococcus pyogenes*. En algunas modalidades, la base de uracilo que precede inmediatamente a la secuencia guía no puede sustituirse (Zetsche, B. y otros, 2015. "Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System" Cell

30 163, 759-771, que se incorpora aquí como referencia en su totalidad para todos los propósitos).

En tercer lugar, los sitios de corte para Cpf1 se escalonan por aproximadamente 3-5 bases, que crean "extremos pegajosos" (Kim y otros, 2016. "Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells" publicado en línea el 6 de

35 junio de 2016). Se cree que estos extremos pegajosos con salientes de 3 a 5 nt facilitan

la ligadura mediada por NHEJ, y mejoran la edición de genes de fragmentos de ADN con extremos coincidentes. Los sitios de corte están en el extremo 3' del ADN objetivo, distal al extremo 5' donde está el PAM. Las posiciones de corte suelen seguir la base 18 en la cadena no hibridada y la base 23 correspondiente en la cadena complementaria hibridada con el ARNcr.

En cuarto lugar, en los complejos Cpf1, la región "semilla" se localiza dentro de los primeros 5 nt de la secuencia guía. Las regiones semilla de ARNcr de Cpf1 son muy sensibles a las mutaciones, e incluso las sustituciones de una sola base en esta región pueden reducir drásticamente la actividad de escisión (ver Zetsche B. y otros 2015 "Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System" Cell 163, 759-771). Críticamente, a diferencia del objetivo Cas9 CRISPR, los sitios de escisión y la región semilla de los sistemas Cpf1 no se superponen. Información adicional sobre el diseño de oligos dirigidos al ARNcr de Cpf1 está disponible en Zetsche B. y otros 2015. ("Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System" Cell 163, 759-771).

(iv) ARN guía (ARNg)

En algunas modalidades, el ARN guía de la presente descripción comprende dos regiones codificantes, que codifican ARNcr y ARNtracr, respectivamente. En otras modalidades, el ARN guía es un híbrido sintético de ARNcr/ARNtracr del ARN guía único (ARNg). En otras modalidades, el ARN guía es un ARNcr para una endonucleasa Cpf1. Los expertos en la técnica apreciarán que, a menos que se indique de cualquier otra manera, todas las referencias a un ARN guía único (ARNsg) en la presente descripción pueden leerse como referencias a un ARN guía (ARNg). Por lo tanto, también se entenderá que las modalidades descritas en la presente descripción que se refieren a un ARN guía único (ARNg) se refieren a un ARN guía (ARNg).

El ARN guía se diseña para reclutar la endonucleasa CRISPR en una región de ADN objetivo. En algunas modalidades, la presente descripción enseña métodos para identificar los sitios objetivos de aterrizaje viables CRISPR, y diseñar ARNs guía para dirigirse a los sitios. Por ejemplo, en algunas modalidades, la presente descripción enseña algoritmos diseñados para facilitar la identificación de sitios de aterrizaje de CRISPR dentro de las regiones de ADN objetivo.

En algunas modalidades, la presente descripción enseña el uso de programas informáticos diseñados para identificar secuencias objetivo de CRISPR candidatas en ambas cadenas de una secuencia de ADN de entrada basado en la longitud de secuencia guía deseada y una secuencia de motivo CRISPR (PAM, motivo adyacente

protoespaciador) para una enzima CRISPR especificada. Por ejemplo, los sitios objetivo para Cpf1 desde *Francisella novicida* U112, con secuencias PAM TTN, pueden identificarse al buscar 5'-TTN-3' tanto en la secuencia de entrada como en el complemento inverso de la entrada. Los sitios objetivo para Cpf1 de la bacteria

5 *Lachnospiraceae* y *Acidaminococcus* sp., con secuencias PAM TTTN, puede identificarse al buscar 5'-TTTN-3' tanto en la secuencia de entrada como en el complemento inverso de la entrada. Igualmente, los sitios objetivo para Cas9 de CRISPR de *S. thermophilus*, con secuencia PAM NNAGAAW, pueden identificarse al buscar 5'-Nx-NNAGAAW-3' tanto en la secuencia de entrada como en el complemento

10 inverso de la entrada. La secuencia PAM para Cas9 de *S. pyogenes* es 5'-NGG-3'. Dado que múltiples apariciones en el genoma del sitio objetivo del ADN pueden conducir a una edición del genoma inespecífico, después de identificar todos los sitios potenciales, las secuencias pueden filtrarse basado en el número de veces que aparecen en el genoma de referencia relevante o en el constructo modular CRISPR.

15 Para aquellas enzimas CRISPR para las que la especificidad de secuencia está determinada por una secuencia "semilla" (tales como los primeros 5 pb de la secuencia guía para la escisión mediada por Cpf1), la etapa de filtrado también puede tener en cuenta las limitaciones de la secuencia semilla.

En algunas modalidades, las herramientas algorítmicas también pueden identificar sitios

20 potenciales fuera del objetivo para una secuencia guía particular. Por ejemplo, en algunas modalidades, puede usarse Cas-Offinder para identificar sitios potenciales fuera del objetivo de Cpf1 (ver Kim y otros, 2016. "Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells" Nature Biotechnology 34, 863-868). También puede usarse cualquier otra herramienta de diseño/identificación de CRISPR

25 disponible públicamente, lo que incluye, por ejemplo, la herramienta crispr.mit.edu del laboratorio de Zhang (ver Hsu, y otros, 2013 "DNA targeting specificity of RNA guided Cas9 nucleases" Nature Biotech 31, 827-832).

En algunas modalidades, se le puede permitir al usuario elegir la longitud de la secuencia de semillas. También se le puede permitir al usuario especificar el número de

30 apariciones de la secuencia semilla: la secuencia PAM en un genoma con el propósito de pasar el filtro. Lo predeterminado es buscar secuencias únicas. El nivel de filtración se modifica al cambiar tanto la longitud de la secuencia de semilla como el número de apariciones de la secuencia en el genoma. El programa puede proporcionar además o de manera alternativa la secuencia de una secuencia guía complementaria a la

secuencia o secuencias objetivo informadas al proporcionar el complemento inverso de la secuencia o secuencias objetivo identificadas.

En el ARN guía, la secuencia de la "secuencia espaciadora/guía" es complementaria a la secuencia del "protoespaciador" en el ADN objetivo. La proteína Cas9 reconoce el "andamio" del ARNg para una estructura de ARNg monocatenaria.

En algunas modalidades, la planta transgénica, parte de la planta, célula vegetal o cultivo de tejido vegetal que se enseña en la presente descripción comprende un constructo recombinante, que comprende al menos una secuencia de ácido nucleico que codifica un ARN guía. En algunas modalidades, el ácido nucleico se une operativamente a un promotor. En otras modalidades, un constructo recombinante comprende además una secuencia de ácido nucleico que codifica una endonucleasa de repeticiones palindrómicas cortas y regularmente interespaciadas (CRISPR). En otras modalidades, el ARN guía es capaz de formar un complejo con dicha endonucleasa CRISPR, y dicho complejo es capaz de unirse y crear una ruptura de doble cadena en una secuencia objetivo genómica de dicho genoma vegetal. En otras modalidades, la endonucleasa CRISPR es Cas9.

En otras modalidades, la secuencia objetivo es un ácido nucleico para *FusR1*, homólogos de *FusR1*, ortólogos de *FusR1* y/o parálogos de *FusR1*, y/o fragmentos y variaciones de los mismos. En algunas modalidades, la presente descripción enseña la edición de genes de *FusR1* en variedades de plátano sensibles a FW susceptibles a patógenos de *Fusarium* mediante el uso de técnicas de ingeniería genética descritas en la presente descripción.

La presente descripción enseña las técnicas de edición de genes dirigidas para modular, estimular y potenciar la resistencia a enfermedades al convertir alelos sensibles a FW en alelos resistentes a FW basado en la información de secuencia proporcionada en la presente descripción. La presente descripción enseña información de secuencia tanto de alelos resistentes a FW como de alelos sensibles a FW. Mediante el uso del sistema CRISPR/Cas, se introducen rasgos resistentes a FW en variedades de plátano sensibles a FW.

En algunas modalidades, los alelos de *FusR1* sensibles a FW deben ser el objetivo de la inactivación. En algunas modalidades, las secuencias de regiones conservadas responsables del rasgo de sensibilidad de FW pueden usarse para hacer maquinarias de edición de genes (tales como proteínas efectoras asociadas a CRISPR, ZFN, TALEN, etc.) para dirigirse a uno o más ortólogos de *FusR1*.

En algunas modalidades, la alteración de la expresión de los alelos endógenos sensibles a FW se lleva a cabo mediante una tecnología de edición de genes. En algunas modalidades, la desactivación de alelos sensibles a FW se lleva a cabo mediante tecnología de edición de genes. En algunas modalidades, la edición de bases de alelos sensibles a FW en alelos resistentes a FW se lleva a cabo mediante tecnología de edición de genes. En algunas modalidades, la tecnología de edición de genes es una ZFN. En otras modalidades, la tecnología de edición de genes es una TALEN. En otras modalidades, la tecnología de edición de genes es un sistema CRISPR/Cas. En otras modalidades, dicho sistema CRISPR comprende una molécula de ácido nucleico y una proteína enzimática, en donde la molécula de ácido nucleico es una molécula de ARN guía (ARNg) y la proteína enzimática es una proteína Cas o un ortólogo Cas. En otras modalidades, al menos dos casetes de expresión se apilan en tándem en el vector de expresión.

En algunas modalidades, las células vegetales modificadas comprenden una o más modificaciones (por ejemplo, inserciones, deleciones o mutaciones de uno o más ácidos nucleicos) en la secuencia de ADN genómico de un gen objetivo endógeno lo que da como resultado la función alterada del gen endógeno, lo que de esta manera modula, estimula, o potencia la resistencia a enfermedades. En dichas modalidades, las células vegetales modificadas comprenden un "gen objetivo endógeno modificado". En algunas modalidades, las modificaciones en la secuencia de ADN genómico provocan mutación, lo que altera de esta manera la función de la proteína FUSR1 sensible a FW en la proteína FUSR1 resistente a FW. En algunas modalidades, las modificaciones en la secuencia de ADN genómico dan como resultado sustituciones de aminoácidos, lo que de esta manera altera la función normal de la proteína codificada. En algunas modalidades, las modificaciones en la secuencia de ADN genómico codifican una proteína endógena modificada con función modulada, alterada, estimulada o potenciada de resistencia a enfermedades/patógenos en comparación con la versión no modificada (es decir, sensible a FW) de la proteína endógena en las accesiones de plátano sensibles a FW.

En algunas modalidades, las células vegetales modificadas descritas en la presente descripción comprenden uno o más genes objetivo endógenos modificados, en donde la una o más modificaciones dan como resultado una función alterada de un producto génico (es decir, una proteína) codificado por el gen objetivo endógeno en comparación con una célula vegetal no modificada. Por ejemplo, en algunas modalidades, una célula vegetal modificada demuestra la expresión de una proteína FUSR1 resistente a FW o

una expresión regulada positivamente de dicha proteína. En algunas modalidades, la expresión del producto génico (tal como *FusR1* resistente a FW modificado genéticamente a partir de *FusR1* sensible a FW) en una célula vegetal modificada se potencia en al menos un 0,5 %, 1 %, 2 %, 3 %, 4 %, 5 % o más en comparación con la expresión del producto génico (tal como *FusR1* sensible a FW) en una célula vegetal no modificada. En otras modalidades, la expresión del producto génico (tal como *FusR1* resistente a FW modificado genéticamente) en una célula vegetal modificada se potencia en al menos un 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 % o más en comparación con la expresión del producto génico (tal como *FusR1* sensible a FW) en una célula vegetal no modificada. En algunas modalidades, las células vegetales modificadas descritas en la presente descripción demuestran una expresión y/o función potenciadas de productos génicos codificados por una pluralidad (por ejemplo, dos o más) de genes objetivo endógenos en comparación con la expresión de los productos génicos en una célula vegetal no modificada. Por ejemplo, en algunas modalidades, una célula vegetal modificada demuestra una expresión y/o función potenciadas de productos génicos de 2, 3, 4, 5, 6, 7, 8, 9, 10 o más genes objetivo endógenos en comparación con la expresión de los productos génicos en una célula vegetal no modificada.

En algunas modalidades, las células vegetales modificadas descritas en la presente descripción comprenden uno o más genes objetivo endógenos modificados, en donde una o más modificaciones de la secuencia de ADN objetivo dan como resultado la expresión de una proteína con función reducida o alterada (por ejemplo, una "proteína endógena modificada") en comparación con la función de la proteína correspondiente expresada en una célula vegetal no modificada (por ejemplo, una "proteína endógena no modificada"). En algunas modalidades, las células vegetales modificadas descritas en la presente descripción comprenden 2, 3, 4, 5, 6, 7, 8, 9, 10 o más genes objetivo endógenos modificados que codifican 2, 3, 4, 5, 6, 7, 8, 9, 10 o más proteínas endógenas modificadas. En algunas modalidades, la proteína endógena modificada demuestra una afinidad de unión potenciada o alterada por otra proteína expresada por la célula vegetal modificada o expresada por otra célula; capacidad de señalización potenciada o alterada; actividad enzimática potenciada o alterada; actividad de unión al ADN potenciada o alterada; o capacidad reducida o alterada para funcionar como una proteína de andamiaje.

35

EJEMPLOS

La presente invención se ilustra adicionalmente mediante los siguientes ejemplos que no deben interpretarse como limitantes. El contenido de todas las referencias, patentes, y solicitudes de patentes publicadas citadas a lo largo de esta solicitud, así como también las Figuras, se incorporan en la presente descripción como referencia en su totalidad para todos los propósitos.

Ejemplo 1: Métodos y materiales para la secuenciación

(1) Material

Se obtuvieron tejidos de hojas de plátano frescos y liofilizados de Bioversity International (Lovaina, Bélgica), Inter-TROP CRB Plantas Tropicales (Guadalupe), y el IITA Genebank (Ibadan, Nigeria), Plant Delights Nursery (Raleigh, NC) y The Flower Bin (Longmont, CO).

