

ELEMENTOS REGULADORES DE PLANTAS Y USOS DE ESTOS

REFERENCIA A SOLICITUD RELACIONADA

[1] La presente solicitud reivindica el beneficio de la solicitud provisional de patente estadounidense n.º 62/969,993, presentada el 4 de febrero de 2020, que se incorpora en su totalidad a la presente mediante referencia.

INCORPORACIÓN DEL LISTADO DE SECUENCIAS

[2] El listado de secuencias, contenido en el archivo con el nombre «MONS479WO_ST25.txt» es de 41,2 kilobytes (medido en Microsoft Windows®), se creó el 11 de enero de 2021, y se presenta junto con la presente por vía electrónica y se incorpora mediante referencia a la presente.

CAMPO DE LA INVENCION

[3] La invención hace referencia al campo de la biología molecular de plantas y la ingeniería genética de plantas. Más específicamente, la invención hace referencia a moléculas de ADN útiles para modular la expresión génica en las plantas.

ANTECEDENTES

[4] Los elementos reguladores son elementos genéticos que regulan la actividad génica al modular la transcripción de una molécula de ADN pasible de transcripción y enlazada de manera operativa. Esos elementos pueden incluir promotores, líderes, intrones y regiones 3' no traducidas y son útiles en los campos de la biología molecular y la ingeniería genética de plantas.

COMPENDIO DE LA INVENCION

[5] La invención proporciona elementos reguladores génicos para el uso en plantas. La invención también proporciona moléculas de ADN recombinantes que comprenden los elementos reguladores. La presente invención también proporciona células de plantas transgénicas, plantas y semillas que comprenden los elementos reguladores. En una modalidad, los elementos reguladores están enlazados de manera operativa a una molécula de ADN pasible de transcripción. En determinadas modalidades, la molécula de ADN pasible de transcripción puede ser heteróloga con respecto a la secuencia reguladora. Por lo tanto, una secuencia de elementos reguladores proporcionada por la invención puede definirse, en modalidades particulares, como enlazada de manera operativa a una molécula de ADN pasible de transcripción heteróloga. La presente invención también proporciona métodos de uso de los elementos reguladores, así como de elaboración y uso de las moléculas de ADN recombinantes que comprenden los elementos reguladores, y las células de

plantas, las plantas y las semillas transgénicas que comprenden los elementos reguladores enlazados de manera operativa a una molécula de ADN pasible de transcripción.

[6] Por lo tanto, en un aspecto, la invención proporciona una molécula de ADN recombinante que comprende una secuencia de ADN que se selecciona dentro del grupo que consiste en: (a) una secuencia con al menos alrededor de 85 por ciento de identidad de secuencia con respecto a cualquiera de las SEQ ID NOs:1-20; (b) una secuencia que comprende cualquiera de las SEQ ID NOs:1-20; y (c) un fragmento de cualquiera de las SEQ ID NOs:1-20, donde el fragmento tiene actividad de regulación de genes; donde la secuencia de ADN está enlazada de manera operativa a una molécula de ADN pasible de transcripción heteróloga. «Molécula de ADN pasible de transcripción heteróloga» significa que la molécula de ADN pasible de transcripción es heteróloga con respecto a la secuencia de polinucleótidos a la que está enlazada de manera operativa. En modalidades específicas, la molécula de ADN recombinante comprende una secuencia de ADN que tiene al menos alrededor de 85 por ciento, al menos alrededor de 86 por ciento, al menos alrededor de 87 por ciento, al menos alrededor de 88 por ciento, al menos alrededor de 89 por ciento, al menos alrededor de 90 por ciento, al menos 91 por ciento, al menos 92 por ciento, al menos 93 por ciento, al menos 94 por ciento, al menos 95 por ciento, al menos 96 por ciento, al menos 97 por ciento, al menos 98 por ciento o al menos 99 por ciento de identidad de secuencia con la secuencia de ADN de cualquiera de las SEQ ID NOs:1-20.

[7] En otro aspecto, en la presente se proporcionan células de plantas transgénicas que comprenden una molécula de ADN recombinante que comprende una secuencia de ADN que se selecciona dentro del grupo que consiste en: (a) una secuencia con al menos alrededor de 85 por ciento de identidad de secuencia con respecto a cualquiera de las SEQ ID NOs:1-20; (b) una secuencia que comprende cualquiera de las SEQ ID NOs:1-20; y (c) un fragmento de cualquiera de las SEQ ID NOs:1-20, donde el fragmento tiene actividad de regulación de genes; donde la secuencia está enlazada de manera operativa a una molécula de ADN pasible de transcripción heteróloga. En determinadas modalidades, la célula de planta transgénica es una célula de planta monocotiledónea. En otras modalidades, la célula de planta transgénica es una célula de planta dicotiledónea.

[8] En incluso otro aspecto, se proporciona además en la presente una planta transgénica, o parte de esta, que comprende una molécula de ADN que comprende

una secuencia de ADN que se selecciona dentro del grupo que consiste en: a) una secuencia con al menos 85 por ciento de identidad de secuencia con respecto a cualquiera de las SEQ ID NO:1-20; b) una secuencia que comprende cualquiera de las SEQ ID NO:1-20; y c) un fragmento de cualquiera de las SEQ ID NO:1-20, donde el fragmento tiene actividad de regulación de genes; donde la secuencia está enlazada de manera operativa a una molécula de ADN pasible de transcripción heteróloga. En modalidades específicas, la planta transgénica es una planta de progenie de cualquier generación que comprende la molécula de ADN recombinante. También se proporciona en la presente una semilla transgénica que comprende la molécula de ADN recombinante que produce esa planta transgénica cuando se cultiva.

[9] En otro aspecto, la invención proporciona un método para producir un producto básico que comprende obtener una planta transgénica o parte de esta que contenga una molécula de ADN recombinante de la invención y producir el producto básico a partir de esta. En una modalidad, el producto básico consiste en semillas, semillas procesadas, concentrado de proteínas, aislado de proteínas, almidón, granos, partes de plantas, aceite de semillas, biomasa, harina y sémola.

[10] En incluso otro aspecto adicional, la invención proporciona un método para producir una planta transgénica que comprende una molécula de ADN recombinante de la invención que comprende transformar una célula de planta con la molécula de ADN recombinante de la invención para producir una célula de planta transformada y regenerar una planta transgénica desde la célula de planta transformada.

BREVE DESCRIPCIÓN DE LAS SECUENCIAS

[11] La SEQ ID NO:1 es una secuencia de ADN de un promotor enlazado de manera operativa a un líder, P-Zm.GRMZM2G487322:2, que deriva de *Zea mays*.

[12] La SEQ ID NO:2 es una secuencia de ADN de una 3' UTR, T-Zm.GRMZM2G487322:2, que deriva de *Zea mays*.

[13] La SEQ ID NO:3 es una secuencia de ADN de un promotor enlazado de manera operativa a un líder, P-Zm.GRMZM2G339781:1, que deriva de *Zea mays*.

[14] La SEQ ID NO:4 es una secuencia de ADN de una 3' UTR, T-Zm.GRMZM2G339781:1, que deriva de *Zea mays*.

[15] La SEQ ID NO:5 es una secuencia de ADN de un grupo de elementos de expresión reguladores o EXP, EXP-Zm.Xet:1, donde se incluye un promotor enlazado de manera operativa al líder (P-Zm.Xet:1), enlazado de manera operativa a un intrón (I-Zm.Xet:1), que deriva de *Zea mays*.

[16] La SEQ ID NO:6 es una secuencia de ADN de un promotor enlazado de manera operativa a un líder, P-Zm.Xet:1, que deriva de *Zea mays*.

[17] La SEQ ID NO:7 es una secuencia de ADN de un intrón, I-Zm.Xet:1, que deriva de *Zea mays*.

5 **[18]** La SEQ ID NO:8 es una secuencia de ADN de una 3' UTR, T-Zm.Xet:1, que deriva de *Zea mays*.

[19] La SEQ ID NO:9 es una secuencia de ADN de un EXP, EXP-Zm.Sat6:1, conformado por un promotor enlazado de manera operativa al líder (P-Zm.Sat6:1), enlazado de manera operativa a un intrón (I-Zm.Sat6:1), que deriva de *Zea mays*.

10 **[20]** La SEQ ID NO:10 es una secuencia de ADN de un promotor enlazado de manera operativa a un líder, P-Zm.Sat6:1, que deriva de *Zea mays*.

[21] La SEQ ID NO:11 es una secuencia de ADN de un intrón, I-Zm.Sat6:1, que deriva de *Zea mays*.

15 **[22]** La SEQ ID NO:12 es una secuencia de ADN de una 3' UTR, T-Zm.Sat6:1, que deriva de *Zea mays*.

[23] La SEQ ID NO:13 es una secuencia de ADN de un EXP, EXP-Zm.GRMZM2G049726:1, conformado por un promotor enlazado de manera operativa al líder (P-Zm.GRMZM2G049726:1), enlazado de manera operativa a un intrón (I-Zm.GRMZM2G049726:1), que deriva de *Zea mays*.

20 **[24]** La SEQ ID NO:14 es una secuencia de ADN de un promotor enlazado de manera operativa a un líder, P-Zm.GRMZM2G049726:1, que deriva de *Zea mays*.

[25] La SEQ ID NO:15 es una secuencia de ADN de un intrón, I-Zm.GRMZM2G049726:1, que deriva de *Zea mays*.

25 **[26]** La SEQ ID NO:16 es una secuencia de ADN de una 3' UTR, T-Zm.GRMZM2G049726:1, que deriva de *Zea mays*.

[27] La SEQ ID NO:17 es una secuencia de ADN de un promotor enlazado de manera operativa a un líder, P-Zm.GRMZM2G141762:1, que deriva de *Zea mays*.

[28] La SEQ ID NO:18 es una secuencia de ADN de un promotor enlazado de manera operativa a un líder, P-Zm.DSUL:1, que deriva de *Zea mays*.

30 **[29]** La SEQ ID NO:19 es una secuencia de ADN de un promotor enlazado de manera operativa a un líder, P-Zm.GRMZM2G512113:1, que deriva de *Zea mays*.

[30] La SEQ ID NO:20 es una secuencia de ADN de una 3' UTR, T-Zm.GRMZM2G512113:1, derivada de *Zea mays*.

[31] La SEQ ID NO:21 es una secuencia de codificación sintética optimizada para la expresión en plantas para la β -glucuronidasa (GUS, GOI-Ec.uidA+St.LS1.nno:1) con un intrón procesable derivado del gen St-LS1 con especificidad de tejido e inducible por luz de la patata (N.º de registro en Genbank: X04753).

5

DESCRIPCIÓN DETALLADA DE LA INVENCIÓN

[32] La invención proporciona elementos reguladores que tienen actividad de regulación de genes en las plantas. Las secuencias de nucleótidos de estos elementos reguladores se proporcionan en las SEQ ID NOs:1-20. Estos elementos reguladores son capaces de afectar la expresión de una molécula de ADN pasible de transcripción enlazada de manera operativa en tejidos de plantas y, por lo tanto, de regular la expresión génica de un transgén enlazado de manera operativa en plantas transgénicas. La invención también proporciona métodos para modificar, producir y usar moléculas de ADN recombinante que contengan los elementos reguladores proporcionados. La invención también proporciona composiciones que incluyen células de plantas transgénicas, plantas, partes de plantas y semillas que contienen las moléculas de ADN recombinante de la invención, y métodos para prepararlas y utilizarlas.