(2) ARN

El ARN total se extrajo de hojas de plátano frescas, congeladas y liofilizadas mediante el uso de un protocolo de Ishihara modificado (Ishihara y otros, 2016). Alrededor de 100 mg de tejido de plátano fresco o congelado se molieron hasta obtener un polvo mediante el uso de un mortero y pistilo limpio, enfriado con hielo seco, que se trató con ARNasa Away™ (Invitrogen, Carlsbad, CA). Alrededor de 20 a 30 mg de tejido de plátano liofilizado se homogeneizaron en un tubo Lysing Matrix D (MP Bio, Santa Ana, CA) sin líquido. Se añadió a cada muestra un mililitro de tampón de lisis de polifenol (800 µl de tampón RLT (Qiagen, Germantown, MD), 200 µl de Fruit-mate (Takara, Mountain View, CA) y 10 µl de β-mercaptoetanol). Las muestras frescas y congeladas se homogeneizaron durante 40 segundos en el ajuste de velocidad 6 de un FastPrep 120 (ThermoFisher Scientific, Waltham, MA), mientras que las muestras liofilizadas se agitaron en vórtex a alta velocidad durante 1 minuto. Todas las muestras se incubaron en hielo durante 4 minutos y después se centrifugaron durante 2 minutos a 8000 x g. El sobrenadante se transfirió a un nuevo tubo de 2,0 ml y se añadió otro 1,0 ml de tampón de lisis de polifenol al sobrenadante. Las muestras se agitaron con vórtex a temperatura alta durante 1 minuto, se incubaron en hielo durante 4 minutos y se centrifugaron durante 2 minutos a 8000 x g. El sobrenadante se dividió entre dos columnas QIAshredder (Qiagen, Germantown, MD) y se centrifugó a velocidad máxima durante 2 minutos hasta que se procesó todo el sobrenadante. Las etapas restantes de la extracción de ARN se llevaron a cabo de acuerdo con el protocolo de Ishihara. El protocolo opcional de limpieza del ARN y digestión con ADNasa en solución también se realizó como se detalla en el protocolo RNeasy Mini (Qiagen, Germantown, MD). La

concentración y la pureza de la muestra se determinaron mediante el uso del espectrofotómetro NanoDrop™ One (ThermoFisher Scientific, Waltham, MA).

(3) ADN

El ADN total se extrajo de hojas de plátano frescas, congeladas y liofilizadas mediante el uso de un protocolo modificado del kit de aislamiento de ADN PowerPlant Pro (MO BIO, Carlsbad, CA). Alrededor de 40 mg de tejido de plátano fresco o congelado se trituraron hasta obtener un polvo mediante el uso de un mortero y pistilo limpio y enfriado con hielo seco que se trató con ARNasa Away™ (Invitrogen, Carlsbad, CA). Alrededor de 10 a 20 mg de tejido de plátano liofilizado se homogeneizaron en un tubo Lysing Matrix D (MP Bio, Santa Ana, CA) sin líquido. Las etapas restantes de la extracción de ADN se llevaron a cabo de acuerdo con el protocolo MO BIO. Se añadió solución de separación fenólica al tampón de lisis y se usaron 250 µl de tampón PD3. La concentración y la pureza de la muestra se determinaron mediante el uso del espectrofotómetro NanoDrop™ One (ThermoFisher Scientific, Waltham, MA).

(4) ADNc

El ADNc se sintetizó a partir de 1,0 µg del ARN total mediante el uso del kit de síntesis de ADNc 1st Strand (Epicenter, Madison, WI). Se usó el cebador adaptador (AP) del kit 3'-RACE de Invitrogen (Invitrogen, Carlsbad, CA) en lugar del cebador poli dT.

(5) Cebadores

Las secuencias de cebadores se diseñaron contra regiones homólogas de genes objetivo supuestos con temperaturas de hibridación de 57 °-64° C mediante el uso del programa OligoAnalyzer Tool (IDT, Coralville, IA). Los cebadores se adquirieron de IDT.

(6) PCR

Las reacciones de PCR se realizaron en reacciones de 25 µl que contenían una concentración final de tampón Phusion® HF 1X, 300 µM de cada dNTP, 0,3 µM de cada cebador directo e inverso, 0,5 unidades de ADN polimerasa de alta fidelidad 1X Phusion® (ThermoFisher Scientific, Waltham, MA) en un termociclador Veriti (Applied Biosystems, Carlsbad, CA). Las condiciones generales de PCR fueron 98 °C durante 2 minutos, seguidas de 35 ciclos de 98 °C durante 10 segundos, de 55 °-62 °C durante 30 segundos (en dependencia del cebador Ta) y 72 °C durante 30 segundos, antes de una extensión final a 72 °C durante 10 minutos y una retención a 4 °C. Los productos de PCR se ejecutaron en un gel de agarosa al 1,5 % y se visualizaron mediante el uso GelRed® Nucleic Acid Stain (Biotium, Hayward, CA) en un Alpha Imager EC (Alpha Innotech, San Leandro, CA).

(7) Clonación

Los fragmentos de PCR se clonaron mediante el uso del kit de clonación de PCR Zero Blunt TOPO (Invitrogen, Carlsbad, CA) mediante el uso de 4 µl de producto de PCR, de acuerdo con el protocolo del fabricante. El vector ligado se transformó en células químicamente competentes Top10 One Shot (Invitrogen, Carlsbad, CA) mediante el uso del protocolo de transformación química. Las células de *E. coli* transformadas se sembraron en placas de agar LB que contenían 50 µg/ml de kanamicina y las placas se cultivaron durante la noche a 37 °C.

(8) PCR de colonias

Las colonias que contenían plásmidos recombinantes se detectaron mediante el uso de PCR con cebadores M13 directos e inversos. Las reacciones de PCR se realizaron en volúmenes de 15 µl que contenían Tris-SO4 60 mM (pH 8,9), sulfato de amonio 18 mM, sulfato de magnesio 2,0 mM, 0,2 mM de cada dNTP, 0,2 µM de cada cebador directo e inverso, 0,3 unidades de platino Taq Hi Fidelity (Invitrogen, Carlsbad, CA) en un termociclador Veriti (Applied Biosystems, Carlsbad, CA). Las colonias se recogieron e inocularon en la reacción de PCR, seguido de una inoculación de 50 µl de LB-kanamicina. Las condiciones de PCR de la colonia fueron 94 °C durante 2 minutos, seguidas de 35 ciclos de 94 °C durante 30 segundos, 50 °C durante 30 segundos y 68 °C durante 1 minuto, antes de una extensión final a 68 °C durante 10 minutos y un mantenimiento a 4 °C. Los productos de PCR se corrieron en un gel de agarosa al 1,5 % y se visualizaron mediante el uso de tinción de ácido nucleico GelRed® (Biotium, Hayward, CA) en un Alpha Imager EC (Alpha Innotech, San Leandro, CA). Se secuenciaron las reacciones de PCR de colonias que produjeron productos del tamaño esperado.

(9) Secuenciación

Se prepararon cinco microlitros de cada producto de PCR para la secuenciación mediante tratamiento enzimático mediante el uso de 2 µl de ExoSAP-IT de alto rendimiento (Affymetrix, Santa Clara, CA). Las reacciones se incubaron a 37 °C durante 15 minutos, seguido de 15 minutos a 80 °C. La plantilla se etiquetó para la secuenciación mediante el uso del kit de secuenciación de ciclo BigDye Terminator v3.1 (Applied Biosystems, Carlsbad, CA) de la siguiente manera: se añadieron 2 µl de la plantilla y 2 µl de un cebador de secuenciación 0,8 µM a una mezcla de tampón de secuenciación BigDye Terminator, mezcla de reacción preparada BigDye Terminator v3.1, y agua, en una reacción de 10 µl. Las condiciones de la reacción de secuenciación de BigDye fueron las siguientes: 96 °C durante 1 minuto, seguido de 25 ciclos de 96 °C durante 10 segundos, 50 °C durante 5 segundos y 60 °C durante 75 segundos. Los terminadores

BigDye no incorporados se eliminaron mediante el uso del kit de purificación BigDye XTerminator (Applied Biosystems, Carlsbad, CA). Las reacciones se secuenciaron mediante el uso del analizador genético Applied Biosystems 3500 (Applied Biosystems, Carlsbad, CA).

5 **(10) Alineamiento de secuencia**

Los archivos de secuencia del ABI 3500 Genetic Analyzer se importaron a Sequencher v4.8 Build 3767 (Gene Codes, Ann Arbor, MI). La secuencia de vectores se recortó mediante el uso de la herramienta Recortar vector. A continuación, las secuencias se alinearon automáticamente y se editaron manualmente para secuenciar los artefactos.

10 **Ejemplo 2: Identificación de diferencias estructurales entre los genes resistentes a la marchitez por *Fusarium* (FW) y los genes sensibles a la marchitez por *Fusarium* (FW)**

En este ejemplo, los genes de resistencia a la marchitez por *Fusarium* se descubrieron mediante el análisis, como se describe más abajo, de secuencias de ADN recuperadas de GenBank. Se descargaron secuencias de nucleótidos de varias especies de plátano (es decir, *Musa itinerans*, *Musa acuminata*, *Musa basjoo*, *Musella lasiocarpa*, *Musa balbisiana*). La secuencia de *FusR1* de *M. itinerans* se obtuvo de múltiples accesiones (ITC1526, ITC1571 y PT-BA-00223), todas las cuales son resistentes a FW. La secuencia de *FusR1* de *M. acuminata* etiquetada como 'resistente a FW' se obtuvo de múltiples accesiones resistentes a FW, lo que incluye ITC0896 (*M. a.* subespecie *banksii*) y PT_BA-00281 (Pisang Bangkahulu). La secuencia de *M. acuminata* etiquetada como 'sensible' es de las accesiones sensibles a FW (ITC0507, ITC0685, PT-BA-00304, PT-BA-00310 y PT-BA-00315). Estas accesiones incluyen múltiples muestras de cultivares de plátano tales como Pisang Madu, Pisang Pipit y Pisang Rojo Uter, todos los cuales se han caracterizado bien como sensibles a FW (Chen y otros, 2019). La secuencia de *M. balbisiana* se obtuvo de varias accesiones sensibles a FW, lo que incluye ITC1016, ITC0545, ITC0080 e ITC0565. *FusR1* de *M. basjoo* es de accesiones resistentes a FW (ITC0061 y PD # 3064). A continuación, se aplicó un análisis bioinformático automatizado a cada comparación por pares y solo aquellas secuencias que contienen un cambio (o cambios) de nucleótidos que producen cambios evolutivamente significativos se conservaron para análisis adicionales. Esto permitió la identificación de genes que han evolucionado para conferir alguna ventaja evolutiva, así como también la identificación de los cambios evolucionados específicos.

35 Puede emplearse cualquiera de varios análisis de evolución molecular diferentes o métodos de tipo Ka/Ks para evaluar cuantitativa y cualitativamente la importancia

evolutiva de los cambios de nucleótidos identificados entre secuencias de genes homólogos de especies relacionadas (Kreitman y Akashi, 1995; Li, 1997). Por ejemplo, la selección positiva en proteínas (es decir, evolución adaptativa a nivel molecular) puede detectarse en genes que codifican proteínas mediante comparaciones por pares de las relaciones de sustituciones de nucleótidos no sinónimas por sitio no sinónimo (Ka)/sustituciones sinónimas por sitio sinónimo (Ks) (Li y otros, 1985; Li, 1993). Puede usarse cualquier comparación de Ka y Ks, aunque es particularmente conveniente y más efectivo comparar estas dos variables como una relación. Las secuencias se identifican al mostrar una diferencia estadísticamente significativa entre Ka y Ks mediante el uso de métodos estadísticos estándar.

En algunos aspectos, el análisis Ka/Ks de Li y otros, (1993) se usa para llevar a cabo la presente descripción, aunque también pueden usarse otros programas de análisis que pueden detectar genes seleccionados positivamente entre especies (Li y otros, 1985; Li, 1993; Messier y Stewart, 1997; Nei, 1987).

El método Ka/Ks, que comprende una comparación de la tasa de sustituciones no sinónimas por sitio no sinónimo con la tasa de sustituciones sinónimas por sitio sinónimo entre regiones de genes codificantes de proteínas homólogas en términos de una relación, se usa para identificar sustituciones de secuencia que pueden ser impulsadas por selección adaptativa en oposición a sustituciones neutrales durante la evolución.

Una sustitución sinónima ("silenciosa") es aquella que, debido a la degeneración del código genético, no modifica la secuencia de aminoácidos codificada; una sustitución no sinónima da como resultado un reemplazo de aminoácidos. El alcance de cada tipo de cambio puede estimarse como Ka y Ks, respectivamente, el número de sustituciones sinónimas por sitio sinónimo y sustituciones no sinónimas por sitio no sinónimo. Los cálculos de Ka/K pueden realizarse manualmente o mediante el uso de un programa informático. Un ejemplo de programas adecuados son Li93 (Li, 1993) o MEGA X: Molecular Evolutionary Genetics Analysis Across Computing Platforms (Kumar y otros, 2018)

Con el propósito de estimar Ka y Ks, se usan secuencias codificantes de proteínas completas o parciales para calcular el número total de sustituciones sinónimas y no sinónimas, así como también sitios sinónimos y no sinónimos. La longitud de la secuencia de polinucleótidos analizada puede ser cualquier longitud apropiada. Preferentemente, la secuencia codificante completa se compara con el fin de determinar todos y cada uno de los cambios significativos. Los programas de computadora disponibles al público, tales como Li93 (Li, 1993), o MEGA X: Molecular Evolutionary

Genetics Analysis Across Computing Platforms (Kumar y otros, 2018) pueden usarse para calcular los valores de Ka y Ks para todas las comparaciones por pares.

Este análisis puede adaptarse aún más para examinar secuencias en una forma de "ventana deslizante" de manera que una pequeña cantidad de cambios importantes no queden enmascarados por toda la secuencia. "Ventana deslizante" se refiere al examen de subsecciones superpuestas consecutivas del gen (las subsecciones pueden tener cualquier longitud).

La comparación de las tasas de sustitución sinónima y no sinónima se representa comúnmente mediante la relación Ka/Ks. Se ha demostrado que Ka/Ks es un reflejo del grado en que la evolución adaptativa ha funcionado en la secuencia en estudio. Pueden usarse segmentos de longitud completa o parciales de una secuencia codificante para el análisis de Ka/Ks. Cuanto mayor sea la relación Ka/Ks, más probable es que una secuencia haya experimentado una evolución adaptativa y las sustituciones no sinónimas sean evolutivamente significativas. Ver, por ejemplo, Messier y Stewart (1997).

Las relaciones Ka/Ks significativamente mayores que uno (1,0) sugieren fuertemente que la selección positiva ha fijado un mayor número de reemplazos de aminoácidos de lo que se puede esperar como resultado del azar solo y contraste con el patrón más comúnmente observado en donde la proporción es menor que o igual a uno (Nei, 1987; Hughes y Nei, 1988; Messier y Stewart, 1994; Kreitman y Akashi, 1995; Messier y Stewart, 1997). Las relaciones inferiores a uno generalmente significan el papel de selección negativa o purificadora, lo que indica que existe una fuerte presión sobre la estructura primaria de las proteínas funcionales y eficaces para que permanezcan sin cambios.

Todos los métodos para calcular las relaciones Ka/Ks se basan en una comparación por pares del número de sustituciones no sinónimas por sitio no sinónimo con el número de sustituciones sinónimas por sitio sinónimo para las regiones codificantes de proteínas de genes homólogos de especies relacionadas. Cada método implementa diferentes correcciones para estimar "múltiples aciertos" (es decir, más de una sustitución de nucleótidos en el mismo sitio). Cada método también usa diferentes modelos de cómo cambian las secuencias de ADN a lo largo del tiempo evolutivo. Por tanto, preferentemente, se usa una combinación de resultados de diferentes algoritmos para aumentar el nivel de sensibilidad para la detección de genes seleccionados positivamente y la confianza en el resultado.

Se entiende que los métodos descritos en la presente descripción podrían conducir a la identificación de secuencias de polinucleótidos de banano que se relacionan funcionalmente con secuencias codificantes de proteínas de banano. Dichas secuencias pueden incluir, pero no se limitan a, secuencias no codificantes o secuencias codificantes que no codifican proteínas. Estas secuencias relacionadas pueden ser, por ejemplo, físicamente adyacentes a las secuencias codificantes de la proteína del banano en el genoma del banano, tales como intrones o secuencias flanqueantes 5' y 3' (lo que incluye elementos de control tales como promotores y potenciadores). Estas secuencias relacionadas pueden obtenerse mediante la búsqueda en una base de datos pública del genoma tal como GenBank o, de manera alternativa, mediante el tamizaje y la secuenciación de una biblioteca genómica apropiada con una secuencia codificante de proteínas como sonda.

Después de que se identificaron los genes candidatos, las secuencias de nucleótidos de los genes en cada par de genes ortólogos se verificaron cuidadosamente mediante técnicas de secuenciación de ADN estándar y después se repitió el análisis Ka/Ks para cada par de genes candidatos cuidadosamente secuenciados. Más específicamente, el programa informático realizó todas las posibles comparaciones por pares entre los supuestos ortólogos de cada gen del banano cultivado, *Musa acuminata* (AAA subgr. Cavendish) en comparación con los ortólogos de las especies silvestres, en búsqueda de relaciones altas de Ka/Ks. El programa informático realizó un BLAST (de forma automatizada) de cada secuencia de ARNm de plátano cultivado contra cada secuencia en el transcriptoma que se secuenció a partir de un pariente silvestre, por ejemplo, *M. balbisiana*. A continuación, el programa informático realizó un análisis de Ka/Ks para cada par de genes (es decir, cada conjunto de ortólogos), al marcar los pares de genes con puntuaciones altas de Ka/Ks.

A continuación, el programa informático comparó cada secuencia de plátano cultivada con cada secuencia de otro pariente silvestre, por ejemplo, *M. basjoo*, nuevamente al hacer una serie de BLAST y después examinar para obtener puntuaciones altas de Ka/Ks. Por tanto, hace esto para la secuencia del transcriptoma de todas las especies silvestres en sucesión. Esto proporciona un conjunto de candidatos (ver más abajo) para un análisis posterior. A continuación, el programa informático comparó cada secuencia de genes en el transcriptoma de *M. balbisiana* con cada secuencia de *M. basjoo*, nuevamente al hacer una serie de BLASTs y después examinar para obtener puntuaciones altas de Ka/Ks. Por tanto, finalmente comparó todos los genes expresados representados en las bibliotecas de ADNc utilizadas de cada especie de plátano con

todos los genes de todas las demás especies de plátano, tanto silvestres como cultivadas, con el objetivo de encontrar todos los genes que muestren evidencia de selección positiva.

Los pares de genes marcados que surgieron se volvieron a secuenciar a continuación de forma individual y cuidadosa en el laboratorio para comprobar la precisión de las lecturas originales de alto rendimiento para eliminar los falsos positivos.

A continuación, se examinó cada par de genes candidatos restantes con una puntuación alta de Ka/Ks para determinar si la comparación era verdaderamente ortóloga o simplemente un falso positivo artificial provocado por una comparación paróloga.

Mediante el uso de la metodología descrita anteriormente, se analizaron las secuencias de genes de plátano disponibles en GenBank para identificar un gen seleccionado positivamente que no se ha relacionado con el rasgo de resistencia a FW en especies de plátano en la técnica. El inventor identificó y seleccionó este gen que se esperaba que produjera resistencia a FW y después lo nombró como *Resistencia a Fusarium 1* (FusR1). Sorprendentemente, el inventor encontró una relación Ka/Ks inusualmente alta de 3,6 entre el ortólogo de *FusR1* del pariente silvestre altamente resistente *M. itinerans* y el *FusR1* del Cavendish sensible a FW (*M. acuminata*).

El inventor obtuvo accesiones de varios tipos de plátanos, lo que incluye cultivares de plátano y razas locales, así como también especies de banano silvestres (no domesticadas) de los géneros *Musa*, *Musella* y *Ensete*. Estos tres géneros comprenden la familia del plátano Musaceae. El inventor hizo esfuerzos sustanciales para obtener múltiples muestras de accesiones de *Musa acuminata* (genoma "A") y *M. balbisiana* (genoma "B"), para muestrear adecuadamente la diversidad taxonómica y geográfica de los bananos. El inventor obtuvo accesiones de la mayoría de las subespecies *acuminata*. Además, para el análisis de grupos externos, el inventor obtuvo accesiones de plantas de familias de plantas conocidas por estar estrechamente relacionadas con Musaceae.

Es bien sabido que algunas especies/variedades de plátano del grupo B son altamente susceptibles a *Foc* -TR4 (Chen y otros, 2019), aunque a veces muestran características agronómicas convenientes como la tolerancia a la sequía. Los plátanos del genoma A muestran un rango de resistencia, tolerancia y sensibilidad a *Fusarium*, en dependencia de la especie particular o variedad cultivada. Como consecuencia, muchas especies de plátano silvestre y variedades de banano cultivadas se han caracterizado cuidadosa y rigurosamente por su resistencia, tolerancia o sensibilidad a TR4 (Li y otros, 2012, Ssali

y otros, 2013, Li y otros, 2015, Wu y otros, 2016, Ribeiro y otros, 2018, Niu y otros, 2018 y Zuo y otros, 2018).

Siempre que fue posible, el inventor optó por preparar tanto ARN (para la conversión en ADNc) como ADN genómico (ADNg). La mayoría de las accesiones se obtuvieron como muestras frescas, congeladas o liofilizadas, y esto usualmente permitió una extracción exitosa de ARN. Para algunas muestras, particularmente cuando son más antiguas o están parcialmente degradadas, solo se pudo aislar el ADNg. Las secuencias de ARNm y/o la secuencia codificante solamente), la secuencia del intrón y algunas secuencias (ver Listado de secuencias) para varias especies de *Musa*, *Musella*, *Ensete*, y de otros grupos se proporcionan aquí como se describe en la Tabla 1 y en los listados de secuencias. Las descripciones detalladas de los métodos se dan en la sección de Métodos y Materiales del Ejemplo 1.

Los plátanos cultivados son el producto de eventos de hibridación entre los plátanos del genoma B (el grupo *Musa balbisiana*) y los plátanos del genoma A (el grupo *M. acuminata*). Es bien sabido que algunas especies/variedades de plátano del grupo B son susceptibles a *Foc*-TR4 (Chen y otros, 2019), aunque a veces poseen rasgos agronómicos convenientes tales como la tolerancia a la sequía (REF). Por el contrario, los plátanos del genoma A muestran un rango de resistencia, tolerancia y sensibilidad a *Fusarium*, en dependencia de la especie particular o variedad cultivada. Algunas especies del grupo del genoma A, tales como *Musa itinerans* y *M. basjoo*, han demostrado ser extremadamente resistentes a *Foc*-TR4 (Li y otros, 2015; Wu y otros, 2016), mientras que algunos cultivares del genoma A tales como Cavendish son exquisitamente sensibles a *Fusarium*.

El análisis de estas secuencias reveló un resultado importante; que es que todas las especies de plátano del genoma "A" (o variedades de plátano cultivadas) que se han caracterizado como resistentes a *Fusarium* comparten secuencias *FusR1* que pertenecen a un grupo común, mientras que las especies/variedades de plátano sensibles a *Fusarium* pertenecen a un grupo diferente. Sorprendentemente, cada accesión del genoma B que el inventor examinó es "sensible a FW", y todas las secuencias de *FusR1* de las accesiones del genoma B están rotas y/o dañadas de alguna manera con alguna combinación de deleciones de pares de bases de secuencias codificantes. A menudo, la deleción tiene un tamaño largo, tal como 82 u 85 pb, sin embargo, el inventor también encontró una deleción de base única consistente. Estas deleciones alteran la secuencia de la proteína inferida al destruir el marco de lectura, lo que generalmente da como resultado una proteína truncada. Además, todas las

secuencias codificantes de *FusR1* del genoma B contienen un intrón no empalmado de 84 pb, que a menudo aparece junto con la delección de 85 pb.

En cuanto a los plátanos con genoma A, el inventor descubrió que todas las accesiones del genoma A que se conoce son resistentes a *Foc*-TR4 comparten un grupo de secuencia de *FusR1* común, mientras que las accesiones de genoma A sensibles a *Foc*-TR4 comparten un grupo de secuencia de *FusR1* diferente.

Esta es una fuerte evidencia de que *FusR1* es responsable de los patrones de resistencia a enfermedades observadas entre las especies resistentes a *Fusarium* frente a las sensibles a *Fusarium*. Los análisis en este ejemplo sugieren que la diferencia de resistencia a la sensibilidad a la raza 4 de *Fusarium* se relaciona fuertemente con las diferencias de secuencia de *FusR1*.

Un respaldo adicional a esto proviene de nuestro examen de las pocas especies de plátano que se han caracterizado como 'tolerantes a la marchitez por *Fusarium*'. Todas estas especies tienen secuencias de *FusR1* que caen en un tercer grupo de secuencias, todos son intermediarias entre los grupos de secuencias resistentes a *Fusarium* y sensibles a *Fusarium*.

La industria platanera se vio obligada en la década de 1950 a pasar de su cultivar principal, Gros Michel, al cultivar Cavendish cuando la raza 1 de *Fusarium* (Enfermedad de Panamá) representó una amenaza crítica para Gros Michel. Se encontró que Cavendish, que es medio hermano de Gros Michel (ambos son especies del genoma "A"), es resistente a la raza 1. Por tanto, los cultivares Cavendish y Gros Michel estrechamente relacionados muestran diferentes perfiles de resistencia a las diversas razas de *Fusarium*. (Ambos son sensibles a *Foc*-TR4, la amenaza actual para la industria platanera).

El inventor secuenció *FusR1* a partir de varias accesiones de *Musa acuminata*. En cada caso, el inventor clonó, como se describe en el Ejemplo 1, el gen *FusR1* y después secuenció múltiples clones del gen *FusR1*. Algunas de estas accesiones de *M. acuminata* se han caracterizado bien por su resistencia/sensibilidad a la marchitez por *Fusarium*. El inventor encontró tres alelos para *FusR1* de *M. acuminata*. La observación crítica es que todas las accesiones resistentes a la marchitez por *Fusarium* comparten secuencias de *FusR1* similares. Los dos alelos de *FusR1* de accesiones de *M. acuminata* resistentes a FW son el alelo *FusR1* resistente a la marchitez por *Fusarium* o simplemente, los "alelos resistentes" (SEQ ID NO: 8 y SEQ ID NO: 10). Por el contrario, todas las accesiones de *M. acuminata* sensibles a FW comparten un alelo diferente, denominado alelo *FusR1* sensible a la marchitez por *Fusarium* (SEQ ID NO:

13). Los alelos FusR1 resistentes a FW difieren en solo unas pocas sustituciones críticas de nucleótidos del alelo sensible a FW. (Ver La Figura 1). Esto sugiere fuertemente que la resistencia/sensibilidad a la marchitez por Fusarium se controla por el alelo FusR1 particular que porta una planta de banano.

5 **Ejemplo 3: Mejoramiento de la resistencia del plátano**

Las versiones tetraploide de cultivares Cavendish sensibles a FW (*M. acuminata*; AAAA) están disponibles o pueden desarrollarse a través de grandes programas de polinización/mejoramiento enfocados en crear, identificar y aislar el porcentaje relativamente bajo de progenie tetraploide que se produce (por ejemplo, Aguilar Morán, J.F., 2013, Improvement of Cavendish Banana cultivars through conventional breeding, Acta Hortic. 986: 205-208; Jenny y otros, En Jacome y otros, Editores, *Mycosphaerella* leaf spot diseases of banana: present status and outlook, Proceedings of the 2nd International Workshop on *Mycosphaerella* leaf spot diseases celebrado en San José, Costa Rica, 20-23 de mayo de 2002, Sesión 4, páginas 199-208) o al someter genotipos diploides AA a poliploidización *in vitro* (Amah y otros, noviembre de 2019, Frontiers in Plant Science, Vol.10, Artículo 1450, 12 páginas).

Las versiones diploides de FusR1 (AA) resistentes a FW de *M. acuminata ssp. Banksia* pueden identificarse o desarrollarse mediante el uso de métodos conocidos por los expertos en la técnica (por ejemplo, Bakry y otros, Capítulo 1, Genetic Improvement in Banana, 50 páginas, En Breeding Plantation Tree Crops: Tropical Species, 2009). Los diploides resultantes se seleccionan para detectar la presencia de SEQ ID NO: 8 y/o SEQ ID NO: 10 (secuencias de ARNm).

Una planta Cavendishtetraploide sensible a FW, tal como un tetraploide de la variedad 'Naine' o 'Williams', puede usarse como parental masculino en cruzamientos con una planta diploide de *M. acuminata ssp. banksia* resistente a FW, tal como un diploide 'ITC0896', usado como progenitor femenino.