[33] Se proporcionan las definiciones y los métodos que siguen para definir mejor la presente invención y guiar a los expertos en la técnica en la práctica de la presente invención. A menos que se indique lo contrario, los términos deben entenderse según el uso convencional que le dan los expertos en la técnica pertinente.

Moléculas de ADN

[34] Tal como se usan en la presente, las expresiones «ADN» o «molécula de ADN» hacen referencia a una molécula de ADN bicatenaria de origen genómico o sintético, es decir, un polímero de bases de desoxirribonucleótidos o una molécula de ADN, que se lee desde el extremo 5' (aguas arriba) hacia el extremo 3' (aguas abajo). Tal como se usa en la presente, la expresión «secuencia de ADN» hace referencia a la secuencia de nucleótidos de una molécula de ADN. La nomenclatura que se usa en la presente corresponde a la del Título 37 del Código de Reglamentos Federales de los Estados Unidos § 1.822, que se indica en las tablas de la norma de la OMPI ST.25 (1998), Apéndice 2, Tablas 1 y 3.

[35] Tal como se usa en la presente, una «molécula de ADN recombinante» es una molécula de ADN que comprende una combinación de moléculas de ADN que no se producirían en esa combinación de manera natural sin la intervención humana. Por

ejemplo, una molécula de ADN recombinante puede ser una molécula de ADN conformada por al menos dos moléculas de ADN heterólogas entre sí, una molécula de ADN que comprenda una secuencia de ADN que se desvíe con respecto a las secuencias de ADN que existen en la naturaleza, una molécula de ADN que comprenda una secuencia de ADN sintética o una molécula de ADN que se haya incorporado al ADN de una célula hospedadora mediante transformación genética o edición de genes.

[36] Se pretende que la referencia en la presente solicitud a una «molécula de ADN aislada», o un término o una expresión equivalente, signifique que la molécula de ADN está presente sola o en combinación con otras composiciones, pero no en su entorno natural. Por ejemplo, los elementos de ácido nucleico, tales como una secuencia codificante, una secuencia intrónica, una secuencia líder no traducida, una secuencia promotora, una secuencia de terminación de la transcripción, y similares, que se encuentran naturalmente en el ADN del genoma de un organismo no se consideran «aislados», siempre y cuando el elemento esté en el genoma del organismo y en la ubicación dentro del genoma en la que se encuentra naturalmente. Sin embargo, cada uno de estos elementos, y subpartes de estos elementos, podrían estar «aislados» dentro del alcance de la descripción, siempre y cuando el elemento no esté en el genoma del organismo ni en la ubicación dentro del genoma en la que se encuentra naturalmente. De manera similar, una secuencia de nucleótidos que codifica una proteína insecticida o cualquier variante insecticida de origen natural de esa proteína sería una secuencia de nucleótidos aislada, siempre y cuando la secuencia de nucleótidos no esté dentro del ADN de la bacteria donde la secuencia que codifica la proteína se encuentra naturalmente. Una secuencia sintética de nucleótidos que codifica la secuencia de aminoácidos de la proteína insecticida de origen natural podría considerarse aislada a los efectos de la presente descripción. A los efectos de la presente descripción, cualquier secuencia de nucleótidos transgénica, es decir, la secuencia de nucleótidos del ADN insertado en el genoma de las células de una planta o bacteria, o presente en un vector extracromosómico, se consideraría una secuencia de nucleótidos aislada ya sea que esté presente dentro del plásmido o estructura similar usada para transformar la célula, dentro del genoma de la planta o bacteria, o presente en cantidades detectables en tejidos, progenie, muestras biológicas o productos básicos derivados de la planta o bacteria.

[37] Tal como se usa en la presente, la expresión «identidad de secuencia» hace referencia al grado en que dos secuencias de polinucleótidos con alineación óptima o dos secuencias de polipéptidos con alineación óptima son idénticas. Se crea una alineación óptima de secuencias al alinear manualmente dos secuencias, p. ej., una secuencia de referencia y otra secuencia, para maximizar la cantidad de coincidencias de nucleótidos en la alineación de secuencias con inserciones, eliminaciones o huecos de nucleótidos internos adecuados. Tal como se usa en la presente, la expresión «secuencia de referencia» hace referencia a una secuencia de ADN proporcionada como las SEQ ID NOs:1-20.

[38] Tal como se usa en la presente, la expresión «porcentaje de identidad de secuencia» o «porcentaje de identidad» o «% de identidad» es la fracción de identidad multiplicada por 100. La «fracción de identidad» para una secuencia con alineación óptima con una secuencia de referencia es la cantidad de coincidencias de nucleótidos en la alineación óptima, dividida por la cantidad total de nucleótidos en la secuencia de referencia, p. ej., la cantidad total de nucleótidos en toda la longitud de la secuencia de referencia completa. Por lo tanto, una modalidad de la invención proporciona una molécula de ADN que comprende una secuencia que, cuando tiene una alineación óptima con respecto a una secuencia de referencia, proporcionada en la presente como las SEQ ID NOs:1-20, tiene al menos alrededor de 85 por ciento de identidad, al menos alrededor de 86 por ciento de identidad, al menos alrededor de 87 por ciento de identidad, al menos alrededor de 88 por ciento de identidad, al menos alrededor de 89 por ciento de identidad, al menos alrededor de 90 por ciento de identidad, al menos alrededor de 91 por ciento de identidad, al menos alrededor de 92 por ciento de identidad, al menos alrededor de 93 por ciento de identidad, al menos alrededor de 94 por ciento de identidad, al menos alrededor de 95 por ciento de identidad, al menos alrededor de 96 por ciento de identidad, al menos alrededor de 97 por ciento de identidad, al menos alrededor de 98 por ciento de identidad, al menos alrededor de 99 por ciento de identidad, o al menos alrededor de 100 por ciento de identidad con respecto a la secuencia de referencia.

Elementos reguladores

[39] Los elementos reguladores tales como promotores, líderes (también conocidos como 5' UTR), potenciadores, intrones y regiones de terminación de la transcripción (o 3' UTR) desempeñan un papel esencial en la expresión general de genes en células vivas. La expresión «elemento regulador», tal como se usa en la presente, hace

referencia a una molécula de ADN que tiene actividad de regulación de genes. La expresión «actividad de regulación de genes», tal como se usa en la presente, hace referencia a la capacidad de afectar la expresión de una molécula de ADN pasible de transcripción y enlazada de manera operativa, por ejemplo, al afectar la transcripción y/o traducción de la molécula de ADN pasible de transcripción y enlazada de manera operativa. Los elementos reguladores, tales como promotores, líderes, potenciadores, intrones y 3' UTR, que funcionan en las plantas son útiles para modificar los fenotipos de plantas mediante ingeniería genética.

[40] Tal como se usa en la presente, un «grupo de elementos de expresión reguladores» o una secuencia «EXP» puede hacer referencia a un grupo de elementos reguladores enlazados de manera operativa, tales como potenciadores, promotores, líderes e intrones. Por ejemplo, un grupo de elementos de expresión reguladores puede estar conformado, por ejemplo, por un promotor enlazado de manera operativa en 5' a una secuencia líder, enlazada de manera operativa en 5' a una secuencia intrónica. Las EXP útiles para poner en práctica la presente invención incluyen las SEQ ID NOS: 5, 9 y 13.

[41] Los elementos reguladores pueden caracterizarse por su patrón de expresión génica, p. ej., efectos positivos y/o negativos tales como la expresión constitutiva o temporal, espacial, de desarrollo, tisular, ambiental, fisiológica, patológica, del ciclo celular y/o con sensibilidad química, y cualquier combinación de estas, así como por indicaciones cuantitativas o cualitativas. Tal como se usa en la presente, un «patrón de expresión génica» es cualquier patrón de transcripción de una molécula de ADN enlazada de manera operativa para convertirla en una molécula de ARN transcrita. La molécula de ARN transcrita puede traducirse para producir una molécula de proteína o puede proporcionar una molécula de ARN no codificante u otra molécula de ARN reguladora, tal como un ARN bicatenario (dsRNA, por sus siglas en inglés), un ARN de transferencia (tRNA, por sus siglas en inglés), un ARN ribosómico (rRNA, por sus siglas en inglés), un microARN (miRNA, por sus siglas en inglés), un ARN interferente pequeño (siRNA, por sus siglas en inglés), y similares.

[42] Tal como se usa en la presente, la expresión «expresión de proteínas» es cualquier patrón de traducción de una molécula de ARN transcrita en una molécula de proteína. La expresión de proteínas puede caracterizarse por sus cualidades temporales, espaciales, de desarrollo o morfológicas, así como por indicaciones cuantitativas o cualitativas.

[43] Un promotor es útil como un elemento regulador para la modulación de la expresión de una molécula de ADN pasible de transcripción y enlazada de manera operativa. Tal como se usa en la presente, el término «promotor» hace referencia, en términos generales, a una molécula de ADN involucrada en el reconocimiento y la unión de la ARN polimerasa II y otras proteínas, tales como factores de transcripción de acción en trans, para iniciar la transcripción. Un promotor puede aislarse inicialmente de la región 5' no traducida (5' UTR) de una copia genómica de un gen, o a los efectos de la presente descripción, los promotores proporcionados en la presente están conformados por un promotor enlazado de manera operativa en 5' al líder. De manera alternativa, los promotores pueden ser moléculas de ADN manipuladas o producidas de manera sintética. Los promotores también pueden ser quiméricos. Los promotores quiméricos se producen a través de la fusión de dos o más moléculas de ADN heterólogas. Los promotores útiles para la puesta en práctica de la presente invención incluyen elementos promotores comprendidos dentro de cualquiera de las SEQ ID NOs:1, 3, 5, 6, 9, 10, 13, 14, 17, 18 y 19, o fragmentos o variantes de estas. En modalidades específicas de la invención, se define de manera adicional que las moléculas de ADN reivindicadas y cualquier variante o derivado de estas tal como se describe en la presente comprenden actividad promotora, es decir, son capaces de actuar como un promotor en una célula hospedadora, tal como en una planta transgénica. En incluso otras modalidades específicas adicionales, se puede definir que un fragmento presenta actividad promotora poseída por la molécula promotora inicial de la cual deriva, o un fragmento puede comprender un «promotor mínimo» que proporcione un nivel basal de transcripción y esté conformado por una caja TATA o secuencia de ADN equivalente para el reconocimiento y la unión del complejo de ARN polimerasa II para el inicio de la transcripción.

[44] En una modalidad, se proporcionan fragmentos de una secuencia promotora descrita en la presente. Los fragmentos de promotores pueden comprender actividad promotora, tal como se describió anteriormente, y pueden ser útiles solos o en combinación con otros promotores y fragmentos de promotores, tal como en promotores quiméricos de construcción, o en combinación con otros elementos de expresión y fragmentos de elementos de expresión. En modalidades específicas, se proporcionan fragmentos de un promotor que comprenden al menos alrededor de 50, al menos alrededor de 75, al menos alrededor de 95, al menos alrededor de 100, al menos alrededor de 125, al menos alrededor de 150, al menos alrededor de 175, al

menos alrededor de 200, al menos alrededor de 225, al menos alrededor de 250, al menos alrededor de 275, al menos alrededor de 300, al menos alrededor de 500, al menos alrededor de 600, al menos alrededor de 700, al menos alrededor de 750, al menos alrededor de 800, al menos alrededor de 900, o al menos alrededor de 1000 nucleótidos contiguos, o más, de una molécula de ADN que tiene actividad promotora tal como se describe en la presente. Los métodos para producir esos fragmentos a partir de una molécula promotora inicial son muy conocidos en la técnica.