Un gran número de la progenie resultante se detecta en busca de plantas triploides (AAA) que comprenden SEQ ID NO: 8 y/o SEQ ID NO: 10 (secuencias de ARNm) y posteriormente se evalúan para determinar los rasgos agronómicos.

Todas las plantas de banano resultantes con resistencia a TR4 pueden mantenerse mediante reproducción asexual y usarse para la producción o en programas de mejoramiento posteriores.

Ejemplo 4: Materiales y métodos para la transformación de plantas

Los sistemas de transformación de plátano usarán material estéril de cepas de banano seleccionadas. Se usará una variedad de metodologías de cultivo y transformación de

tejidos para aumentar la probabilidad de éxito. Ver, por ejemplo, los protocolos de transformación descritos en Ploetz (2015, *Phytopathology* 105:1512-1521), Patente de los Estados Unidos núm. 7,534,930; Patente de los Estados Unidos núm. 6,133,035; Sagi y otros, *Bio/Technology* 13, 481-485, 1995; May y otros, *Bio/Technology* 13, 485-492, 1995; Vishnevetsky y otros, *Transgenic Res.* 20(1):61-71, 2011; Paul y otros (2011); Zhong y otros, *Plant Physiol.* 110, 1097-1107, 1996; Dugdale y otros, *Journal of General Virology* 79: 2301-2311, 1998; Mohan y Swennen (editores), 2004, *Banana improvement: cellular, molecular biology, and induced mutations*, Science Publishers, Inc.; y, Remy y otros, 2013, *Genetically modified bananas: Past, present and future*, *Acta Horticulturae* 974:71-80, cada uno de los cuales se incorpora en la presente descripción como referencia en su totalidad.

Estas metodologías se centrarán en las condiciones de cultivo de tejidos, al identificar diferentes tipos de tejidos para regeneración/disparo, formulaciones de medios, cepas de agrobacterias, casetes de selección, construcción de vectores de control y administración, administración de genes, marcadores seleccionables y sustratos de tejido/célula objetivo para el suministro y transformación del ADN. Los experimentos iniciales desplegarán vectores de control mediante el uso de marcadores visuales y casetes de selección para optimizar rápidamente la dirección experimental y detectar potenciales eventos transgénicos. Los experimentos paralelos estarán dirigidos a optimizar la eficiencia de transformación y usar genes de interés (GOI).

Se realizarán modificaciones en las formulaciones de los medios, los vectores y los procesos de transformación para mejorar la eficiencia de los procesos y la transformación. Los vectores de transformación que contienen genes clave de interés continuarán transformándose para producir eventos adicionales de sobreexpresión o desactivación. Los vectores que se usarán según sea necesario incluyen, pero no se limitan a, vectores apilados de múltiples genes, vectores de genes policistrónicos y vectores de edición CRISPR de múltiples ARNg para probar la eficacia en plátano. Las pruebas se realizarán en eventos T0 para mostrar la presencia y el número de copias del gen marcador seleccionable o del GOI. Además, el análisis de la expresión de ARNm se usará según sea necesario para cualquier GOI claves. El material vegetal supuestamente transformado se usará para pruebas o análisis posteriores.

Las tecnologías CRISPR se describen en detalle en otra parte de la presente descripción, lo que incluye las referencias a las composiciones y procedimientos para usar CRISPR para editar genomas de plantas, tal como el genoma del plátano. Las composiciones y procedimientos detallados para usar CRISPR para eliminar un gen en

plantas que da lugar a un fenotipo de interés (por ejemplo, resistencia a patógenos fúngicos como *Fusarium*) se proporcionan en los documentos WO 2019/118342 (PCT/US2018/064735), WO 2018/220581 (PCT/IB2018/053903) y US 2019/0032070 (US 16/072,706), cada uno de los cuales se incorpora de manera específica y completa como referencia en la presente descripción.

Una vez que los sitios objetivo para inactivar un gen candidato (por ejemplo, gen(es) *FusR1* sensible(s) a FW endógeno(s)) se tamizan in silico y se seleccionan, se construirán vectores CRISPR/Cas9 para las mutaciones objetivo en el gen candidato que se encuentran en las plantas de interés para la transformación de los vectores en la planta de interés (es decir, variedades de plátano sensibles a FW, tales como la variedad Cavendish estéril, triploide, ampliamente cultivada, y su progenie).

Los vectores CRISPR/Cas9 se transformarán en plantas de interés tales como variedades de plátano, especialmente plátanos sensibles a FW mediante el uso de protocolos mediados por agrobacterias que se conocen en la técnica (ver, por ejemplo, Ma y otros, 2015) y/o desarrollados o refinados por el inventor. En consecuencia, se realizará el cultivo de tejidos y la regeneración de plantas transformadas.

Las plantas transformadas con los vectores CRISPR/Cas9 se regenerarán y someterán a prueba para verificar la introducción de los vectores CRISPR/Cas9 en las células vegetales de interés. Como control para la inducción de indeles, también se usará en este experimento un constructo que exprese Cas9 de tipo silvestre.

La inactivación del(de los) gen(es) candidato(s) se examinará en todas las plantas transformadas. La inactivación se estudiará mediante (1) PCR cuantitativo para comprobar la supresión y/o silenciamiento del gen candidato o (2) amplificación por PCR y posterior secuenciación de Sanger y/o secuenciación profunda de alto rendimiento.

Además, la sustitución o sustituciones de aminoácidos provocadas por el cambio de marco introducido a la región del genoma objetivo se analizará mediante secuenciación de proteínas con espectrometría de masas.

Las plantas transformadas obtenidas se cultivarán en el invernadero controlado y/o en condiciones de campo. Se observará en las plantas transformadas, verificadas con la inserción, delección o sustitución de aminoácidos de interés, la resistencia mejorada a FW, el mal de Panamá o la infección por *Fusarium oxysporum f. sp. cubense* raza tropical 4.

Ejemplo 5: Transformación del plátano

Las plantas de plátano susceptibles a *Fusarium oxysporum* raza 4 (también conocida como raza tropical 4 o TR4) pueden transformarse en plantas resistentes a TR4 al

transformarlas con una secuencia de nucleótidos que codifica la resistencia mediante el uso de las tecnologías de transformación de plátano proporcionadas en el **Ejemplo 4** y las secuencias de nucleótidos de *FusR1* que codifican la resistencia a TR4 como se proporciona en la presente descripción. Por ejemplo, un cultivar de plátano Cavendish susceptible a TR4 puede transformarse con uno de los alelos de *FusR1* que codifican para la resistencia a TR4 como se proporciona en la presente descripción. Como ejemplo adicional, un cultivar de plátano Cavendish susceptible a TR4 puede transformarse con una o más de las siguientes secuencias codificantes de nucleótidos que codifican la resistencia a TR4: SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 9 SEQ ID NO: 11, SEQ ID NO: 18, SEQ ID NO: 21 y/o SEQ ID NO: 24.

Por ejemplo, el cultivar de plátano Cavendish 'Grand Nain' (AAA) puede transformarse con la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 9 y/o SEQ ID NO: 11 mediante el uso de los protocolos de transformación establecidos en la Patente de los Estados Unidos núm. 7,534,930 ('Transgenic Disease Resistant Banana'), que se incorpora en la presente descripción en su totalidad para todo lo que describe.

En resumen, las flores masculinas inmaduras de un cultivar de plátano Cavendish, tal como 'Grand Nain' o 'Williams', se usan para producir callos embriogénicos. Se construye un constructo de ácido nucleico que comprende la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 18, SEQ ID NO: 21 y/o SEQ ID NO: 24, unido operativamente a una secuencia promotora 35S. O, de manera alternativa, la secuencia promotora del alelo 1 resistente a FW de *FusR1* de *M. acuminata* (SEQ ID NO: 31) podría usarse para impulsar la expresión de los alelos de resistencia. Este constructo se introduce en los callos embriogénicos mediante el uso de bombardeo de microproyectiles. Las plántulas bombardeadas se regeneran a partir de los callos embriogénicos y las plántulas se someten a análisis de PCR para determinar qué plántulas se transformaron con el gen o genes de resistencia a TR4. Los extractos de cultivo de tejidos de las plantas resultantes que expresan positivamente el gen o genes de resistencia a TR4 se someten a prueba para determinar su capacidad para suprimir el crecimiento de TR4. Además, las supuestas plantas transformadas se someten a prueba para determinar su resistencia a TR4. Las plantas resistentes a TR4 se aíslan y clonan. Las plantas resistentes a TR4 pueden usarse en programas de mejoramiento para transferir los genes resistentes como se establece en el **Ejemplo 3**.

Cuando una planta transformada expresa la SEQ ID NO: 2 o la SEQ ID NO: 5; y también expresa la SEQ ID NO: 9 o la SEQ ID NO: 11, esa planta transformada habría apilado genes de resistencia a TR4 dado que comprende dos ácidos nucleicos diferentes que

codifican la resistencia a TR4. Como se analizó anteriormente y se presentó en la Tabla 1, la SEQ ID NO: 2 y la SEQ ID NO: 5 son secuencias codificantes del alelo 1 y del alelo 2 de *FusR1*, respectivamente, que codifican la resistencia obtenida de *M. itinerans*. Por el contrario, la SEQ ID NO: 9 y la SEQ ID NO: 11 son secuencias codificantes del alelo 1 y del alelo 2 de *FusR1*, respectivamente, que codifican la resistencia obtenida de *M. acuminata ssp. banksia*. Por tanto, una planta transformada que expresa ambos tipos de genes de resistencia tendría una resistencia apilada, o piramidal, a la raza tropical 4 del mal de Panamá.

Todas las plantas de plátano resultantes/seleccionadas con resistencia a TR4 pueden mantenerse mediante reproducción asexual y usarse para la producción o en programas de mejoramiento posteriores.

Ejemplo 6: Transformación del plátano con inicio con una variedad que comprende resistencia

Pueden producirse plantas de plátano transformadas resistentes a la raza tropical 4 del mal de Panamá mediante el uso de los procedimientos descritos en el Ejemplo 5, donde la planta inicial no transformada también tiene resistencia a TR4 y/o a una o más enfermedades adicionales. De esta forma, la planta transformada resultante puede tener genes de resistencia múltiples o apilados. Por ejemplo, el cultivar de partida usado en los procedimientos de transformación del Ejemplo 5 puede ser un cultivar Cavendish con el gen de resistencia RGA2 (Dale y otros, 2017). Por tanto, un cultivar Cavendish que comprende la secuencia codificante de RGA2 puede transformarse para expresar la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 9 y/o SEQ ID NO: 11, SEQ ID NO: 18, SEQ ID NO: 21 y/o SEQ ID NO: 24 y, de esta manera, tiene genes de resistencia a TR4 apilados.

Todas las plantas de plátano resultantes/seleccionadas con resistencia a TR4 pueden mantenerse mediante reproducción asexual y usarse para la producción o en programas de mejoramiento posteriores.

Ejemplo 7: Inactivación de la expresión de genes de susceptibilidad de *FusR1*

Además de o, de manera alternativa en lugar de, transformar las plantas de acuerdo con el **Ejemplo 5** o el **Ejemplo 6**, las secuencias de nucleótidos de los alelos de *FusR1* que codifican la susceptibilidad a TR4 en *M. acuminata* (por ejemplo, SEQ ID NO: 14) pueden inactivarse mediante el uso de una TALEN, una meganucleasa, una nucleasa con dedos de zinc, una nucleasa asociada a CRISPR u otras herramientas apropiadas de edición de genes.

En uno de dichos métodos, puede usarse un ARN guía junto con una nucleasa asociada a CRISPR apropiada, lo que incluye en donde el ARN guía comprende un dominio de direccionamiento variable que es complementario a la totalidad o a una secuencia parcial de SEQ ID NO: 14. Por ejemplo, puede introducirse una ruptura de doble cadena en una secuencia endógena que codifica un alelo de *FusR1* sensible a FW en *M. acuminata* (SEQ ID NO: 14) en una célula de plátano mediante el uso de una SEQ ID NO: 14 modificada, en donde la SEQ ID NO: 14 comprende una alteración de ácido nucleico que inactiva la función génica de SEQ ID NO: 14.

Para obtener detalles sobre cómo construir y usar dicha nucleasa asociada a CRISPR y ARN guía en plantas, ver, por ejemplo, la Publicación de Solicitud de Patente de los Estados Unidos núm. 2019/0032070 A1 y el documento WO 2019/118342 A1, cada uno de los cuales se incorpora como referencia en su totalidad. Para usar CRISPR como una herramienta de edición de genes en plátano, lo que incluye para silenciar genes de susceptibilidad a enfermedades, ver, por ejemplo, el documento WO 2018/220581 A1 (Compositions and Methods for Increasing Shelf-Life of Banana); Tripahi y otros, 2019, CRISPR/Cas9 editing of endogenous banana streak virus in the B genome of *Musa* spp. overcomes a major challenge in banana breeding, *Communications Biology* 2, artículo 46, 11 páginas; y, Ntui y otros, enero de 2020, Robust CRISPR/Cas9 mediated genome editing tool for banana and plantain (*Musa* spp.), vol. 21, 10 páginas.

La célula vegetal modificada puede generarse/regenerarse en una ~~planta de plátano~~ [planta de banano](#) que puede mantenerse mediante reproducción asexual.

Todas las plantas de plátano resultantes/seleccionadas con la inactivación génica para la susceptibilidad a TR4 pueden mantenerse mediante reproducción asexual y usarse para la producción o en programas de mejoramiento posteriores.

Ejemplo 8: Edición de genes de plátanos susceptibles a TR4

Las plantas de plátano susceptibles a *Fusarium oxysporum* raza 4 (también conocida como raza tropical 4 o TR4) pueden modificarse en plantas resistentes a TR4 mediante el uso de herramientas de edición de genes/selección de genes para cambiar sus secuencias de ácidos nucleicos endógenas que codifican la susceptibilidad en secuencias de nucleótidos que codifican la resistencia mediante el uso de las tecnologías de edición de genes de plátano proporcionadas en el **Ejemplo 4** y las secuencias de nucleótidos de *FusR1* que codifican la resistencia a TR4 como se proporciona en la presente descripción. Por ejemplo, la secuencia de ácido nucleico endógena que codifica la susceptibilidad a TR4 en un cultivo de plátano Cavendish puede alterarse basado en la secuencia de ácido nucleico de uno de los alelos de *FusR1*

que codifica la resistencia a TR4 como se proporciona en la presente descripción. Como ejemplo adicional, la secuencia de ácido nucleico que codifica la susceptibilidad a TR4 en un cultivar de plátano Cavendish puede alterarse basado en una o más de las siguientes secuencias de codificación de nucleótidos que codifican la resistencia a TR4:

5 SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 18, SEQ ID NO: 21 y/o SEQ ID NO: 24.

Por ejemplo, el cultivar de plátano Cavendish 'Grand Nain' (AAA) puede modificarse basado en las secuencias de ácido nucleico que codifican la resistencia a TR4 como se establece en la presente descripción (es decir, basado en la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 18, SEQ ID NO: 21 y/o SEQ ID NO: 24) mediante el uso de herramientas modernas de edición de genes. Ver la Figura 1.

En algunos ejemplos generales, el gen *FusR1* de susceptibilidad a FW endógeno de SEQ ID NO: 14 se modifica mediante uno o más de los siguientes cambios basado en su alineamiento con genes *FusR1* resistentes a FW de SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 11, SEQ ID NO: 18, SEQ ID NO: 21 y/o SEQ ID NO: 24. Ver la Figura 1.

En algunos ejemplos específicos, la SEQ ID NO: 14 se modifica mediante los siguientes cambios basado en su alineamiento con la SEQ ID NO: 9 (ver alineamiento de secuencia, Figura 1): la T correspondiente a la posición 148 se reemplaza por G (148T>G); la T correspondiente a la posición 323 se reemplaza por A (323T>A); la G correspondiente a la posición 344 se reemplaza por C (344G>C); y/o, la A correspondiente a la posición 347 se reemplaza por T (347A>T). En un ejemplo, la única sustitución realizada es 344G>C. En un ejemplo, se realizan las siguientes tres sustituciones: 323T>A, 344G>C y 347A>T. En otro ejemplo más, se realizan las cuatro sustituciones: 148T>G, 323T>A, 344G>C y 347A>T. Ver la Figura 1.

En algunos ejemplos generales, todas y cada una de las sustituciones de ácidos nucleicos se realizan en las secuencias de ácidos nucleicos que codifican proteínas FUSR1 susceptibles a FW de modo que los ácidos nucleicos modificados resultantes codifican proteínas FUSR1 resistentes a FW. Ver la Figura 1 y la Figura 2.

En algunos ejemplos específicos, la secuencia de ácido nucleico endógeno que codifica la proteína FUSR1 susceptible a FW de SEQ ID NO: 15 se modifica por uno o más cambios de ácido nucleico basado en su alineamiento con la proteína FUSR1 resistente a FW de SEQ ID NO: 12 para producir los siguientes cambios de proteína: la leucina correspondiente a la posición 50 se reemplaza por Valina (50L>V); la valina correspondiente a la posición 108 se reemplaza por ácido glutámico (108V>E); la

arginina en la posición 115 se reemplaza por prolina (115R>P); y/o, el ácido aspártico en la posición 116 se reemplaza por valina (116D>V). En un ejemplo, la única sustitución de proteínas que se realiza es 115R>P. En otro ejemplo, las únicas sustituciones de proteínas que se realizan son 108V>E, 115R>P y 116D>V. En otro ejemplo más, se realizan las cuatro sustituciones de proteínas: 50L>V, 108V>E, 115R>P y 116D>V. Ver la Figura 2.

Los protocolos de edición de genes específicos de plátano de las siguientes publicaciones proporcionan los protocolos para realizar las sustituciones necesarias de pares de bases de nucleótidos en plátano: Shao y otros, 2020, Using CRISPR/Cas9 genome editing system to create MaGA20ox2 gene-modified semi-dwarf banana, *Plant Biotechnology Journal*, 18:17-19; Kaur y otros, 2017, CRISPR/Cas9-mediated efficient editing in phytoene desaturase (PDS) demonstrates precise manipulation in banana cv. Rasthali genome, *Functional & Integrative Genomics*, 18(1):89-99; Otang y otros, 2020, Robust CRISPR/Cas9 mediated genome editing tool for banana and plantain (*Musa* spp.), *Current Plant Biology*, 21, 10 páginas; Tripathi y otros, 2019, CRISPR/Cas9 editing of endogenous banana streak virus in the B genome of *Musa* spp. overcomes a major challenge in banana breeding, *Communications Biology*, 2:46, 11 páginas; y la Patente de los Estados Unidos núm. 7,381,556, cada una de las cuales se incorpora en su totalidad como referencia en la presente descripción para todo lo que enseña.

En resumen, las flores masculinas inmaduras de un cultivar de plátano Cavendish, tal como 'Grand Nain' o 'Williams', se usan para producir callos embriogénicos y/o una suspensión de células embriogénicas. Se prepara un constructo por CRISPR/Cas9 de acuerdo con los procedimientos descritos en una o más de las publicaciones científicas y de patentes enumeradas anteriormente, en donde el constructo se construye basado en los alineamientos de secuencia proporcionados en la Figura 1. El constructo se suministra en los callos embriogénicos o en la suspensión de células embriogénicas y se generan plántulas bien enraizadas. Los regenerados aleatorios se seleccionan y tamizan para detectar la presencia del gen Cas9 por PCR mediante el uso de cebadores. Las plántulas bien enraizadas de los eventos positivos para PCR de Cas9 y las plantas de control se aclimatan y se colocan en macetas en el invernadero. Se realizan análisis moleculares para confirmar la edición de genes en los genes endógenos *FusR1*.

Las plantas con el genoma editado y las plantas de control se evalúan para determinar los rasgos agronómicos y se evalúa la resistencia a TR4. Se clonan las plantas resultantes con los genes editados que expresan positivamente la(s) proteína(s) de resistencia a TR4 y muestran resistencia a TR4. Las plantas resistentes a TR4 con los

genes editados pueden usarse en programas de mejoramiento para transferir los genes resistentes como se establece en el **Ejemplo 3**.

Todas las plantas de plátano resultantes/seleccionadas con resistencia a TR4 pueden mantenerse mediante reproducción asexual y usarse para la producción o en programas de mejoramiento posteriores.

Otras modalidades numeradas de la descripción

Otra materia objeto contemplada por la presente invención se expone en las siguientes modalidades numeradas:

1. Una molécula de ácido nucleico aislada que comprende la secuencia de ácido nucleico de SEQ ID NO: 14 que codifica la susceptibilidad a *Fusarium oxysporum* raza 4 cuando se expresa en una planta, en donde la SEQ ID NO: 14 está modificada por una, dos, tres o cuatro sustituciones de ácido nucleico de modo que la secuencia de ácido nucleico resultante codifica la resistencia a *Fusarium oxysporum* raza 4 cuando se expresa en una planta.
2. La molécula de ácido nucleico aislada de la modalidad 1, en donde las sustituciones de ácido nucleico comprenden reemplazar una T correspondiente a la posición 148 de la SEQ ID NO: 14 por una G (148T>G).
3. La molécula de ácido nucleico aislada de la modalidad 1, en donde las sustituciones de ácido nucleico comprenden reemplazar una T correspondiente a la posición 323 de la SEQ ID NO: 14 por una A (323T>A).
4. La molécula de ácido nucleico aislada de la modalidad 1, en donde las sustituciones de ácido nucleico comprenden reemplazar una G correspondiente a la posición 344 de la SEQ ID NO: 14 por una C (344G>C).
5. La molécula de ácido nucleico aislada de la modalidad 1, en donde las sustituciones de ácido nucleico comprenden reemplazar una A correspondiente a la posición 347 de la SEQ ID NO: 14 por una T (347A>T).
6. La molécula de ácido nucleico aislada de la modalidad 1, en donde las sustituciones de ácido nucleico comprenden reemplazar una T correspondiente a la posición 323 por una A (323T>A), reemplazar una G correspondiente a la posición 344 por una C (344G>C) y reemplazar una A correspondiente a la posición 347 por una T (347A>T), y en donde todas las posiciones se basan en la SEQ ID NO: 14.
7. La molécula de ácido nucleico aislada de la modalidad 1, en donde la SEQ ID NO: 14 codifica una secuencia de aminoácidos de SEQ ID NO: 15 y en donde las sustituciones de ácido nucleico dan como resultado el reemplazo de una leucina correspondiente a la posición 50 de la SEQ ID NO: 15 por una valina (50L>V).

8. La molécula de ácido nucleico aislada de la modalidad 1, en donde la SEQ ID NO: 14 codifica una secuencia de aminoácidos de la SEQ ID NO: 15 y en donde las sustituciones de ácido nucleico dan como resultado el reemplazo de una valina correspondiente a la posición 108 de la SEQ ID NO: 15 por un ácido glutámico (108V>E).
- 5 9. La molécula de ácido nucleico aislada de la modalidad 1, en donde la SEQ ID NO: 14 codifica una secuencia de aminoácidos de SEQ ID NO: 15 y en donde las sustituciones de ácido nucleico dan como resultado el reemplazo de una arginina correspondiente a la posición 115 de la SEQ ID NO: 15 por una prolina (115R>P).
- 10 10. La molécula de ácido nucleico aislada de la modalidad 1, en donde la SEQ ID NO: 14 codifica una secuencia de aminoácidos de SEQ ID NO: 15 y en donde las sustituciones de ácido nucleico dan como resultado el reemplazo de un ácido aspártico correspondiente a la posición 116 de la SEQ ID NO: 15 por una valina (116D>V).
- 15 11. La molécula de ácido nucleico aislada de la modalidad 1, en donde la SEQ ID NO: 14 codifica una secuencia de aminoácidos de SEQ ID NO: 15 y en donde las sustituciones de ácido nucleico dan como resultado el reemplazo de una valina correspondiente a la posición 108 de la SEQ ID NO: 15 por un ácido glutámico (108V>E), una arginina correspondiente a la posición 115 de SEQ ID NO: 15 por una prolina (115R>P) y un ácido aspártico correspondiente a la posición 116 de la SEQ ID NO: 15 por una valina (116D>V).
- 20 12. La molécula de ácido nucleico aislada de las modalidades 1-11, en donde la expresión se produce en una célula vegetal, tejido vegetal, cultivo de células vegetales, cultivo de tejido vegetal o planta completa.
13. La molécula de ácido nucleico aislada de la modalidad 12, en donde la expresión se produce en una célula, tejido, cultivo celular, cultivo de tejido o planta completa de *Musa*.
- 25 14. La molécula de ácido nucleico aislada de la modalidad 13, en donde la expresión se produce en una célula, tejido, cultivo celular, cultivo de tejido o planta completa de *Musa acuminata*.
15. Un constructo de ácido nucleico que comprende la molécula de ácido nucleico aislada de las modalidades 1-11, en donde la secuencia de ácido nucleico está unida operativamente a un promotor capaz de dirigir la expresión de la secuencia de ácido nucleico.
- 30 16. El constructo de ácido nucleico de la modalidad 15, en donde el promotor es un promotor vegetal.
17. El constructo de ácido nucleico de la modalidad 15, en donde el promotor es un promotor 35S.
- 35

18. El constructo de ácido nucleico de la modalidad 15, en donde el promotor es codificado por la SEQ ID NO: 31.
19. Un vector de transformación que comprende el constructo de ácido nucleico de las modalidades 15-18.
- 5 20. Un método para transformar una célula vegetal que comprende introducir el vector de transformación de la modalidad 19 en una célula vegetal, por lo cual la célula vegetal transformada expresa la secuencia de ácido nucleico que codifica la resistencia a *Fusarium oxysporum* raza 4.
21. El método de la modalidad 20, en donde la célula vegetal es una célula vegetal de
- 10 *Musa*.
22. El método de la modalidad 20, en donde la célula vegetal es una célula vegetal de *Musa acuminata*.
23. El método de las modalidades 20-22, que comprende además producir tejido vegetal transformado a partir de la célula vegetal transformada.
- 15 24. El método de la modalidad 23, que comprende además producir una plántula transformada a partir del tejido vegetal transformado.
25. El método de la modalidad 24, que comprende además producir un clon de la plántula transformada.
26. El método de las modalidades 24 o 25, que comprende además hacer crecer la
- 20 plántula transformada o el clon de la plántula transformada en una planta transformada madura.
27. El método de la modalidad 26, en donde la planta transformada madura es una planta de *Musa* y la planta de *Musa* transformada madura es capaz de producir frutos.
28. El método de la modalidad 27, que comprende además producir clones de la planta
- 25 de *Musa* transformada madura.
29. El método de la modalidad 27 o 28, que comprende además el uso de la planta de *Musa* transformada madura o el clon de la planta de *Musa* transformada madura en un método de mejoramiento.
30. Una molécula de aminoácidos aislada que comprende una secuencia de
- 30 aminoácidos de SEQ ID NO: 15 que codifica una proteína que cuando se produce en una planta da como resultado susceptibilidad a *Fusarium oxysporum* raza 4, en donde la SEQ ID NO: 15 está modificada por uno, dos, tres o cuatro sustituciones de aminoácidos de modo que codifica una proteína que cuando se produce en una planta da como resultado resistencia a *Fusarium oxysporum* raza 4.