[45] Las composiciones derivadas de cualquiera de los elementos promotores comprendidos dentro de cualquiera de las SEQ ID NOs:1, 3, 5, 6, 9, 10, 13, 14, 17, 18 y 19 tales como eliminaciones 5' o internas, por ejemplo, pueden producirse mediante el uso de métodos conocidos en la técnica para mejorar o alterar la expresión, incluso mediante la remoción de elementos que tengan efectos positivos o negativos en la expresión; la duplicación de elementos que tengan efectos positivos o negativos en la expresión; y/o la duplicación o remoción de elementos que tengan efectos específicos a nivel tisular o celular en la expresión. Las composiciones derivadas de cualquiera de los elementos promotores comprendidos dentro de cualquiera de las SEQ ID NOs:1, 3, 5, 6, 9, 10, 13, 14, 17, 18 y 19, que comprenden eliminaciones 3' en las que se remueve el elemento de caja TATA o su secuencia equivalente, y la secuencia aguas abajo, pueden usarse, por ejemplo, para generar elementos potenciadores. Se pueden realizar eliminaciones adicionales para remover cualquier elemento que tenga efectos positivos o negativos; específicos a nivel tisular; específicos a nivel celular; o específicos a nivel de tiempo (tal como, de modo no taxativo, el ritmo circadiano) en la expresión. Cualquiera de los elementos promotores comprendidos dentro de cualquiera de las SEQ ID NOs:1, 3, 5, 6, 9, 10, 13, 14, 17, 18 y 19 y sus fragmentos o potenciadores derivados pueden usarse para generar composiciones de elementos reguladores de la transcripción quiméricos.

[46] De acuerdo con la invención, un promotor o fragmento de promotor puede analizarse en busca de la presencia de elementos promotores conocidos, es decir, las características de secuencias de ADN, tales como una caja TATA y otros motivos de sitios de unión de factores de transcripción conocidos. La identificación de esos elementos promotores conocidos puede ser usada por el experto en la técnica para diseñar variantes del promotor que tengan un patrón de expresión similar al del promotor original.

[47] Tal como se usa en la presente, el término «líder» hace referencia a una molécula de ADN aislada de la región 5' no traducida (5' UTR), un gen y definida generalmente como un segmento de nucleótidos entre el sitio de inicio de la transcripción (TSS, por sus siglas en inglés) y el sitio de inicio de la secuencia de codificación de proteínas. De manera alternativa, los líderes pueden ser elementos de ADN manipulados o producidos de manera sintética. Se puede usar un líder como un elemento regulador 5' para la modulación de la expresión de una molécula de ADN pasible de transcripción y enlazada de manera operativa. Las moléculas líderes pueden usarse con un promotor heterólogo o con su promotor natural. Los líderes útiles para la puesta en práctica de la presente invención incluyen elementos líderes comprendidos dentro de cualquiera de las SEQ ID NOS:1, 3, 5, 6, 9, 10, 13, 14, 17, 18 y 19, o fragmentos o variantes de estas. En modalidades específicas, esas secuencias de ADN pueden definirse como capaces de actuar como un líder en una célula hospedadora, lo cual incluye, por ejemplo, una célula de planta transgénica. En una modalidad, esas secuencias se decodifican como que comprenden actividad de líder.

[48] Las secuencias líderes (también denominadas 5' UTR) comprendidas dentro de cualquiera de las SEQ ID NOS:1, 3, 5, 6, 9, 10, 13, 14, 17, 18 y 19 pueden estar conformadas por elementos reguladores, o pueden adoptar estructuras secundarias que pueden tener un efecto en la transcripción o traducción de una molécula de ADN pasible de transcripción y enlazada de manera operativa. Las secuencias líderes comprendidas dentro de cualquiera de las SEQ ID NOS:1, 3, 5, 6, 9, 10, 13, 14, 17, 18 y 19 pueden usarse de acuerdo con la invención para generar elementos reguladores quiméricos que afecten la transcripción o traducción de una molécula de ADN pasible de transcripción y enlazada de manera operativa.

[49] Tal como se usa en la presente, el término «intrón» hace referencia a una molécula de ADN que puede aislarse o identificarse a partir de un gen y puede definirse generalmente como una región despalmada durante el procesamiento del ARN mensajero (mRNA, por sus siglas en inglés) antes de la traducción. De manera alternativa, un intrón puede ser un elemento de ADN manipulado o producido de manera sintética. Un intrón puede contener elementos potenciadores que efectúen la transcripción de genes enlazados de manera operativa. Se puede usar un intrón como elemento regulador para modular la expresión de una molécula de ADN pasible de transcripción y enlazada de manera operativa. Una construcción puede comprender un intrón, y el intrón puede o no ser heterólogo con respecto a la molécula de ADN

pasible de transcripción. Ejemplos de intrones en la técnica incluyen el intrón de actina de arroz y el intrón HSP70 de maíz.

[50] En las plantas, la inclusión de algunos intrones en las construcciones génicas conduce al aumento del mRNA y la acumulación de proteína con respecto a construcciones que carecen del intrón. Este efecto ha sido denominado «potenciamiento mediado por intrones» (IME, por sus siglas en inglés) de la expresión génica. Se han identificado intrones conocidos por estimular la expresión en las plantas en genes del maíz (p. ej., tubA1, Adh1, Sh1 y Ubi1), en genes del arroz (p. ej., tpi) y en genes de plantas dicotiledóneas como los de la petunia (p. ej., rbcS), la papa (p. ej., st-ls1), y de la *Arabidopsis thaliana* (p. ej., ubq3 y pat1). Se ha demostrado que las eliminaciones o mutaciones dentro de los sitios de corte y empalme de un intrón reducen la expresión génica, lo que indica que el corte y empalme puede ser necesario para el IME. Sin embargo, mutaciones puntuales dentro de los sitios de corte y empalme del gen pat1 de la *A. thaliana* han mostrado IME en plantas dicotiledóneas. Se ha demostrado que múltiples usos del mismo intrón en una planta presentan desventajas. En esos casos, es necesario tener un conjunto de elementos de control básico para la construcción de elementos de ADN recombinante adecuados. Los ejemplos de intrones útiles en la puesta en práctica de la presente invención se presentan como las SEQ ID NOs:7, 11 y 15.

[51] Tal como se usan en la presente, las expresiones «molécula de terminación de la transcripción 3'», «región 3' no traducida» o «3' UTR» hacen referencia a una molécula de ADN que se usa durante la transcripción para obtener la región no traducida de la parte 3' de una molécula de mRNA. La región 3' no traducida de una molécula de mRNA puede generarse mediante la escisión específica y poliadenilación 3', lo que también se conoce como cola de poliA. Una 3' UTR puede estar enlazada de manera operativa y ubicada aguas abajo con respecto a una molécula de ADN pasible de transcripción y puede incluir una señal de poliadenilación y otras señales reguladoras capaces de afectar la transcripción, el procesamiento de mRNA o la expresión génica. Se cree que las colas de poliA funcionan en estabilidad de mRNA y en el inicio de la traducción. Los ejemplos de moléculas de terminación de la transcripción en 3' en la técnica son la región 3' de la nopalina sintasa, la región 3' de hsp17 del trigo, la región 3' de la subunidad pequeña de rubisco del guisante, la región 3' de E6 del algodón y la 3' UTR de la coixina.

[52] Las 3' UTR normalmente tienen un uso beneficioso para la expresión recombinante de moléculas de ADN específicas. Una 3' UTR débil tiene el potencial de generar revisión, lo que puede afectar la expresión de la molécula de ADN ubicada en los casetes de expresión circundantes. Un control adecuado de la terminación de la transcripción puede evitar la revisión en secuencias de ADN (p. ej., otros casetes de expresión) que se ubican aguas abajo y puede permitir además un reciclado eficiente de la ARN polimerasa para mejorar la expresión génica. La terminación eficiente de la transcripción (la liberación de ARN Polimerasa II del ADN) es un requisito previo para el reinicio de la transcripción y de ese modo afecta directamente el nivel de transcripción general. Después de la terminación de la transcripción, el mRNA maduro se libera del sitio de síntesis y se transporta como molde al citoplasma. Los mRNA eucariotas se acumulan como formas de poli(A) *in vivo*, lo que hace difícil detectar los sitios de terminación de la transcripción mediante métodos convencionales. Sin embargo, la predicción de las 3' UTR funcionales y eficientes mediante métodos bioinformáticos es difícil dado que no hay secuencias de ADN conservadas que permitan la predicción fácil de una 3' UTR efectiva.

[53] Desde un punto de vista práctico, normalmente es beneficioso que una 3' UTR usada en un casete de expresión posea las características detalladas a continuación. En primer lugar, la 3' UTR debería ser capaz de terminar de manera eficiente y eficaz la transcripción del transgén y evitar la revisión de la transcripción en cualquier secuencia de ADN circundante, que puede comprender otro casete de expresión como en el caso de múltiples casetes de expresión que residen en un ADN de transferencia (T-DNA, por sus siglas en inglés), o el ADN cromosómico circundante en el que se ha insertado el T-DNA. En segundo lugar, la 3' UTR no debería causar una reducción en la actividad de transcripción a instancias del promotor, del líder, de los potenciadores y de los intrones que se usan para impulsar la expresión de la molécula de ADN. Finalmente, en la biotecnología vegetal, la 3' UTR se usa a menudo para el cebado de reacciones de amplificación del ARN transcrito inverso que se extrae de la planta transformada y se usa para: (1) evaluar la actividad de transcripción o la expresión del casete de expresión una vez integrado en el cromosoma de la planta; (2) evaluar la cantidad de copias de inserciones dentro del ADN de la planta; y (3) evaluar la cigosidad de la semilla resultante después de la reproducción. La 3' UTR también se usa en reacciones de amplificación de ADN extraído de la planta transformada para caracterizar la integridad del casete insertado. Las 3' UTR útiles

para poner en práctica la presente invención se presentan como las SEQ ID NOs: 2, 4, 8, 12, 16 y 20.

[54] Tal como se usa en la presente, el término «potenciador» o «elemento potenciador» hace referencia a un elemento regulador de acción en *cis*, también conocido como elemento *cis*, que confiere un aspecto del patrón de expresión general, pero habitualmente es insuficiente por sí solo para impulsar la transcripción, de una molécula de ADN pasible de transcripción y enlazada de manera operativa. A diferencia de los promotores, los elementos potenciadores no incluyen habitualmente un sitio de inicio de la transcripción (TSS, por sus siglas en inglés) ni una caja TATA o una secuencia de ADN equivalente. Un promotor o fragmento de promotor puede comprender naturalmente uno o más elementos potenciadores que afecten la transcripción de una secuencia de ADN enlazada de manera operativa. Un elemento potenciador también puede fusionarse con un promotor para producir un elemento *cis* promotor quimérico, que confiera un aspecto de la modulación general de la expresión génica.