31. La molécula de aminoácido aislada de la modalidad 30, en donde las sustituciones de aminoácidos comprenden reemplazar una leucina correspondiente a la posición 50 de la SEQ ID NO: 15 por una valina (50L>V).
- 5 32. La molécula de aminoácido aislada de la modalidad 30, en donde las sustituciones de aminoácidos comprenden reemplazar una valina correspondiente a la posición 108 de la SEQ ID NO: 15 por un ácido glutámico (108V>E).
33. La molécula de aminoácido aislada de la modalidad 30, en donde las sustituciones de aminoácidos comprenden reemplazar una arginina correspondiente a la posición 115 de la SEQ ID NO: 15 por una prolina (115R>P).
- 10 34. La molécula de aminoácido aislada de la modalidad 30, en donde las sustituciones de aminoácidos comprenden reemplazar un ácido aspártico correspondiente a la posición 116 de la SEQ ID NO: 15 por una valina (116D>V).
35. La molécula de aminoácido aislada de la modalidad 30, en donde las sustituciones de aminoácidos comprenden reemplazar una valina correspondiente a la posición 108 de la SEQ ID NO: 15 por un ácido glutámico (108V>E), una arginina correspondiente a la posición 115 de la SEQ ID NO: 15 por una prolina (115R>P), y un ácido aspártico correspondiente a la posición 116 de la SEQ ID NO: 15 por una valina (116D>V).
- 15 36. El segmento de molécula de aminoácido aislada de las modalidades 30-35, en donde la producción se produce en una célula vegetal, tejido vegetal, cultivo de células vegetales, cultivo de tejido vegetal o planta completa.
- 20 37. El segmento de molécula de aminoácido aislada de la modalidad 36, en donde la producción se produce en una célula, tejido, cultivo celular, cultivo de tejido o planta completa de *Musa*.
38. El segmento de molécula de aminoácido aislada de la modalidad 36, en donde la producción se produce en una célula, tejido, cultivo celular, cultivo de tejido o planta completa de *Musa acuminata*.
- 25 39. Un constructo de ácido nucleico que comprende una secuencia de ácido nucleico que codifica la resistencia a *Fusarium oxysporum* raza 4 cuando se expresa en una planta, en donde dicha secuencia de ácido nucleico se selecciona del grupo que consiste en la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 18, SEQ ID NO: 21, y SEQ ID NO: 24, y en donde la secuencia de ácido nucleico está unida operativamente a un promotor capaz de dirigir la expresión de la secuencia de ácido nucleico.
- 30 40. El constructo de ácido nucleico de la modalidad 39, en donde el promotor es un promotor vegetal.
- 35

41. El constructo de ácido nucleico de la modalidad 39, en donde el promotor es un promotor 35S.
42. El constructo de ácido nucleico de la modalidad 39, en donde el promotor es codificado por la SEQ ID NO: 31.
- 5 43. Un vector de transformación que comprende el constructo de ácido nucleico de las modalidades 39-42.
44. Un método para transformar una célula vegetal que comprende introducir el vector de transformación de la modalidad 43 en una célula vegetal, por lo cual la célula vegetal transformada expresa la secuencia de ácido nucleico que codifica la resistencia a
- 10 *Fusarium oxysporum* raza 4.
45. El método de la modalidad 44, en donde la célula vegetal es una célula vegetal de *Musa*.
46. El método de la modalidad 44, en donde la célula vegetal es una célula vegetal de *Musa acuminata*.
- 15 47. El método de las modalidades 44 - 46, que comprende además producir tejido vegetal transformado a partir de la célula vegetal transformada.
48. El método de la modalidad 47, que comprende además producir una plántula transformada a partir del tejido vegetal transformado.
49. El método de la modalidad 48, que comprende además producir un clon de la
- 20 plántula transformada.
50. El método de las modalidades 48 o 49, que comprende además hacer crecer la plántula transformada o el clon de la plántula transformada en una planta transformada madura.
51. El método de la modalidad 50, en donde la planta transformada madura es una
- 25 planta de *Musa* y la planta de *Musa* transformada madura es capaz de producir frutos.
52. El método de la modalidad 51, que comprende además producir clones de la planta de *Musa* transformada madura.
53. El método de las modalidades 51 o 52, que comprende además el uso de la planta de *Musa* transformada madura o el clon de la planta de *Musa* transformada madura en
- 30 un método de mejoramiento.
54. Un método de mejoramiento de plátano que comprende cruzar una primera planta de *Musa* que comprende una secuencia de ácido nucleico que codifica la resistencia a *Fusarium oxysporum* raza 4 con una segunda planta de *Musa* que es susceptible a *Fusarium oxysporum* raza 4 y seleccionar la progenie resultante del cruzamiento basado
- 35 en su resistencia a *Fusarium oxysporum* raza 4, en donde dicha secuencia de ácido

nucleico que codifica la resistencia a *Fusarium oxysporum* raza 4 se selecciona del grupo que consiste en la SEQ ID NO: 2, SEQ ID NO: 5, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 18, SEQ ID NO: 21, y SEQ ID NO: 24.

5 55. El método de mejoramiento de plátano de la modalidad 54, que comprende además producir clones de la progenie resultante del cruzamiento en donde los clones se seleccionan basado en su resistencia a *Fusarium oxysporum* raza 4.

56. El método de mejoramiento de plátano de la modalidad 54, en donde la primera y segunda plantas de *Musa* son de diferentes especies de *Musa*. El método de mejoramiento de plátano de la modalidad 54, en donde la primera y segunda plantas de
10 *Musa* son de la misma especie de *Musa*. El método de mejoramiento de plátano de la modalidad 54, en donde la primera y/o la segunda planta de *Musa* es una planta de *Musa acuminata*.

57. El método de mejoramiento de plátano de la modalidad 54, en donde la progenie del cruzamiento que muestra resistencia a *Fusarium oxysporum* raza 4 se selecciona
15 mediante el uso de marcadores moleculares que están diseñados basado en la secuencia de ácido nucleico que codifica la resistencia a *Fusarium oxysporum* raza 4 que está presente en la primera planta de *Musa* usada en el cruzamiento.

58. Un método para obtener una célula vegetal de *Musa acuminata* con un gen endógeno silenciado que codifica la susceptibilidad a *Fusarium oxysporum* raza 4,
20 donde el método comprende introducir una ruptura de doble cadena en al menos un sitio en un gen exógeno codificado por la SEQ ID NO: 14 para producir una célula vegetal de *Musa acuminata* con un gen endógeno silenciado que codifica la susceptibilidad a *Fusarium oxysporum* raza 4.

59. El método de la modalidad 58 comprende además generar una planta de *Musa acuminata* a partir de la célula vegetal de *Musa acuminata* con un gen endógeno
25 silenciado que codifica la susceptibilidad a *Fusarium oxysporum* raza 4 para producir una planta de *Musa acuminata* con un gen endógeno silenciado que codifica la susceptibilidad a *Fusarium oxysporum* raza 4.

60. El método de la modalidad 59, que comprende además el uso de la planta de *Musa acuminata* con un gen endógeno silenciado que codifica la susceptibilidad a *Fusarium*
30 *oxysporum* raza 4 en un programa de mejoramiento de plátano.

61. El método de la modalidad 20 o 44, en donde la célula vegetal es la célula vegetal de *Musa acuminata* de la modalidad 59 con un gen endógeno silenciado que codifica la susceptibilidad a *Fusarium oxysporum* raza 4.

62. El método de la modalidad 58, en donde la ruptura de doble cadena es inducida por una nucleasa seleccionada del grupo que consiste en una TALEN, una meganucleasa, una nucleasa con dedos de zinc y una nucleasa asociada a CRISPR.
63. El método de conformidad con la reivindicación 62, en donde la ruptura de doble cadena es inducida por una nucleasa asociada a CRISPR y donde se proporciona un ARN guía.
64. Un método para producir una célula vegetal resistente a *Fusarium oxysporum* raza 4 que comprende introducir al menos una modificación genética en una o más secuencias de ácido nucleico endógenas que codifican la susceptibilidad a *Fusarium oxysporum* raza 4, en donde la modificación genética confiere resistencia a *Fusarium oxysporum* raza 4 a la célula vegetal.
65. El método de la modalidad 64, en donde la al menos una modificación genética se introduce mediante una TALEN, una meganucleasa, una nucleasa con dedos de zinc o una nucleasa asociada a CRISPR.
66. El método de la reivindicación 64, en donde la al menos una modificación genética se introduce mediante una nucleasa asociada a CRISPR y un ARN guía asociado.
67. El método de la modalidad 64, en donde la al menos una modificación genética se selecciona de la lista que consiste en reemplazar una T correspondiente a la posición 148 de la SEQ ID NO: 14 por una G (148T>G), reemplazar una T correspondiente a la posición 323 de la SEQ ID NO: 14 por una A (323T>A), reemplazar una G correspondiente a la posición 344 de la SEQ ID NO: 14 por una C (344G>C), y reemplazar una A correspondiente a la posición 347 de la SEQ ID NO: 14 por una T (347A>T).
68. El método de la modalidad 64, en donde la al menos una modificación genética da como resultado un cambio en un aminoácido seleccionado del grupo que consiste en reemplazar una leucina correspondiente a la posición 50 de la SEQ ID NO: 15 por una valina (50L>V), reemplazar una valina correspondiente a la posición 108 de la SEQ ID NO: 15 por un ácido glutámico (108V>E), reemplazar una arginina correspondiente a la posición 115 de la SEQ ID NO: 15 por una prolina (115R>P), y reemplazar un ácido aspártico correspondiente a la posición 116 de la SEQ ID NO: 15 por una valina (116D>V).
69. El método de las modalidades 64-68, en donde la célula vegetal es una célula vegetal de *Musa*.
70. El método de las modalidades 64-68, en donde la célula vegetal es una célula vegetal de *Musa acuminata*.

71. El método de las modalidades 64-70, que comprende además producir tejido vegetal transformado a partir de la célula vegetal transformada.
72. El método de la modalidad 71, que comprende además producir una plántula transformada a partir del tejido vegetal transformado.
- 5 73. El método de la modalidad 72, que comprende además producir un clon de la plántula transformada.
74. El método de las modalidades 71 o 72, que comprende además hacer crecer la plántula transformada o el clon de la plántula transformada en una planta transformada madura.
- 10 75. El método de la modalidad 74, en donde la planta transformada madura es una planta de *Musa* y la planta de *Musa* transformada madura es capaz de producir frutos.
76. El método de la modalidad 75, que comprende además producir clones de la planta de *Musa* transformada madura.
77. El método de las modalidades 75 o 76, que comprende además el uso de la planta
- 15 de *Musa* transformada madura o el clon de la planta de *Musa* transformada madura en un método de mejoramiento.

INCORPORACIÓN COMO REFERENCIA

- Todas las referencias, artículos, publicaciones, patentes, publicaciones de patentes y
- 20 solicitudes de patente citadas en la presente descripción dentro del texto anterior y/o citadas más abajo se incorporan como referencia en su totalidad para todos los propósitos. Sin embargo, la mención de cualquier referencia, artículo, publicación, patente, publicación de patente y solicitud de patente citada en la presente descripción no es, y no debe tomarse como reconocimiento o cualquier forma de sugerencia de que
- 25 constituyen una técnica anterior válida o forman parte del conocimiento general común en cualquier país del mundo.

DOCUMENTOS DE PATENTE DE LOS ESTADOS UNIDOS

- 7,534,930 B2 5/2009 Vishnevetsky y otros
- 6,274,319 8/2001 Messier y Sikela
- 30 9,834,783 12/2017 Messier

OTRAS PUBLICACIONES

- Armenteros, J.J.A.A. 2017. DeepLoc: prediction of protein subcellular localization using deep learning. *Bioinformatics* **33**(21):3387–3395.
- Armenteros y otros, 2019. SignalP 5.0 improves signal peptide predictions using deep
- 35 neural networks. *Nat Biotechnol* **37**:420–423.

- Bai, T-T. y otros, 2013. Transcriptome and Expression Profile Analysis of Highly Resistant and Susceptible Banana Roots Challenged with *Fusarium oxysporum* f. sp. *cubense* Tropical Race 4. *PLOS | One* Published: 23 de septiembre, 2013.
- Barbosa, J.A.R.G. y otros, 2007. Crystal Structure of the Bowman-Birk Inhibitor from *Vigna unguiculata* Seeds in Complex with β -Trypsin at 1.55 Å Resolution and Its Structural Properties in Association with Proteinases. *Biophysical Journal*. **92**(5): 1638-1650.
- Chen, A., y otros, 2019. Assessing Variations in Host Resistance to *Fusarium oxysporum* f. sp. *cubense* Race 4 in *Musa* Species, With a Focus on the Subtropical Race 4. *Front. Microbiol.* **10**.
- Christelová, P. y otros, 2017. Molecular and cytological characterization of the global *Musa* germplasm collection provides insights into the treasure of banana diversity. *Biodivers. Conserv.* **26**: 801.
- Dale, J. y otros, 2017. Transgenic Cavendish bananas with resistance to *Fusarium* wilt tropical race 4. *Nature Communications*. **8**: Número de artículo 1496.
- Davey, M.W. y otros, 2013. A draft *Musa balbisiana* genome sequence for molecular genetics in polyploid, inter- and intra-specific *Musa* hybrids. *BMC Genomics* **14**: 683.
- D'Hont, A. y otros, 2012. The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. *Nature* **488**:213–217.
- Dita, M. y otros, 2018. *Fusarium* Wilt of banana: current knowledge on epidemiology and research needs toward sustainable disease management. *Front Plant Sci.* **9**:1468.
- Heslop-Harrison, J.S. y Schwarzacher, T. 2007. Domestication, Genomics and the Future for Banana. *Annals of Botany* **100**(5):1073–1084.
- Hiller, K. y otros, 2004. PrediSi: prediction of signal peptides and their cleavage positions. *Nucleic Acids Res.* **32**(emisión en servidor Web):W375-9.
- Hippolyte, I. y otros, 2012. Foundation characteristics of edible *Musa* triploids revealed from allelic distribution of SSR markers. *Annals of Botany* **109**(5):937–951.
- Hughes, A.L y Nei, M. 1988 *Nature* **335**:167-170.
- Ishihara y otros, 2016. An improved method for RNA extraction from woody legume species *Acacia koa* A. Gray and *Leucaena leucocephala* (Lam.) de Wit. *Int. J. For. Wood Sci.* **3**(1): 31-35.
- Kreitman, M. y Akashi, H. 1995. Molecular evidence for natural selection. *Annu. Rev. Ecol. Syst.* **26**:403-422.
- Kumar, S., y otros, 2018. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Molecular Biology and Evolution* **35**:1547-1549.

- Li, C.-Y. y otros, 2012. Transcriptome profiling of resistant and susceptible Cavendish banana roots following inoculation with *Fusarium oxysporum* f. sp. *cubense* tropical race 4. *BMC Genomics* **13**:374.
- Li, W.-H. y otros, 1985. A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Mol. Biol. Evol.* **2**: 150-174.
- Li, W.-H. 1993. Unbiased estimation of the rates of synonymous and nonsynonymous substitution. *J. Mol. Evol.* **36**: 9699.
- Li, W.-H., 1997. *Molecular Evolution*. Sunderland, Massachusetts: Sinauer Associates.
- Li, W.M. y otros, 2015. Resistance sources to *Fusarium oxysporum* f. sp. *cubense* tropical race 4 in banana wild relatives. *Plant Pathology* **64**:1061-1067.
- Ma, X, y otros, 2015 A Robust CRISPR/Cas9 System for Convenient, High-Efficiency Multiplex Genome Editing in Monocot and Dicot Plants. *Mol. Plant.* **8**:1274–1284.
- Messier, W. y Stewart, C.-B. 1994 *Current Biol.* **4**:911-913.
- Messier, W. y Stewart, C.-B. 1997. *Nature* **385**:151-154.
- Nei M. y Kumar S. 2000. *Molecular Evolution and Phylogenetics*. Oxford University Press, New York.
- Paul, J.-Y. y otros, 2011. Apoptosis-related genes confer resistance to Fusarium wilt in transgenic 'Lady Finger' bananas. *Plant Biotechnology Journal*.
- Niu, Y. y otros, 2018. Comparative digital gene expression analysis of tissue-cultured plantlets of highly resistant and susceptible banana cultivars in response to *Fusarium oxysporum*. *Int. J. Mol. Sci.* **19**. doi: 10.3390/ijms19020350.
- Peraza-Echeverria, S. y otros, 2009. Molecular cloning and *in silico* analysis of potential *Fusarium* resistance genes in banana. *Mol. Breeding.* **23**(3): 431-443.
- Ploetz, R.C. 2015. Fusarium Wilt of banana. *Phytopathology Review*.
- Raboin, L-M. y otros, 2005. Diploid Ancestors of Triploid Export Banana Cultivars: Molecular Identification of 2n Restitution Gamete Donors and n Gamete Donors. *Mol Breeding* **16**:333.
- Reese M.G. 2001. Application of a time-delay neural network to promoter annotation in the *Drosophila melanogaster* genome. *Comput. Chem.* **26**(1): 51-56.
- Ribeiro, L.R. y otros, 2018. Sources of resistance to *Fusarium oxysporum* f. sp. *cubense* in banana germplasm. *Rev. Bras. Frutic.* **40**:1. Epub 08 Feb, 2018.
- Rouard, M. y otros, 2018. Three New Genome Assemblies Support a Rapid Radiation in *Musa acuminata* (Wild Banana). *Genome Biology and Evolution* **10**(12): 3129–3140.

Con formato: Español (España)

- Solovyev V.V. y Salamov A.A. 1997. The Gene-Finder computer tools for analysis of human and model organisms genome sequences. En *Proceedings of the Fifth International Conference on Intelligent Systems for Molecular Biology* (eds. Rawling C., Clark D., Altman R., Hunter L., Lengauer T., Wodak S.), Halkidiki, Greece, AAAI Press, 294-302.
- Solovyev V.V. 2001. Statistical approaches in Eukaryotic gene prediction. En *Handbook of Statistical Genetics* (eds. Balding D. y otros), John Wiley & Sons, Ltd., p. 83-127.
- Solovyev V.V. y Shahmuradov I.A. 2003. PromH: Promoters identification using orthologous genomic sequences. *Nucleic Acids Res.* **31**(13): 3540-3545.
- Stokstad, E. 2019. Banana fungus puts Latin America on alert. *Science* **365**(6450): 207-208.
- Ssali, R. y otros, 2013. Inheritance of resistance to *Fusarium oxysporum* f. sp. *cubense* race 1 in bananas. *Euphytica* **194**: 425. Van der Berg, N. y otros, 2007. Tolerance in banana to *Fusarium* wilt is associated with early up-regulation of cell wall-strengthening genes in the roots. *Molecular Plant Pathology*. **8**(3): 333-341.
- Venkataaramana, R.K. y otros, 2015. Insights into *Musa balbisiana* and *Musa acuminata* species divergence and development of genic microsatellites by transcriptomics approach. *Plant Gene* **4**: 78-82.
- Wang, Y. y otros, 2017. Differential gene expression in banana roots in response to *Fusarium* wilt. *Canadian Journal of Plant Pathology* **39**(2): 163-175. doi.org/10.1080/07060661.2017.1342693.
- Wu, W. y otros, 2016. Whole genome sequencing of a banana wild relative *Musa itinerans* provides insights into lineage-specific diversification of the *Musa* genus. *Scientific Reports* **6**: Número de artículo: 31586.
- Zhang, L. y otros, (2018) Identification and evaluation of resistance to *Fusarium oxysporum* f. sp. *cubense* tropical race 4 in *Musa acuminata* Pahang. *Euphytica* **214**: 106.
- Zuo, C. y otros, 2018. Germplasm screening of *Musa* spp. for resistance to *Fusarium oxysporum* f. sp. *cubense* tropical race 4 (Foc-TR4). *Eur J Plant Pathol.* **151**:723.

Con formato: Español (España)

RESUMEN

La presente descripción proporciona composiciones y métodos para proporcionar una amplia resistencia a patógenos fúngicos, tales como un hongo *Fusarium*, y plantas derivadas de ellos.

5

10

15

FIGURA 1

Alelo 1 de <i>M. itinerans</i> (SEQ ID NO: 2)	atggctggaggaggcaaaagggtgaagcgctgctctcttctactttgtgac	
Alelo 1 resistente de <i>M. acuminata</i> (SEQ ID NO: 9)	atggctggaggaggcaaaagggtgaagcgctgctctcttctactttgtgac	
Alelo sensible de <i>M. acuminata</i> (SEQ ID NO: 14)	atggctggaggaggcaaaagggtgaagcgctgctctcttctactttgtgac	
Acc. de <i>M. balbisiana</i> ITC1016 (SEQ ID NO: 27)	atggctggaggaggcaaaagggtgaagcgctgctctcttctactttgtgac	
Alelo 1 de <i>M. basjoo</i> (SEQ ID NO: 18)	atggctggaggaggcaaaagggtgaagcgctgctctcttctactttgtgac	50
Alelo 1 de <i>M. itinerans</i> (SEQ ID NO: 2)	gctgctcgtgatgttggtggccttcttcgccaccgactcctcggcgcccc	
Alelo 1 resistente de <i>M. acuminata</i> (SEQ ID NO: 9)	gctgctcgtgatgttggtggccttcttcgccaccgactcctcggcgcccc	
Alelo sensible de <i>M. acuminata</i> (SEQ ID NO: 14)	gctgctcgtgatgttggtggccttcttcgccaccgactcctcggcgcccc	
Acc. de <i>M. balbisiana</i> ITC1016 (SEQ ID NO: 27)	gctgctcgtgatgttggtggccttcttcgccaccgactcctcggcgcccc	
Alelo 1 de <i>M. basjoo</i> (SEQ ID NO: 18)	gctgctcgtgatgttggtggccttcttcgccaccgactcctcggcgcccc	100
Alelo 1 de <i>M. itinerans</i> (SEQ ID NO: 2)	GTGTCAACACCCCGTCCGCACTCCCTCGCCAGAGCGGTACTGAGTGGGTTG	
Alelo 1 resistente de <i>M. acuminata</i> (SEQ ID NO: 9)	GTGTCAACACCCCGTCCGCAATCCCTCGCCAGAGCGGCACTGAGTGGGTTG	
Alelo sensible de <i>M. acuminata</i> (SEQ ID NO: 14)	GTGTCAACACCCCGTCCGCAATCCCTCGCCAGAGCGGCACTGAGTGGGTTG	
Acc. de <i>M. balbisiana</i> ITC1016 (SEQ ID NO: 27)	GTGTCAACACCCCGTCCGCACTCCCTCGCCAGAGCGGCACTGAGTGGGTTG	
Alelo 1 de <i>M. basjoo</i> (SEQ ID NO: 18)	GTGTCAACACCCCGTCCGCAATCCCTCGCCAGAGCGGCACTGAGTGGGTTG	150
Alelo 1 de <i>M. itinerans</i> (SEQ ID NO: 2)	GAGGCAAGGCGAGATGGGCGGTGTGCAGATGCATCTGTCTCTCATTTA	
Alelo 1 resistente de <i>M. acuminata</i> (SEQ ID NO: 9)	GAGGCAAGGCGAGATGAGCCGTGTGTGCAGATGCGCGTGTCTCTCATTTA	
Alelo sensible de <i>M. acuminata</i> (SEQ ID NO: 14)	GAGGCAAGGCGAGATGAGCCGTGTGTGCAGATGCGCGTGTCTCTCATTTA	
Acc. de <i>M. balbisiana</i> ITC1016 (SEQ ID NO: 27)	GAGGTAAGGCGAGATGCGCGGTGTGTGCATGCGTGTCTCTCTCATTTA	
Alelo 1 de <i>M. basjoo</i> (SEQ ID NO: 18)	GAGGCAAGGCGAGATGAGCCGTGTGTGCAGATGCGCGTGTCTCTCATTTA	200
	* * ** * * * * *	

FIGURA 1 continuación

Alelo 1 de <i>M. itinerans</i> (SEQ ID NO: 2)	CCACACTACTTGGTGGCTTTGGAGCGGCGTATGGCAAGGCTCTTGGCCCTT	250
Alelo 1 resistente de <i>M. acuminata</i> (SEQ ID NO: 9)	9) CCCACTACTTGGTGGCTTTGGCGCGCAATAGGCAAGGCTCTTGGCCCTT	
Alelo sensible de <i>M. acuminata</i> (SEQ ID NO: 14)	14) CCCACTACTTGGTGGCTTTGGCGCGCAATAGGCAAGGCTCTTGGCCCTT	
Acc. de <i>M. balbisiana</i> ITC1016 (SEQ ID NO: 27)	CCACACTCTCTTTTGGCTTTTGGCGGCGGTATGGCAAGGCTCTTGGCCCTT	
Alelo 1 de <i>M. basjoo</i> (SEQ ID NO: 18)	CCACACTACTTGGTGGCTTTGGCGGCGCAATAGGCAAGGCTCTTGGCCCTT	
	* ** * * * *	
Alelo 1 de <i>M. itinerans</i> (SEQ ID NO: 2)	CCGCTGCGACCAACTGGCGAGTGTCTCTCAACGAGTGCACCTTGCCTCGAT	
Alelo 1 resistente de <i>M. acuminata</i> (SEQ ID NO: 9)	9) CCGCTGCGACAACCTGGCGAGTGTCTCTCAACGAGTGCACCTTGCCTCGAT	
Alelo sensible de <i>M. acuminata</i> (SEQ ID NO: 14)	} CCGCTGCGACAACCTGGCGAGTGTCTCTCAACGAGTGCACCTTGCCTCGAT	
Acc. de <i>M. balbisiana</i> ITC1016 (SEQ ID NO: 27)	CCGCTGCGACCAACTGGCGAGTGTCTCTCAACGAGTGCACCTTGCCTCGAT	
Alelo 1 de <i>M. basjoo</i> (SEQ ID NO: 18)	CCGCTGCGACAACCTGGCGAGTGTCTCTCAACGAGTGCACCTTGCCTCGAT	300
	* * * * *	
Alelo 1 de <i>M. itinerans</i> (SEQ ID NO: 2)	CAGTGGACTACAGGCTGCCAGGCGCACTCTCTGTGGCTGGCTTGATGG	
Alelo 1 resistente de <i>M. acuminata</i> (SEQ ID NO: 9)	9) CTTATGGACCCCAAGGTCTGCCAGGCCCAACTCTCTGTCTCTGGCTGTGGC	
Alelo sensible de <i>M. acuminata</i> (SEQ ID NO: 14)	CTTATGGACCCCAAGGTCTGCCAGGCCCAACTCTCTGTCTCTGGCTGTGGC	
Acc. de <i>M. balbisiana</i> ITC1016 (SEQ ID NO: 27)	CGTGTGACCCCAAGGCTGCCAGGCGCACTCTCTGTGGC-CGCTCGATGC	
Alelo 1 de <i>M. basjoo</i> (SEQ ID NO: 18)	CTTATGGACCCCAAGGTCTGCCAGGCCCAACTCTCTGTCTCTGGCTGTGGC	350
	*** ** * * * *	
Alelo 1 de <i>M. itinerans</i> (SEQ ID NO: 2)	CGTCCCAAACTAGAGCGGTGCGACAGTGGCGCATCGAAGAAACCGGTG	
Alelo 1 resistente de <i>M. acuminata</i> (SEQ ID NO: 9)	9) AGCCCCCAAAAGTAGAGCCGCGCAGTGGGCTATCGAAGAAACCGGTG	
Alelo sensible de <i>M. acuminata</i> (SEQ ID NO: 14)	AGCCCCCAAAAGTAGAGCCGCGCAGTGGGCGCATCGAAGAAACCGGTG	
Acc. de <i>M. balbisiana</i> ITC1016 (SEQ ID NO: 27)	AGCCCCCAAAAGTag	
Alelo 1 de <i>M. basjoo</i> (SEQ ID NO: 18)	AGCCCCCAAAAGTAGAGCCGCGCAGTGGGCTATCGAAGAAACCGGTG	400
	* * * * *	

FIGURA 1 continuación

Alelo 1 de <i>M. itinerans</i> (SEQ ID NO: 2)	GGAAATTAGCGATGATGGTGTGA	
Alelo 1 resistente de <i>M. acuminata</i> (SEQ ID NO: 9)	GGAAATTAGCGATGATGGTGTGA	
Alelo sensible de <i>M. acuminata</i> (SEQ ID NO: 14)	GGAAATTAGCGATGATGGTGTGA	
Acc. de <i>M. balbisiana</i> ITC1016 (SEQ ID NO: 27)		
Alelo 1 de <i>M. basjoo</i> (SEQ ID NO: 18)	GGAAATTAGCGATGATGGTGTGA	423

M. itinerans 1 (SEQ ID NO: 3)
M. itinerans 2 (SEQ ID NO: 6)
M. acuminata resistente (SEQ ID NO: 12)
M. basjoo (SEQ ID NO: 19)
M. acuminata sensible (SEQ ID NO: 15)
M. balbisiana (SEQ ID NO: 32)
Musella lasiocarpa (SEQ ID NO: 25)

M. itinerans 1 (SEQ ID NO: 3)
M. itinerans 2 (SEQ ID NO: 6)
M. acuminata resistente (SEQ ID NO: 12)
M. basjoo (SEQ ID NO: 19)
M. acuminata sensible (SEQ ID NO: 15)
M. balbisiana (SEQ ID NO: 32)
Musella lasiocarpa (SEQ ID NO: 25)

M. itinerans 1 (SEQ ID NO: 3)
M. itinerans 2 (SEQ ID NO: 6)
M. acuminata resistente (SEQ ID NO: 12)
M. basjoo (SEQ ID NO: 19)
M. acuminata sensible (SEQ ID NO: 15)
M. balbisiana (SEQ ID NO: 32)
Musella lasiocarpa (SEQ ID NO: 25)

M. itinerans 1 (SEQ ID NO: 3)
M. itinerans 2 (SEQ ID NO: 6)
M. acuminata resistente (SEQ ID NO: 12)
M. basjoo (SEQ ID NO: 19)
M. acuminata sensible (SEQ ID NO: 15)
M. balbisiana (SEQ ID NO: 32)
Musella lasiocarpa (SEQ ID NO: 25)