[55] Se cree que muchos elementos promotores-potenciadores se unen a proteínas de unión al ADN y/o afectan la topología del ADN, mediante lo cual se producen conformaciones locales que permiten o restringen de manera selectiva el acceso de la ARN polimerasa al molde de ADN o que facilitan la apertura selectiva de la doble hélice en el sitio de inicio de la transcripción. Un elemento potenciador puede funcionar para unirse a factores de transcripción que regulen la transcripción. Algunos elementos potenciadores se unen a más de un factor de transcripción, y los factores de transcripción pueden interactuar con afinidades diferentes con más de un dominio potenciador. Los elementos potenciadores pueden identificarse mediante una cantidad de técnicas, que incluyen el análisis de eliminación, es decir, la eliminación de uno o más nucleótidos del extremo 5' o internos de un promotor, el análisis de proteínas de unión al ADN mediante el uso de seguimiento de huella de la DNasa I, interferencia de metilación, ensayos de cambio de movilidad electroforética, seguimiento de huella genómica *in vivo* mediante reacción en cadena de la polimerasa (PCR, por sus siglas en inglés) mediada por ligación y otros ensayos convencionales o mediante análisis de similitud de secuencias de ADN mediante el uso de motivos de elementos *cis* conocidos o elementos potenciadores como una secuencia objetivo o un motivo objetivo con métodos de comparación de secuencias de ADN convencionales, tales como BLAST. La estructura fina de un dominio potenciador

puede estudiarse de manera adicional mediante la mutagénesis (o sustitución) de uno o más nucleótidos o mediante otros métodos convencionales conocidos en la técnica. Los elementos potenciadores pueden obtenerse mediante síntesis química o al aislarse de los elementos reguladores que incluyen esos elementos, y pueden
5 sintetizarse con nucleótidos flanqueantes adicionales que contengan sitios enzimáticos de restricción útiles para facilitar la manipulación de subsecuencias. Por lo tanto, el diseño, la construcción y el uso de elementos potenciadores de acuerdo con los métodos descritos en la presente para modular la expresión de moléculas de ADN pasibles de transcripción y enlazadas de manera operativa están comprendidos
10 en la invención. Los potenciadores pueden derivar de cualquiera de los promotores comprendidos en las SEQ ID NOs:1, 3, 5, 6, 9, 10, 13, 14, 17, 18 y 19.

[56] Tal como se usa en la presente, el término «quimérico/a» hace referencia a una molécula de ADN individual producida mediante la fusión de una primera molécula de ADN con una segunda molécula de ADN, donde ni la primera ni la segunda molécula
15 de ADN se encontrarían normalmente en esa configuración, es decir, fusionadas entre sí. La molécula de ADN quimérica es, por lo tanto, una nueva molécula de ADN que de otro modo no se encontraría normalmente en la naturaleza. Tal como se usa en la presente, la expresión «promotor quimérico» hace referencia a un promotor producido a través de esa manipulación de las moléculas de ADN. Un promotor quimérico puede
20 combinar dos o más fragmentos de ADN; por ejemplo, la fusión de un promotor con un elemento potenciador. Por lo tanto, el diseño, la construcción y el uso de promotores quiméricos de acuerdo con los métodos descritos en la presente para modular la expresión de moléculas de ADN pasibles de transcripción y enlazadas de manera operativa están comprendidos en la presente invención.

[57] Los elementos reguladores quiméricos pueden diseñarse para que comprendan diversos elementos constituyentes que puedan enlazarse de manera operativa mediante diversos métodos conocidos en la técnica, tales como la digestión y ligación con enzimas de restricción, la clonación independiente de la ligación, el ensamblaje modular de productos de PCR durante la amplificación o la síntesis
30 química directa del elemento regulador, así como otros métodos conocidos en la técnica. Los diversos elementos reguladores quiméricos resultantes pueden estar conformados por los mismos elementos constituyentes, o variantes de estos, pero diferir en la secuencia de ADN o las secuencias de ADN que comprenden la secuencia o las secuencias de ADN enlazadoras que permiten que las partes constituyentes se

enlacen de manera operativa. En la invención, las secuencias de ADN proporcionadas como las SEQ ID NOs:1-20 pueden proporcionar secuencias de referencia de elementos reguladores, donde los elementos constituyentes que comprenden la secuencia de referencia pueden unirse mediante métodos conocidos en la técnica y pueden comprender sustituciones, eliminaciones y/o inserciones de uno o más nucleótidos o mutaciones que se produzcan de manera natural en la transformación de células de plantas y bacterianas.

[58] Tal como se usa en la presente, el término «variante» hace referencia a una segunda molécula de ADN, tal como un elemento regulador, que es similar en composición, aunque no idéntica, a una primera molécula de ADN, y donde la segunda molécula de ADN sigue manteniendo la funcionalidad general, es decir, el mismo patrón de expresión o uno similar, por ejemplo, a través de una actividad de transcripción más o menos equivalente, de la primera molécula de ADN. Una variante puede ser una versión más corta o truncada de la primera molécula de ADN o una versión alterada de la secuencia de la primera molécula de ADN, tal como una con diferentes sitios de enzimas de restricción y/o eliminaciones, sustituciones o inserciones internas. Una «variante» también puede abarcar un elemento regulador que tenga una secuencia de nucleótidos que comprenda una sustitución, eliminación o inserción de uno o más nucleótidos de una secuencia de referencia, donde el elemento regulador derivado tiene una actividad de transcripción o traducción mayor, menor o equivalente a la de la molécula reguladora original correspondiente. Las «variantes» de elementos reguladores también abarcarán variantes que surjan a partir de mutaciones que se produzcan de manera natural en la transformación de células de plantas y bacterianas. En la presente invención, una secuencia de polinucleótidos proporcionada como las SEQ ID NOs:1-20 puede usarse para crear variantes similares en composición, aunque no idénticas, a la secuencia de ADN del elemento regulador original, y al mismo tiempo mantener la funcionalidad general, es decir, el mismo patrón de expresión o uno similar, del elemento regulador original. La producción de esas variantes de la invención se encuentra dentro del conocimiento de la técnica a la luz de la descripción y está comprendida dentro del alcance de la invención.

[59] La eficacia de las modificaciones, duplicaciones o eliminaciones descritas en la presente en los aspectos de expresión deseados de un transgén particular puede evaluarse de manera empírica en ensayos en plantas estables y temporales, tales

como los que se describen en los ejemplos de trabajo de la presente, de manera de validar los resultados. que pueden variar dependiendo de los cambios hechos y la finalidad del cambio en la molécula de ADN inicial.

Construcciones

5 **[60]** Tal como se usa en la presente, el término «construcción» significa cualquier molécula de ADN recombinante tal como un plásmido, cósmido, virus, fago o molécula de ADN o ARN lineal o circular, que derive de cualquier fuente, capaz de una integración genómica o replicación autónoma, que comprenda una molécula de ADN donde al menos una molécula de ADN se haya enlazado a otra molécula de ADN de
10 manera funcionalmente operativa, es decir, enlazada de manera operativa. Tal como se usa en el presente, el término «vector» significa cualquier construcción que pueda usarse con el fin de una transformación, es decir, la introducción de ADN o ARN heterólogo en una célula hospedadora. Una construcción incluye normalmente uno o más casetes de expresión. Tal como se usa en la presente, un «casete de expresión»
15 hace referencia a una molécula de ADN que comprende al menos una molécula de ADN pasible de transcripción y enlazada de manera operativa a uno o más elementos reguladores, normalmente al menos un promotor y una 3' UTR.

[61] Tal como se usa en la presente, la expresión «enlazado/a de manera operativa» hace referencia a una primera molécula de ADN unida a una segunda
20 molécula de ADN, donde la primera y la segunda molécula de ADN están dispuestas de forma tal que la primera molécula de ADN afecte la función de la segunda molécula de ADN. Las dos moléculas de ADN pueden o no ser parte de una única molécula de ADN contigua y pueden o no ser adyacentes. Por ejemplo, un promotor está enlazado de manera operativa a una molécula de ADN pasible de transcripción si el promotor
25 modula la transcripción de la molécula de ADN pasible de transcripción de interés en una célula. Un líder, por ejemplo, está enlazado de manera operativa a la secuencia de ADN cuando es capaz de afectar la transcripción o traducción de la secuencia de ADN.

[62] Las construcciones de la invención pueden proporcionarse, en una modalidad,
30 como construcciones dobles de borde de plásmido de inducción tumoral (Ti, por sus siglas en inglés) que tienen las regiones del borde derecho (RB o AGRtu.RB) y borde izquierdo (LB o AGRtu.LB) del plásmido Ti que se aíslan de *Agrobacterium tumefaciens* que comprenden un T-DNA que, junto con las moléculas de transferencia proporcionadas por las células de *A. tumefaciens*, permiten la integración del T-DNA

en el genoma de una célula de planta (véase, p. ej., la patente estadounidense 6,603,061). Las construcciones también pueden contener los segmentos de ADN de estructura principal de plásmido que proporcionan la función de replicación y la selección de antibióticos en las células bacterianas, p. ej., un origen de replicación de *Escherichia coli* tal como ori322, un origen de replicación de rango de hospedador amplio tal como oriV u oriRi, y una región de codificación para un marcador seleccionable tal como Spec/Strp que codifica Tn7 aminoglicósido adeniltransferasa (*aadA*), lo que confiere resistencia a la espectinomicina o estreptomicina, o un gen marcador seleccionable de gentamicina (Gm, Gent). Para la transformación en plantas, la cepa bacteriana hospedadora generalmente es *A. tumefaciens* ABI, C58, o LBA4404; sin embargo, otras cepas conocidas por los expertos en la técnica de la transformación de plantas pueden funcionar en la invención.

[63] Se conocen métodos en la técnica para ensamblar e introducir construcciones en una célula de tal manera que la molécula de ADN pasible de transcripción se transcriba en una molécula de ARNm funcional traducida y expresada como una proteína. Para la puesta en práctica de la invención, los expertos en la técnica conocen bien las composiciones y los métodos convencionales para preparar y usar construcciones y células hospedadoras. Los vectores típicos útiles para la expresión de ácidos nucleicos en plantas superiores son bien conocidos en la técnica e incluyen vectores derivados del plásmido Ti de *Agrobacterium tumefaciens* y el vector de control de transferencia pCaMVN.

[64] Varios elementos reguladores pueden estar incluidos en una construcción, incluido cualquiera de los que se proporciona en la presente. Cualquiera de esos elementos reguladores puede proporcionarse en combinación con otros elementos reguladores. Esas combinaciones pueden diseñarse o modificarse para producir características reguladoras convenientes. En una modalidad, las construcciones de la invención comprenden al menos un elemento regulador enlazado de manera operativa a una molécula de ADN pasible de transcripción y enlazada de manera operativa a una 3' UTR.

[65] Las construcciones de la invención pueden incluir cualquier promotor o líder proporcionado en la presente o conocido en la técnica. Por ejemplo, un promotor de la invención puede estar enlazado de manera operativa a un líder 5' no traducido y heterólogo, tal como uno derivado de un gen de proteína de choque térmico. De manera alternativa, un líder de la invención puede estar enlazado de manera operativa

a un promotor heterólogo tal como el promotor de la transcripción del virus del mosaico de la coliflor 35S.

[66] Los casetes de expresión también pueden incluir una secuencia de codificación de péptidos de tránsito que codifique un péptido útil para el direccionamiento subcelular de una proteína enlazada de manera operativa, particularmente a un cloroplasto, leucoplasto u otro organelo plástido; una mitocondria; un peroxisoma; una vacuola; o una ubicación extracelular. Muchas proteínas localizadas en cloroplastos se expresan a partir de genes nucleares como precursores y se dirigen al cloroplasto mediante un péptido de tránsito hasta el cloroplasto (CTP, por sus siglas en inglés).