M. itinerans 1 (SEQ ID NO: 3)
M. itinerans 2 (SEQ ID NO: 6)
M. acuminata resistente (SEQ ID NO: 12)
M. basjoo (SEQ ID NO: 19)
M. acuminata sensible (SEQ ID NO: 15)
M. balbisiana (SEQ ID NO: 32)
Musella lasiocarpa (SEQ ID NO: 25)

maggkrg^aeassllvllvmllaffatdssarvtprph
maggkrg^aeassllvllvmllaffatdssarvtprph
maggkrg^aeassllvllvmllaffatdssarvtprpq
maggkrg^aeassllvllvmllaffatnssarvtprpq
maggkrg^aeassllvllvmllaffatnssarvtprpq
maggkrg^aeassllvllvmllaffatnssarvtprpq
maggkrg^aeassllvllvmllaffatdssarvarprph
maggkrg^aeassllvllvmllaffatdssarvtprpq

40

*
SLARAVLSALEARADGPCRCICPLIYPTWCVCSGVWQG
SLARAVLSALEGRADGPCRCICPLIYPTWCVCSGVWQG
SLARAALSAGGARQDEPCRCACPLIYPTWCI CGGIWQG
SLARAALSAGGARQDEPCRCACPLIYPTWCI CGGIWQG
SLARAALSAGGARQDEPCRCACPLIYPTWCI CGGIWQG
SLARAALSAGVRQDAPCTCVCPLIYPPFCFGGVWQG
SLARVALSALGVRQDEPCRCICPRIYTAWCISGAWQG

80

*
SCPSACTNCECLINECTCLDHVDYKCAQDSCGWLDG-VP
SCPSACTNCECLINECTCLDHVDYKCAEADSGWLGD-VP
SCPSACNNQCIVLINECTCLDMDPKVCANSFPWA-A-P
SCPSACNNQCIVLINECTCLMDPKVCANSPWPVA-A-P
SCPSACNNQCIVLINECTCLDMDPKVCANSFPWA-A-P
SCPSACTNCEVLINECTCIDRVDPKVCESGASMO--PP
SCPSACTCKCOLNECHCEDDIYDYNACLADSPWLDAAP

120

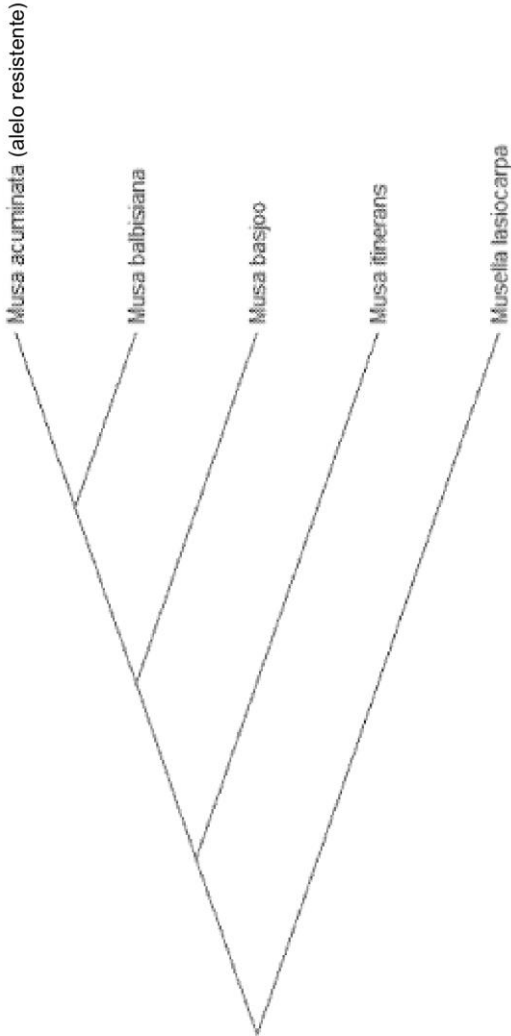
*

KLPFSQQWAIETGGKLAMV
KLRPSQQWAIETGGKLAMV
KVBPQQWAIETGGKLAMV
KVBPQQWAIETGGKLAMV
KVBPQQWAIETGGKLAMV

K
K
K
K
K

141

FIGURA 3



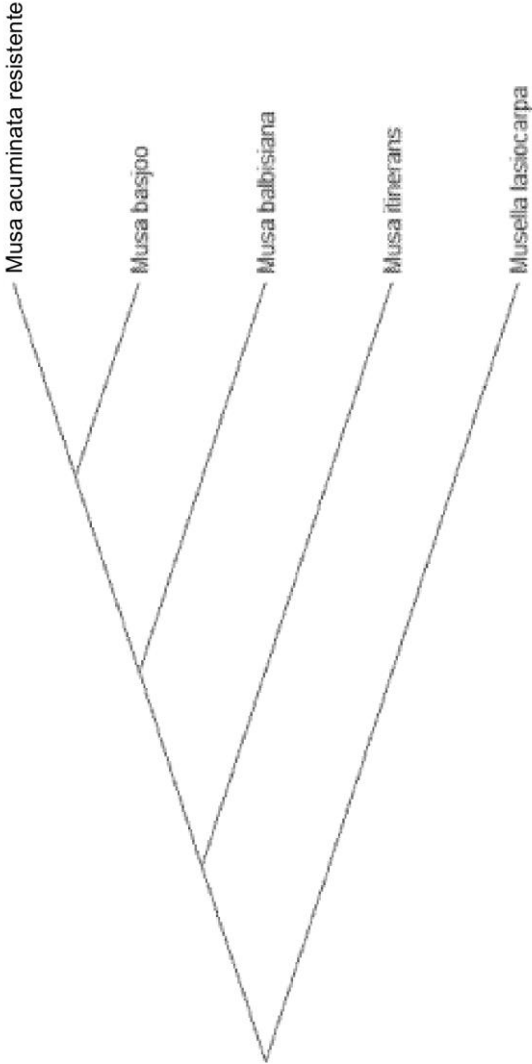


FIGURA 4

FIGURA 5

Acc. de <i>M. balbisiana</i> ITC1016 (SEQ ID NO: 26)	atgcttgagagagcctgaggggtgagcgtgctcttcttactttgtgac	50
Acc. de <i>M. balbisiana</i> ITC0545 (SEQ ID NO: 28)	atgcttgagagagcctgaggggtgagcgtgctcttcttactttgtgac	
Acc. de <i>M. balbisiana</i> ITC0080 (SEQ ID NO: 29)	atgcttgagagagcctgaggggtgagcgtgctcttcttactttgtgac	
Acc. de <i>M. balbisiana</i> ITC1527 (SEQ ID NO: 30)	atgcttgagagagcctgaggggtgagcgtgctcttcttactttgtgac	
Acc. de <i>M. balbisiana</i> ITC1016 (SEQ ID NO: 26)	gctgctgctgacgttgttggccttcttgcgcacccgactccttcggcagcccc	
Acc. de <i>M. balbisiana</i> ITC0545 (SEQ ID NO: 28)	gctgctgctgacgttgttggccttcttgcgcacccgactccttcggcagcccc	
Acc. de <i>M. balbisiana</i> ITC0080 (SEQ ID NO: 29)	gctgctgctgacgttgttggccttcttgcgcacccgactccttcggcagcccc	
Acc. de <i>M. balbisiana</i> ITC1527 (SEQ ID NO: 30)	gctgctgctgacgttgttggccttcttgcgcacccgactccttcggcagcccc	100
Acc. de <i>M. balbisiana</i> ITC1016 (SEQ ID NO: 26)	gtgtcgacaccccgctcgcactccttcgcgcagaggtaggtagataaatatg	
Acc. de <i>M. balbisiana</i> ITC0545 (SEQ ID NO: 28)	gtgtcgacaccccgctcgcactccttcgcgcagaggtaggtagataaatatg	
Acc. de <i>M. balbisiana</i> ITC0080 (SEQ ID NO: 29)	gtgtcgacaccccgctcgcactccttcgcgcagaggtaggtagataaatatg	
Acc. de <i>M. balbisiana</i> ITC1527 (SEQ ID NO: 30)	gtgtcgacaccccgctcgcactccttcgcgcagaggtaggtagataaatatg	150
Acc. de <i>M. balbisiana</i> ITC1016 (SEQ ID NO: 26)	<u>catgcgaacttggtatat - gattgggctgagatcgaggcatcggttaatt</u>	
Acc. de <i>M. balbisiana</i> ITC0545 (SEQ ID NO: 28)	<u>catgcgaacttggtatat - gattgggctgagatcgaggcatcggttaatt</u>	
Acc. de <i>M. balbisiana</i> ITC0080 (SEQ ID NO: 29)	<u>catgcgaacttggtatat - gattgggctgagatcgaggcatcggttaatt</u>	
Acc. de <i>M. balbisiana</i> ITC1527 (SEQ ID NO: 30)	<u>catgcgaacttggtatat - gattgggctgagatcgaggcatcggttaatt</u>	200

FIGURA 5 continuación

Acc. de <i>M. balbisiana</i> ITC1016 (SEQ ID NO: 26)	CGCTCTTCATGCTGACGGGCACTGAGTGCCTGTTGGGGTTAAGGCAAGATG	250
Acc. de <i>M. balbisiana</i> ITC0545 (SEQ ID NO: 28)	CCGTCCTTCATGCTGACGGGCACTGAGTGCCTGTTGGGGTTAAGC:::	
Acc. de <i>M. balbisiana</i> ITC0080 (SEQ ID NO: 29)	CGCTCTTCATGCTGACGGGCACTGAGTGCCTGTTGGGGTTAAGGCAAGATG	
Acc. de <i>M. balbisiana</i> ITC1527 (SEQ ID NO: 30)	CGCTCTTCATGCTGACGGGCACTGAGTGCCTGTTGGGGTTAAGGCAAGATG	250
Acc. de <i>M. balbisiana</i> ITC1016 (SEQ ID NO: 26)	CGCCGTCGTGCACATGCGTCTCTCTCTCTCATTTACCCACCTCCTTTTTC	
Acc. de <i>M. balbisiana</i> ITC0545 (SEQ ID NO: 28)	:::	
Acc. de <i>M. balbisiana</i> ITC0080 (SEQ ID NO: 29)	CGCCGTCGTGCACATGCGTCTCTCTCTCTCATTTACCCACCTCCTTTTTC	
Acc. de <i>M. balbisiana</i> ITC1527 (SEQ ID NO: 30)	CGCCGTCGTGCACATGCGTCTCTCTCTCTCATTTACCCACCTCCTTTTTC	300
Acc. de <i>M. balbisiana</i> ITC1016 (SEQ ID NO: 26)	TTTTGCGGCGGTATGGCAGGCTCTGCGCTCCGCTCCGCTGCACCAACTG	
Acc. de <i>M. balbisiana</i> ITC0545 (SEQ ID NO: 28)	:::	
Acc. de <i>M. balbisiana</i> ITC0080 (SEQ ID NO: 29)	TTTTGCGGCGGTATGGCAGGCTCTGCGCTCCGCTCCGCTGCACCAACTG	
Acc. de <i>M. balbisiana</i> ITC1527 (SEQ ID NO: 30)	TTTTGCGGCGGTATGGCAGGCTCTGCGCTCCGCTCCGCTGCACCAACTG	350
Acc. de <i>M. balbisiana</i> ITC1016 (SEQ ID NO: 26)	CGAGTGTGTCTCAACGAGTGCACCTTGCATCGATCGTGTGGACCCCAAGG	
Acc. de <i>M. balbisiana</i> ITC0545 (SEQ ID NO: 28)	CGAGTGTGTCTCAACGAGTGCACCTTGCATCGATCGTGTGGACCCCAAGG	
Acc. de <i>M. balbisiana</i> ITC0080 (SEQ ID NO: 29)	CGAGTGTGTCTCAACGAGTGCACCTTGCATCGATCGTGTGGACCCCAAGG	
Acc. de <i>M. balbisiana</i> ITC1527 (SEQ ID NO: 30)	CGAGTGTGTCTCAACGAGTGCACCTTGCATCGATCGTGTGGACCCCAAGG	400

FIGURA 5 continuación

Acc. de <i>M. balbisiana</i> ITC1016 (SEQ ID NO: 26)	CCTGCGAGGCCGACTCCTGTG: CCGGCT----	CGATGCAGCCCCCAAAGT	
Acc. de <i>M. balbisiana</i> ITC0545 (SEQ ID NO: 28)	CCTGCGAGGCCGACTCCTGTG: CCGGCTG	CGATGCAGCCCCCAAAGT	
Acc. de <i>M. balbisiana</i> ITC0080 (SEQ ID NO: 29)	CCTGCGTGGCCGACTCCTGTG: CCGGCT----	CGATGCAGCCCCCAAAGT	
Acc. de <i>M. balbisiana</i> ITC1527 (SEQ ID NO: 30)	CCTGCGAGGCCGACTCCTGTG: CCGGCTG	CGATGCAGCCCCCAAAGT	450
Acc. de <i>M. balbisiana</i> ITC1016 (SEQ ID NO: 26)	AGAGCCGTCGACAGTGGGCGACCGAAGAAACCGTGGGAAATTAGGGA		
Acc. de <i>M. balbisiana</i> ITC0545 (SEQ ID NO: 28)	AGAGCCGTCGACAGTGGGCGACCGAAGAAACCGTGGGAAATTAGGGA		
Acc. de <i>M. balbisiana</i> ITC0080 (SEQ ID NO: 29)	AGAGCCGTCGACAGTGGGCGACCGAAGAAACCGTGGGAAATTAGGGA		
Acc. de <i>M. balbisiana</i> ITC1527 (SEQ ID NO: 30)	AGAGCCGTCGACAGTGGGCGACCGAAGAAACCGTGGGAAATTAGGGA		500
Acc. de <i>M. balbisiana</i> ITC1016 (SEQ ID NO: 26)	CGATGGTGTgaTCCAATTGTGTTGTGA		
Acc. de <i>M. balbisiana</i> ITC0545 (SEQ ID NO: 28)	CGATGGTGTgaTCCA		
Acc. de <i>M. balbisiana</i> ITC0080 (SEQ ID NO: 29)	CGATGGTGTgaTCCAATTGTGTTGTGA		
Acc. de <i>M. balbisiana</i> ITC1527 (SEQ ID NO: 30)	CGATGGTGTgaTCCAATTGTGTTGT		528