Los ejemplos de esas proteínas de cloroplastos aislados incluyen, de modo no taxativo, las asociadas con la subunidad pequeña (SSU, por sus siglas en inglés) de ribulosa-1,5,-bisfosfato carboxilasa, ferredoxina, ferredoxina oxidoreductasa, la proteína I y proteína II del complejo captador de luz, tiorredoxina F, y enolpiruvil shikimato fosfato sintasa (EPSPS, por sus siglas en inglés). Los péptidos de tránsito hasta el cloroplasto se describen, por ejemplo, en la patente estadounidense n.º 7,193,133. Se ha demostrado que las proteínas que no son cloroplastos pueden dirigirse al cloroplasto por la expresión de un CTP heterólogo enlazado de manera operativa al transgén que codifica una proteína que no es un cloroplasto.

Moléculas de ADN pasibles de transcripción

[67] Tal como se usa en la presente, la expresión «molécula de ADN pasible de transcripción» hace referencia a cualquier molécula de ADN pasible de transcripción en una molécula de ARN, lo cual incluye, de modo no taxativo, aquellas que tienen secuencias de codificación de proteínas, aquellas que codifican ARN guía y aquellas que producen moléculas de ARN que tienen secuencias útiles para la supresión génica. El tipo de molécula de ADN puede incluir, de modo no taxativo, una molécula de ADN de la misma planta, una molécula de ADN de otra planta, una molécula de ADN de un organismo diferente o una molécula de ADN sintético, tal como una molécula de ADN que contenga un mensaje no codificante de un gen, o una molécula de ADN que codifique una versión artificial, sintética o modificada de otro modo de un transgén. Los ejemplos de moléculas de ADN pasibles de transcripción para su incorporación en construcciones de la invención incluyen, p. ej., moléculas de ADN o genes de una especie distinta de la especie en la que se incorpore la molécula de ADN o genes que se originen o estén presentes en la misma especie, pero que se

incorporen en las células receptoras mediante métodos de ingeniería genética en lugar de técnicas de reproducción clásicas.

[68] Un «transgén» hace referencia a una molécula de ADN pasible de transcripción heteróloga con respecto a una célula hospedadora al menos con respecto a su ubicación en el genoma de la célula hospedadora y/o una molécula de ADN pasible de transcripción incorporada artificialmente en el genoma de una célula hospedadora en la generación actual o cualquier generación anterior de la célula.

[69] Un elemento regulador, tal como un promotor de la invención, puede estar enlazado de manera operativa a una molécula de ADN pasible de transcripción heteróloga con respecto al elemento regulador. Tal como se usa en la presente, el término «heterólogo» hace referencia a la combinación de dos o más moléculas de ADN cuando una combinación de ese tipo no se encuentra normalmente en la naturaleza. Por ejemplo, las dos moléculas de ADN pueden derivar de especies diferentes y/o las dos moléculas de ADN pueden derivar de genes diferentes, p. ej., genes diferentes de la misma especie o los mismos genes de especies diferentes. Por lo tanto, un elemento regulador es heterólogo con respecto a una molécula de ADN pasible de transcripción y enlazada de manera operativa si una combinación de ese tipo normalmente no se encuentra en la naturaleza, es decir, la molécula de ADN pasible de transcripción no se produce de manera natural en enlace operativo con el elemento regulador.

[70] La molécula de ADN pasible de transcripción puede generalmente ser cualquier molécula de ADN para la cual se desea la expresión de una transcripción. Esa expresión de una transcripción puede dar como resultado la traducción de la molécula de mRNA resultante, y, por lo tanto, la expresión de proteínas. De manera alternativa, por ejemplo, una molécula de ADN pasible de transcripción puede diseñarse para causar en última instancia la disminución de la expresión de una proteína o un gen específico. En una modalidad, esto se puede lograr mediante el uso de una molécula de ADN pasible de transcripción orientada en la dirección no codificante. El experto en la técnica está familiarizado con el uso de esa tecnología en la dirección no codificante. Cualquier gen puede someterse a una regulación negativa de esta manera, y, en una modalidad, una molécula de ADN pasible de transcripción puede diseñarse para la supresión de un gen específico a través de la expresión de una molécula de dsRNA, siRNA o miRNA.

[71] Por lo tanto, una modalidad de la invención es una molécula de ADN recombinante que comprende un elemento regulador de la invención, tal como se proporciona en las SEQ ID NOs:1-20, con un enlace operativo a una molécula de ADN pasible de transcripción heteróloga de manera de modular la transcripción de la molécula de ADN pasible de transcripción en un nivel deseado o en un patrón deseado cuando la construcción está integrada en el genoma de una célula de planta transgénica. En una modalidad, la molécula de ADN transcribible comprende una región de codificación de proteína de un gen y en otra modalidad la molécula de ADN pasible de transcripción comprende una región antisentido de un gen.

10 **Genes de interés agronómico**

[72] Una molécula de ADN pasible de transcripción puede ser un gen de interés agronómico. Tal como se usa en la presente, la expresión «gen de interés agronómico» hace referencia a una molécula de ADN pasible de transcripción que, cuando se expresa en un tejido, una célula o un tipo de célula de planta particular, confiere una característica conveniente. El producto de un gen de interés agronómico puede actuar dentro de la planta con el fin de causar un efecto sobre la morfología, la fisiología, el crecimiento, el desarrollo, el rendimiento, la composición de los granos, el perfil nutricional, la resistencia a enfermedades o plagas y/o la tolerancia al entorno o a químicos que tiene una planta, o puede actuar como un agente pesticida en la dieta de una plaga que se alimenta de la planta. En una modalidad de la invención, un elemento regulador de la invención se incorpora en una construcción de forma tal que el elemento regulador esté enlazado de manera operativa a una molécula de ADN pasible de transcripción que sea un gen de interés agronómico. En una planta transgénica que contiene una construcción de ese tipo, la expresión del gen de interés agronómico puede conferir un rasgo agronómico beneficioso. Un rasgo agronómico beneficioso puede incluir, por ejemplo, de modo no taxativo, la tolerancia a herbicidas, el control de insectos, el rendimiento modificado, la resistencia a enfermedades, la resistencia a patógenos, el crecimiento y desarrollo de plantas con modificaciones, el contenido de almidón modificado, el contenido de aceite modificado, el contenido de ácido graso modificado, el contenido de proteína con modificación, la maduración de la fruta modificada, la nutrición animal y humana aumentada, las producciones de biopolímeros, la resistencia a dificultades del entorno, los péptidos farmacéuticos, las cualidades de procesamiento mejoradas, el sabor mejorado, la utilidad de producción

de semillas híbridas, la producción de fibra mejorada y la producción de biocombustible conveniente.

[73] Los ejemplos de genes de interés agronómico conocidos en la técnica incluyen aquellos para la resistencia a herbicidas (patentes estadounidenses n.º 6,803,501; 5 6,448,476; 6,248,876; 6,225,114; 6,107,549; 5,866,775; 5,804,425; 5,633,435; y 5,463,175), un mayor rendimiento (patentes estadounidenses n.º USRE38,446; 6,716,474; 6,663,906; 6,476,295; 6,441,277; 6,423,828; 6,399,330; 6,372,211; 6,235,971; 6,222,098; y 5,716,837), el control de insectos (patentes estadounidenses n.º 6,809,078; 6,713,063; 6,686,452; 6,657,046; 6,645,497; 6,642,030; 6,639,054; 10 6,620,988; 6,593,293; 6,555,655; 6,538,109; 6,537,756; 6,521,442; 6,501,009; 6,468,523; 6,326,351; 6,313,378; 6,284,949; 6,281,016; 6,248,536; 6,242,241; 6,221,649; 6,177,615; 6,156,573; 6,153,814; 6,110,464; 6,093,695; 6,063,756; 6,063,597; 6,023,013; 5,959,091; 5,942,664; 5,942,658; 5,880,275; 5,763,245; y 5,763,241), la resistencia a enfermedades fúngicas (patentes estadounidenses n.º 15 6,653,280; 6,573,361; 6,506,962; 6,316,407; 6,215,048; 5,516,671; 5,773,696; 6,121,436; 6,316,407; y 6,506,962), la resistencia a virus (patentes estadounidenses n.º 6,617,496; 6,608,241; 6,015,940; 6,013,864; 5,850,023; y 5,304,730), la resistencia a nematodos (patente estadounidense n.º 6,228,992), la resistencia a enfermedades bacterianas (patente estadounidense n.º 5,516,671), el crecimiento y 20 desarrollo de plantas (patentes estadounidenses n.º 6,723,897 y 6,518,488), la producción de almidón (patentes estadounidenses n.º 6,538,181; 6,538,179; 6,538,178; 5,750,876; 6,476,295), la producción de aceites con modificaciones (patentes estadounidenses n.º 6,444,876; 6,426,447; y 6,380,462), la producción alta de aceite (patentes estadounidenses n.º 6,495,739; 5,608,149; 6,483,008; y 25 6,476,295), el contenido de ácidos grasos con modificaciones (patentes estadounidenses n.º 6,828,475; 6,822,141; 6,770,465; 6,706,950; 6,660,849; 6,596,538; 6,589,767; 6,537,750; 6,489,461; y 6,459,018), una producción alta de proteínas (patentes estadounidenses n.º 6,380,466), la maduración de los frutos (patentes estadounidenses n.º 5,512,466), la nutrición mejorada de animales y seres 30 humanos (patentes estadounidenses n.º 6,723,837; 6,653,530; 6,5412,59; 5,985,605; y 6,171,640), los biopolímeros (patentes estadounidenses n.º USRE37,543; 6,228,623; y 5,958,745 y 6,946,588), la resistencia a dificultades del entorno (patente estadounidense n.º 6,072,103), los péptidos farmacéuticos y péptidos secretables (patentes estadounidenses n.º 6,812,379; 6,774,283; 6,140,075; y 6,080,560), los

rasgos de procesamiento mejorados (patente estadounidense n.º 6,476,295), la capacidad de digestión mejorada (patente estadounidense n.º 6,531,648), el contenido bajo de rafinosa (patente estadounidense n.º 6,166,292), la producción de enzimas industriales (patente estadounidense n.º 5,543,576), el sabor mejorado (patente estadounidense n.º 6,011,199), la fijación de nitrógeno (patente estadounidense n.º 5,229,114), la producción de semillas híbridas (patente estadounidense n.º 5,689,041), la producción de fibra (patentes estadounidenses n.º 6,576,818; 6,271,443; 5,981,834; y 5,869,720) y la producción de biocombustible (patente estadounidense n.º 5,998,700).

[74] De manera alternativa, un gen de interés agronómico puede afectar las características o los fenotipos de plantas que se mencionaron anteriormente al codificar una molécula de ARN que cause la modulación dirigida de la expresión génica de un gen endógeno, por ejemplo, mediante mecanismos no codificantes (véase, p. ej. la patente estadounidense n.º 5,107,065); de ARN inhibidor («RNAi», incluida la modulación de la expresión génica mediante mecanismos mediados por miRNA, siRNA, siRNA de acción en *trans* y sRNA de fase, p. ej., tal como se describe en las solicitudes publicadas US 2006/0200878 y US 2008/0066206, y en la solicitud de patente estadounidense 11/974,469); o mecanismos mediados por la supresión conjunta. El ARN también podría ser una molécula de ARN catalizadora (p. ej., una ribozima o un ribointerruptor; véase, p. ej., U.S. 2006/0200878) diseñada para escindir un producto de mRNA endógeno deseado. Se conocen métodos en la técnica para construir e introducir construcciones en una célula de tal manera que la molécula de ADN pasible de transcripción se transcriba en una molécula capaz de provocar la supresión génica.

Marcadores seleccionables

[75] Los transgenes de marcadores seleccionables también pueden usarse con los elementos reguladores de la invención. Tal como se usa en la presente, la expresión «transgén de marcador seleccionable» hace referencia a cualquier molécula de ADN pasible de transcripción cuya expresión en una planta transgénica, un tejido o una célula, o falta de esta, pueda analizarse o calificarse de alguna manera. Los genes marcadores seleccionables, y sus técnicas de selección y examinación asociadas, para el uso en la puesta en la práctica de la invención se conocen en la técnica e incluyen, de modo no taxativo, moléculas de ADN pasibles de transcripción que codifican la β -glucuronidasa (GUS), la proteína verde fluorescente (GFP), proteínas

que confieren resistencia antibiótica y proteínas que confieren tolerancia a los herbicidas. Se proporciona un ejemplo de un transgén de marcador seleccionable como la SEQ ID NO:21.

Edición genómica

- 5 **[76]** Varias modalidades hacen referencia a una construcción de ADN recombinante que comprende uno o más casetes de expresión que comprenden una secuencia con al menos alrededor de 85 por ciento de identidad de secuencia con respecto a cualquiera de las SEQ ID NOs: 1-20 o un fragmento de estas que se enlaza de manera operativa a una secuencia de ADN heteróloga que codifica una enzima de
- 10 modificación genómica con especificidad de sitio y/o una o más proteínas asociadas cualesquiera para llevar a cabo la modificación del genoma. Este o estos casetes de expresión de nucleasas pueden estar presentes en la misma molécula o vector que un molde donante para la edición con molde (en *cis*) o en una molécula o vector separado (en *trans*). Se conocen en la técnica varios métodos de edición que
- 15 involucran diferentes enzimas de modificación genómica con especificidad de secuencia (o complejos de proteínas y/o ARN guía) que modifican el ADN genómico. En algunas modalidades, una enzima de modificación genómica con especificidad de sitio modifica el genoma al inducir una ruptura de cadena doble (DSB, por sus siglas en inglés) o un corte en un locus o sitio genómico deseado. En algunas modalidades,
- 20 durante el proceso de reparación de la DSB o del corte introducido por la enzima de modificación genómica, un ADN molde donante puede integrarse al genoma en el sitio de la DSB o el corte. En algunas modalidades, durante el proceso de reparación de la DSB o del corte introducido por la enzima de modificación genómica, se puede introducir una mutación de inserción o delección (indel) en el genoma. En algunas
- 25 modalidades, una enzima de modificación genómica con especificidad de sitio comprende una citidina desaminasa. En algunas modalidades, una enzima de modificación genómica con especificidad de sitio comprende una adenina desaminasa. En la presente descripción, las enzimas de modificación genómica con especificidad de sitio incluyen endonucleasas, recombinasas, transposasas,
- 30 desaminasas, helicasas, transcriptasas inversas y cualquier combinación de estas.
- [77]** Varias modalidades hacen referencia a un elemento regulador de genes tal como se describe en la presente enlazado de manera operativa a una molécula de ADN pasible de transcripción heteróloga que codifica uno o más componentes de un sistema de edición genómica. Los sistemas de edición genómica pueden usarse para

introducir una o más inserciones, eliminaciones, sustituciones, modificaciones de bases, translocaciones o inversiones en un genoma de una célula hospedadora. En algunas modalidades, un elemento regulador de genes tal como se describe en la presente está enlazado de manera operativa a una molécula de ADN pasible de transcripción heteróloga que codifica un dominio de unión de ADN con especificidad de secuencia, tal como una proteína efectora CRISPR-Cas, una proteína de dedo de cinc o una proteína activadora de la transcripción (TAL). En algunas modalidades, el dominio de unión de ADN con especificidad de secuencia puede ser una proteína de fusión. En algunas modalidades, un elemento regulador de genes tal como se describe en la presente está enlazado de manera operativa a una molécula de ADN pasible de transcripción heteróloga que codifica una proteína efectora CRISPR-Cas. En algunas modalidades, la proteína efectora CRISPR-Cas se selecciona de entre un sistema CRISPR-Cas Tipo I, un sistema CRISPR-Cas Tipo II, un sistema CRISPR-Cas Tipo III, un sistema CRISPR-Cas Tipo IV, un sistema CRISPR-Cas Tipo V o un sistema CRISPR-Cas Tipo VI. En algunas modalidades, un elemento regulador de genes tal como se describe en la presente está enlazado de manera operativa a una molécula de ADN pasible de transcripción heteróloga que codifica un ARN guía. Tal como se usa en la presente, un «ARN guía» o «gRNA» hace referencia a un ARN que reconoce una secuencia de ADN objetivo y dirige, o «guía», una proteína efectora CRISPR a la secuencia de ADN objetivo. Un ARN guía está conformado por una región complementaria al ADN objetivo (denominada crRNA) y una región que se une a la proteína efectora CRISPR (denominada tracrRNA). Un ARN guía puede ser una sola molécula de ARN (sgRNA) o dos moléculas de ARN separadas (un gRNA de 2 piezas). En algunas modalidades, un gRNA puede comprender además un molde de ARN (pegRNA) para una transcriptasa inversa.

[78] Varias modalidades se vinculan con un elemento regulador de genes tal como se describe en la presente enlazado de manera operativa a una molécula de ADN pasible de transcripción heteróloga que codifica uno o más componentes de un sistema de edición genómica CRISPR-Cas que comprende una proteína efectora CRISPR-Cas y un ARN guía. Los ejemplos de proteínas efectoras CRISPR-Cas incluyen, de modo no taxativo, la proteína efectora Cas9, C2c1, C2c3, C2c4, C2c5, C2c8, C2c9, C2c10, Cas12a (también denominada Cpf1), Cas12b, Cas12c, Cas12d, Cas12e, Cas12h, Cas12i, Cas12g, Cas13a, Cas13b, Cas13c, Cas13d, Cas1, Cas1B, Cas2, Cas3, Cas3', Cas3'', Cas4, Cas5, Cas6, Cas7, Cas8, Cas9 (también conocida

como Csn1 y Csx12), Cas10, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4 (dinG), Csf5, Cas14a, Cas14b y Cas14c. En algunas modalidades, un elemento regulador de genes tal como se describe en la presente está enlazado de manera operativa a una proteína efectora CRISPR-Cas que comprende una mutación en su sitio activo de nucleasa (p. ej., RuvC, HNH, p. ej., sitio RuvC de un dominio de nucleasa Cas12a; p. ej., sitio RuvC y/o sitio HNH de un dominio de nucleasa Cas9). Una proteína efectora CRISPR-Cas que tiene una mutación en su sitio activo de nucleasa y, por lo tanto, ya no comprende actividad de nucleasa, se denomina comúnmente «muerta», por ejemplo, dCas. En algunas modalidades, un polipéptido o dominio de proteína efectora CRISPR-Cas que tiene una mutación en su sitio activo de nucleasa puede tener una actividad alterada o actividad reducida en comparación con la misma proteína efectora CRISPR-Cas sin la mutación. En algunas modalidades, un elemento regulador de genes tal como se describe en la presente está enlazado de manera operativa a una proteína efectora CRISPR-Cas que tiene una mutación en su sitio activo de nucleasa para generar una actividad de nickasa enlazada de manera operativa a una enzima transcriptasa inversa.

Transformación celular

[79] La invención también está dirigida a un método de producción de plantas y células transformadas que comprendan uno o más elementos reguladores enlazados de manera operativa a una molécula de ADN pasible de transcripción.

[80] El término «transformación» hace referencia a la introducción de una molécula de ADN en un hospedador receptor. Tal como se usa en la presente, el término «hospedador» hace referencia a bacterias, hongos o plantas, lo cual incluye cualquier célula, tejido, órgano o progenie de las bacterias, los hongos o las plantas. Los tejidos y las células de plantas de interés particular incluyen protoplastos, callos, raíces, tubérculos, semillas, tallos, hojas, plántulas, embriones y polen.

[81] Tal como se usa en la presente, el término «transformado/a» hace referencia a una célula, un tejido, un órgano o un organismo en el que se haya introducido una molécula de ADN externa, tal como una construcción. La molécula de ADN introducida puede integrarse al ADN genómico de la célula, el tejido, el órgano o el organismo receptor de forma tal que la molécula de ADN introducida sea heredada por la progenie posterior. Un organismo o una célula «transgénico/a» o «transformado/a»

también puede incluir progenie del organismo o de la célula y progenie producida a partir de un programa de fitomejoramiento que emplee ese organismo transgénico como un original en una cruce y presente un fenotipo alterado como resultado de la presencia de una molécula de ADN externa. La molécula de ADN introducida también
5 puede introducirse transitoriamente en la célula receptora de forma tal que la progenie posterior no herede la molécula de ADN introducida. El término «transgénico/a» hace referencia a una bacteria, un hongo o una planta que contiene una o más moléculas de ADN de carácter heterólogo.

[82] Existen muchos métodos muy conocidos por los expertos en la técnica para
10 introducir moléculas de ADN en las células de plantas. El proceso generalmente comprende las etapas de selección de una célula hospedadora adecuada, transformación de la célula hospedadora con un vector y obtención de la célula hospedadora transformada. Los métodos y materiales para transformar las células de las plantas mediante la introducción de una construcción de planta en un genoma de
15 planta en la puesta en práctica de la presente invención pueden incluir cualquiera de los métodos muy conocidos y demostrados. Los métodos adecuados incluyen, de modo no taxativo, infección bacteriana (p. ej., *Agrobacterium*), vectores BAC binarios, suministro directo de ADN (p. ej., mediante transformación mediada por PEG, absorción de ADN mediada por inhibición/desecación, electroporación, agitación con
20 fibras de carburo de silicio y aceleración de partículas recubiertas con ADN), edición de genes (p. ej., los sistemas CRISPR-Cas), entre otros.

[83] Las células hospedadoras pueden ser cualquier célula u organismo, tal como una célula de planta, célula de alga, alga, célula de hongo, hongo, célula bacteriana o célula de insecto. En modalidades específicas, las células hospedadoras y las células
25 transformadas pueden incluir células de plantas de cultivo.

[84] Posteriormente, una planta transgénica puede regenerarse a partir de una célula de planta transgénica de la invención. Mediante el uso de técnicas de fitomejoramiento convencionales o autopolinización, la semilla puede producirse a partir de esta planta transgénica. Esa semilla, y la planta de progenie resultante
30 cultivada a partir de esa semilla, contendrán la molécula de ADN recombinante de la invención y, por lo tanto, serán transgénicas.

[85] Las plantas transgénicas de la invención pueden autopolinizarse para proporcionar semillas para las plantas transgénicas homocigotas de la invención (homocigotas con respecto a la molécula de ADN recombinante) o cruzarse con

plantas no transgénicas o plantas transgénicas diferentes para proporcionar semillas para plantas transgénicas heterocigotas de la invención (heterocigotas con respecto a la molécula de ADN recombinante). En la presente, se hace referencia a esas plantas transgénicas homocigotas y heterocigotas como «plantas de progenie». Las plantas de progenie son plantas transgénicas que descienden de la planta transgénica original y que contienen la molécula de ADN recombinante de la invención. Las semillas producidas mediante el uso de una planta transgénica de la invención pueden cosecharse y usarse para cultivar generaciones de plantas transgénicas, es decir, plantas de progenie de la invención, que comprenden la construcción de la presente invención y que expresan un gen de interés agronómico. Las descripciones de los métodos de fitomejoramiento que se usan comúnmente para diferentes cultivos pueden encontrarse en uno de varios libros de referencia, véase, p. ej., Allard, *Principles of Plant Breeding*, John Wiley & Sons, NY, U. of CA, Davis, CA, 50-98 (1960); Simmonds, *Principles of Crop Improvement*, Longman, Inc., NY, 369-399 (1979); Sneep y Hendriksen, *Plant breeding Perspectives*, Wageningen (ed), Center for Agricultural Publishing and Documentation (1979); Fehr, *Soybeans: Improvement, Production and Uses*, 2.^a Edition, Monograph, 16:249 (1987); Fehr, *Principles of Variety Development, Theory and Technique*, (Vol. 1) y *Crop Species Soybean* (Vol. 2), Iowa State Univ., Macmillan Pub. Co., NY, 360-376 (1987).

[86] Las plantas transformadas pueden analizarse en busca de la presencia del gen o de los genes de interés y el nivel de expresión y/o el perfil conferido por los elementos reguladores de la invención. Los expertos en la técnica son conscientes de los numerosos métodos disponibles para los análisis de las plantas transformadas. Por ejemplo, los métodos para el análisis de las plantas incluyen, de modo no taxativo, la transferencia Southern o Northern, enfoques basados en PCR, análisis bioquímicos, métodos de análisis fenotípicos, evaluaciones de campo y ensayos de inmunodiagnóstico. La expresión de una molécula de ADN pasible de transcripción puede medirse mediante el uso de reactivos y métodos TaqMan® (Applied Biosystems, Foster City, CA) como lo describe el fabricante y los tiempos de ciclo de PCR determinados mediante el uso de la matriz de prueba TaqMan®. De manera alternativa, los reactivos Invader® (Third Wave Technologies, Madison, WI) y los métodos descritos por el fabricante pueden usarse para evaluar la expresión transgénica.

[87] La invención también proporciona partes de una planta de la invención. Las partes de las plantas incluyen, de modo no taxativo, hojas, tallos, raíces, tubérculos, semillas, endospermo, óvulos y polen. Las partes de plantas de la invención pueden ser viables, no viables, regenerables y/o no regenerables. La invención también incluye y proporciona células de plantas transformadas que comprenden una molécula de ADN de la invención. Las células de plantas transformadas o transgénicas de la invención incluyen células de plantas regenerables y/o no regenerables.

[88] La invención también proporciona un producto básico que se produce a partir de una planta transgénica o parte de esta que contiene la molécula de ADN recombinante de la invención. Los productos básicos de la invención contienen una cantidad detectable de ADN que comprende una secuencia de ADN seleccionada dentro del grupo que consiste en las SEQ ID NOS:1-20. Tal como se usa en la presente, un «producto básico» hace referencia a cualquier composición o producto conformado por material derivado de una planta transgénica, semilla, célula de planta o parte de planta que contiene la molécula de ADN recombinante de la invención. Los productos básicos incluyen, de modo no taxativo, semillas procesadas, granos, partes de plantas y sémola. Un producto básico de la invención contendrá una cantidad detectable de ADN correspondiente a la molécula de ADN recombinante de la invención. La detección de uno o más de estos ADN en una muestra puede usarse para determinar el contenido o la fuente del producto básico. Se puede usar cualquier método estándar de detección de moléculas de ADN, incluidos los métodos de detección descritos en la presente.

[89] La invención puede entenderse más fácilmente a través de la referencia a los ejemplos que siguen, que se proporcionan con fines ilustrativos, y no se pretende que limiten la invención, a menos que se especifique. Los expertos en la técnica deberían darse cuenta de que las técnicas descritas en los ejemplos que siguen representan técnicas descubiertas por los inventores que funcionan bien en la puesta en práctica de la invención. Sin embargo, los expertos en la técnica deberían darse cuenta, a la luz de la presente descripción, que es posible realizar muchos cambios a las modalidades específicas descritas y aun así obtener un resultado similar sin apartarse del espíritu y alcance de la invención; por lo tanto, todo objeto establecido o mostrado debe interpretarse como ilustrativo y no en sentido taxativo.

EJEMPLOS

Ejemplo 1

Identificación y clonación de los elementos reguladores

[90] En este ejemplo, se describe la identificación, síntesis y clonación de elementos de expresión reguladores derivados de *Zea mays*.

[91] Se identificaron genes con preferencia por la expresión en el polen mediante el uso de datos de transcriptomas públicos y privados. Se realizó un examen bioinformático de cada locus génico y se identificaron los correspondientes promotores, líderes, intrones y 3' UTR de los genes. Los EXP, promotores/líderes, intrones y 3' UTR identificados se presentan en la Tabla 1 a continuación.

Tabla 1. Grupos de elementos de expresión reguladores, promotores, líderes, intrones y 3' UTR.

Notación	SEQ ID NO:	Tamaño (pb)	Descripción y/o elementos reguladores de EXP enlazados en la dirección 5' → 3' (SEQ ID NO):
P-Zm.GRMZM2G487322:2	1	1973	Promotor + Líder
T-Zm.GRMZM2G487322:2	2	501	3' UTR
P-Zm.GRMZM2G339781:1	3	1998	Promotor + Líder
T-Zm.GRMZM2G339781:1	4	501	3' UTR
EXP-Zm.Xet:1	5	2196	EXP: P-Zm.Xet:1 (SEQ ID NO:6, I-Zm.Xet:1 (SEQ ID NO:7)
P-Zm.Xet:1	6	1998	Promotor
I-Zm.Xet:1	7	191	Intrón
T-Zm.Xet:1	8	498	3' UTR
EXP-Zm.Sat6:1	9	2183	EXP: P-Zm.Sat6:1 (SEQ ID NO:10, I-Zm.Sat6:1 (SEQ ID NO:11)
P-Zm.Sat6:1	10	1998	Promotor
I-Zm.Sat6:1	11	178	Intrón
T-Zm.Sat6:1	12	498	3' UTR
EXP-Zm.GRMZM2G049726:1	13	2145	EXP: P-Zm.GRMZM2G049726:1 (SEQ ID NO:14), I-Zm.GRMZM2G049726:1 (SEQ ID NO:15)
P-Zm.GRMZM2G049726:1	14	1998	Promotor
I-Zm.GRMZM2G049726:1	15	140	Intrón

Notación	SEQ ID NO:	Tamaño (pb)	Descripción y/o elementos reguladores de EXP enlazados en la dirección 5' → 3' (SEQ ID NO):
T-Zm.GRMZM2G049726:1	16	501	3' UTR
P-Zm.GRMZM2G141762:1	17	1998	Promotor + Líder
P-Zm.DSUL:1	18	1568	Promotor + Líder
P-Zm.GRMZM2G512113:1	19	2000	Promotor + Líder
T-Zm.GRMZM2G512113:1	20	704	3' UTR

[92] Los EXP, los promotores/líderes y las 3' UTR que se identifican se sintetizaron y clonaron mediante el uso de métodos conocidos en la técnica en construcciones de vectores de transformación de plantas de carácter binario, en un casete de expresión usado para impulsar la expresión de la β -glucuronidasa (GUS) para evaluar su actividad en plantas de maíz transformadas de manera estable, tal como se describirá en el Ejemplo 2 a continuación.

Ejemplo 2

Análisis de elementos reguladores que impulsan la expresión de GUS en plantas de maíz transformadas de manera estable

[93] Las plantas de maíz se transformaron con un vector, específicamente un vector de expresión en plantas que contiene elementos reguladores de prueba que impulsan la expresión del transgén de la β -glucuronidasa (GUS). Las plantas resultantes se analizaron en busca de la expresión de la proteína GUS, para evaluar el efecto del elemento regulador seleccionado en la expresión.

[94] Las plantas de maíz se transformaron con construcciones de expresión de GUS en plantas. Los elementos reguladores se clonaron en un vector de expresión en plantas de base mediante el uso métodos estándares conocidos en la técnica. Los vectores de expresión de plantas resultantes contenían una región del borde izquierdo de *Agrobacterium tumefaciens* (B-AGRtu.borde izquierdo), un primer casete de selección de transgenes usado para la selección de células de plantas transformadas que confiere resistencia al herbicida glifosato, un segundo casete de transgenes para evaluar la actividad de los elementos reguladores donde se incluye un promotor y un líder, opcionalmente enlazados de manera operativa en 5' a un intrón en determinados diseños de vectores, con enlace operativo a una secuencia de codificación para GUS donde se incluye un intrón procesable (SEQ ID NO:21), con enlace operativo a una 3'

UTR, y una región del borde derecho de *Agrobacterium tumefaciens* (B-AGRtu.borde derecho).

[95] Las configuraciones de los elementos de expresión para los casetes de expresión de GUS para cada construcción se presentan en la Tabla 2. La Construcción-1 a Construcción-5 comprenden casetes de expresión de GUS conformados por el promotor y líder naturales enlazados de manera operativa (indicados como «P-») o el promotor, líder e intrón naturales enlazados de manera operativa (indicados como «EXP-»). Los casetes de expresión de GUS de la Construcción-1 y Construcción-2 no comprenden un intrón. De la Construcción-1 a la Construcción-5 y la Construcción-8 comprenden la 3' UTR natural del mismo promotor/líder o promotor/líder/intrón del locus génico. En la Construcción-6 y la Construcción-7, el promotor y líder naturales (indicados como «P-») estaban enlazados de manera operativa a una 3' UTR y un intrón con posibilidad de expresión en plantas que no formaban parte del gen natural. La Construcción-8 también comprende un intrón expresable en plantas que no formaba parte del gen natural.

Tabla 2. Configuración del casete de expresión de GUS para la Construcción-1 a la Construcción-7.

Construcción	EXP o Promotor/Líder	EXP o Promotor/Líder, 3' UTR SEQ ID NO:	3' UTR	UTR 3', SEQ ID NO:
Construcción-1	P- Zm.GRMZM2G487322: 2 (No intron)	1	T- Zm.GRMZM2G487 322:2	2
Construcción-2	P- Zm.GRMZM2G339781: 1 (No Intron)	3	T- Zm.GRMZM2G339 781:1	4
Construcción-3	EXP-Zm.Xet:1	5	T-Zm.Xet:1	8
Construcción-4	EXP-Zm.Sat6:1	9	T-Zm.Sat6:1	12
Construcción-5	EXP- Zm.GRMZM2G049726: 1	13	T- Zm.GRMZM2G049 726:1	16

Construcción-6	P- Zm.GRMZM2G141762: 1 + Intrón de planta	17	3' UTR de planta	
Construcción-7	P-Zm.DSUL:1 + Intrón de planta	18	3' UTR de planta	
Construcción-8	P- Zm.GRMZM2G512113: 1 + Intrón de planta	19	T- Zm.GRMZM2G512 113:1	20

[96] Las células de plantas de maíz se transformaron mediante el uso de las construcciones de vectores de transformación binarias descritas anteriormente mediante la transformación mediada por *Agrobacterium*, como se sabe bien en la técnica. Se indujo que las células de plantas transformadas resultantes formaran plantas de maíz enteras.

[97] El análisis cualitativo y cuantitativo de GUS se usó para evaluar la actividad del elemento de expresión en órganos y tejidos de plantas seleccionados en las plantas transformadas. Para el análisis cualitativo de la expresión de GUS mediante tinción histoquímica, se incubaron tejidos de montaje completo o seccionados con solución de tinción de GUS que contenía 1 mg/mL de X-Gluc (5-bromo-4-cloro-3-indolil-b-glucurónido) durante 5 h a 37 °C y se destiñeron con EtOH al 35 % y ácido acético al 50 %. La expresión de GUS se determinó cualitativamente mediante la inspección visual de los órganos o tejidos de plantas seleccionados en busca de coloración azul bajo un microscopio de disección o compuesto.

[98] Para el análisis cuantitativo de la expresión de GUS mediante ensayos enzimáticos, se extrajo la proteína total de tejidos seleccionados de plantas de maíz transformadas. Se incubaron uno o dos microgramos de proteína total con el sustrato fluorogénico, 4-metileumbeliferil- β -D-glucurónido (MUG) a una concentración de 1 mM en un volumen de reacción total de 50 microlitros. Después de 1 h de incubación a 37 °C, la reacción se detuvo mediante la adición de 350 microlitros de solución de bicarbonato de sodio 200 mM. El producto de reacción, 4-metilumbeliferona (4-MU), es fluorescente al máximo con un pH elevado, donde el grupo hidroxilo se ioniza. La adición de la solución básica de carbonato de sodio detiene simultáneamente el ensayo y ajusta el pH para cuantificar el producto fluorescente 4-MU. La cantidad de 4-MU formada se estimó mediante la medición de su fluorescencia mediante el uso de un lector de microplacas Omega FLUOstar (BMG LABTECH) (excitación a 355 nm,

emisión a 460 nm). Los valores de actividad de GUS se proporcionan en nmoles de 4-MU/hora/mg de proteína total.

[99] Se obtuvieron muestras de los siguientes tejidos en busca de la expresión de GUS en la generación R₀: hoja y raíz en etapa V4; hoja y raíz en etapa V7; hoja, flor/antera y polen en etapa VT; mazorca/seda en etapa R1; y embrión de semilla y endospermo de semilla en etapa R3 21 días después de la polinización (DAP, por sus siglas en inglés).

[100] Las Tablas 3 y 4 muestran la expresión de GUS cuantitativa media para los tejidos de los que se obtuvieron muestras, donde «bdl» indica un nivel por debajo de la detección y «NA» indica que no se sometió a ensayo. Los intervalos de expresión de GUS para la flor/antera y polen en etapa VT, y la mazorca/seda en etapa R1 también se presentan en la Tabla 4.

Tabla 3. Expresión media de GUS de plantas de maíz transformadas de manera estable para la hoja y raíz en etapa V4; hoja y raíz en etapa V7; hoja en etapa VT; y embrión y endospermo 21 DAP.

Construcción	Raíz V4	Hoja V4	Raíz V7	Hoja V7	Hoja VT	Embrión 21 DAP	Endospermo 21 DAP
Construcción-1	23	25	30	24	24	40	22
Construcción-2	39	37	39	40	22	49	37
Construcción-3	22	27	75	59	bdl	NA	NA
Construcción-4	36	40	37	36	22	48	41
Construcción-5	22	30	27	28	25	28	23
Construcción-6	103	62	39	53	42	329	332
Construcción-7	70	49	48	35	42	41	50
Construcción-8	46	bdl	25	bdl	bdl	20	71

Tabla 4. GUS media e intervalo de expresión de plantas de maíz transformadas de manera estable para flores/anteras y polen VT, y mazorca/seda R1.

Construcción	Flor, anteras VT	Flor, polen VT	Mazorca, seda R1	Intervalo de flor, anteras VT	Intervalo de flor, polen VT	Intervalo de mazorca, seda R1
Construcción-1	433	1202	23	27-2731	189 - 3998	23-24
Construcción-2	434	1040	bdl	21-1228	47 - 1965	bdl

Construcción-3	60	274	29	27-114	53 - 606	21-39
Construcción-4	27	110	bdl	21-39	77 - 139	bdl
Construcción-5	25	91	23	21-28	44 - 241	44-241
Construcción-6	1182	NA	NA	634-1897	NA	NA
Construcción-7	987	NA	1265	509-1279	NA	568-2151
Construcción-8	63	NA	bdl	31-114	NA	bdl

[101] Tal como se puede ver en las Tablas 3 y 4, muchos de los casetes de expresión de GUS en cada una de las construcciones demostraron una mayor expresión en las flores/antera y/o el polen VT con la excepción de la Construcción-6 y Construcción-7. Con respecto a la Construcción-6, se midió de manera cuantitativa una expresión alta en la antera VT y en el embrión y endospermo 21 DAP. Para la Construcción-7, se midió de manera cuantitativa una expresión alta de GUS en la antera VT y la mazorca/seda R1. Para la Construcción-8, la expresión de la antera VT fue menor en relación con la de la Construcción-6 y la Construcción-7. No se determinó de manera cuantitativa expresión de GUS en el polen VT para la Construcción-6, Construcción-7 y Construcción-8. La Construcción-1 y Construcción-2 mostraron una expresión alta de GUS en polen VT con una expresión más alta de GUS en la flor/antera VT que en otros tejidos.

[102] Los tejidos seleccionados de los eventos transformados también se observaron microscópicamente para determinar el aspecto de la expresión cualitativa observada en la tinción de GUS de los tejidos. La Tabla 5 a continuación resume las observaciones hechas de estos tejidos.

Tabla 5. Expresión de GUS cualitativa de plantas de maíz transformadas de manera estable.

Construcción	Expresión de GUS cualitativa
Construcción-1	La tinción en la antera está asociada a polen.
Construcción-2	La tinción en la antera está asociada a polen.
Construcción-3	La tinción en la antera está asociada a polen.
Construcción-4	La tinción en la antera está asociada a polen.
Construcción-5	La tinción en la antera está asociada a polen.
Construcción-6	Tinción visible en polen VT, gluma/pálea y pedicelo; mesófilo de hoja V7, vaina del haz, haz vascular y células de guarda; corteza radicular, endodermis, epidermis y periciclo V7.

Construcción	Expresión de GUS cualitativa
Construcción-7	Tinción visible en polen VT, gluma/pálea y pedicelo; seda R1; mesófilo de hoja V7, vaina del haz, haz vascular y células de guarda; corteza radicular, endodermis, epidermis y periciclo V7.
Construcción-8	Tinción visible en polen VT; punta de raíz V4; capa de células de transferencia del endospermo basal R3, aleurona, embrión, endospermo, pedicelo, pericarpio y escutelo.

[103] Tal como se puede ver en la Tabla 5, la expresión de GUS para la Construcción-1 a la Construcción-5 se observó principalmente en el polen VT y, por lo tanto, proporcionó una expresión más alta en la flor/antera de VT cuando se midió cuantitativamente. Los elementos de expresión en los casetes de expresión de GUS de la Construcción-1 a Construcción-5 son, por tanto, elementos de expresión «preferidos por el polen». Con respecto a la expresión de GUS de la Construcción-6 y la Construcción-7, la expresión de GUS se observó no solo en el polen de VT, sino también en otros tejidos, tales como las células de la hoja y raíz V7. Curiosamente, la expresión de GUS de la Construcción-7 también se observó en la seda R1. Para la Construcción-8, fue claramente visible la tinción del polen VT, así como la tinción en la punta de raíz V4 y en los tejidos del embrión y endospermo 21DAP R3.

[104] Por lo tanto, los elementos de expresión comprendidos dentro de los casetes de expresión de GUS de la Construcción-1 a la Construcción-5 muestran un patrón de expresión preferido por el polen. Los elementos de expresión comprendidos dentro de los casetes de expresión de GUS de la Construcción-6, Construcción-7 y Construcción-8, aunque muestran expresión en el polen, también se expresan en otros tejidos de las plantas de maíz transformadas de manera estable.

[105] Habiéndose ilustrado y descrito los principios de la presente invención, debería ser evidente para los expertos en la técnica que la invención puede modificarse en su disposición y detalles sin apartarse de esos principios. Se reivindican todas las modificaciones que se encuentran dentro del espíritu y alcance de las reivindicaciones. Todas las publicaciones y documentos de patentes publicados que se citan en la presente se incorporan mediante esta referencia en la misma medida que si se indicara de manera específica e individual la incorporación mediante referencia de cada publicación o solicitud de patente particular.

REIVINDICACIONES:

1. Una molécula de ADN recombinante que comprende una secuencia de ADN que se selecciona dentro del grupo que consiste en:
 - a) una secuencia con al menos 85 por ciento de identidad de secuencia con respecto a cualquiera de las SEQ ID NOS:1-20;
 - b) una secuencia que comprende cualquiera de las SEQ ID NOS:1-20; y
 - c) un fragmento de cualquiera de las SEQ ID NOS: 1-20, donde el fragmento tiene actividad de regulación de genes;
 donde dicha secuencia está enlazada de manera operativa a una molécula de ADN pasible de transcripción heteróloga.
2. La molécula de ADN recombinante de la reivindicación 1, donde dicha secuencia tiene al menos 90 por ciento de identidad de secuencia con respecto a la secuencia de ADN de cualquiera de las SEQ ID NOS:1-20.
3. La molécula de ADN recombinante de la reivindicación 1, donde dicha secuencia tiene al menos 95 por ciento de identidad de secuencia con respecto a la secuencia de ADN de cualquiera de las SEQ ID NOS:1-20.
4. La molécula de ADN recombinante de la reivindicación 1, donde la secuencia de ADN comprende actividad de regulación de genes.
5. La molécula de ADN recombinante de la reivindicación 1, donde la molécula de ADN pasible de transcripción heteróloga comprende un gen de interés agronómico.
6. La molécula de ADN recombinante de la reivindicación 5, donde el gen de interés agronómico confiere tolerancia a herbicidas en las plantas.
7. La molécula de ADN recombinante de la reivindicación 5, donde el gen de interés agronómico confiere resistencia a las plagas en las plantas.
8. La molécula de ADN recombinante de la reivindicación 1, donde la molécula de ADN pasible de transcripción heteróloga codifica un dsRNA, un miRNA o un siRNA.
9. Una célula de planta transgénica que comprende una molécula de ADN recombinante que comprende una secuencia que se selecciona dentro del grupo que consiste en:
 - a) una secuencia con al menos 85 por ciento de identidad de secuencia con respecto a cualquiera de las SEQ ID NOS:1-20;
 - b) una secuencia que comprende cualquiera de las SEQ ID NOS:1-20; y
 - c) un fragmento de cualquiera de las SEQ ID NOS: 1-20, donde el fragmento tiene actividad de regulación de genes;

donde dicha secuencia está enlazada de manera operativa a una molécula de ADN pasible de transcripción heteróloga.

10. La célula de planta transgénica de la reivindicación 9, donde dicha célula de planta transgénica es una célula de planta monocotiledónea.

5 11. La célula de planta transgénica de la reivindicación 9, donde dicha célula de planta transgénica es una célula de planta dicocotiledónea.

12. Una planta transgénica, o parte de esta, que comprende la molécula de ADN recombinante de la reivindicación 1.

10 13. Una planta de progenie de la planta transgénica de la reivindicación 12, o una parte de esta, donde la planta de progenie o parte de esta comprende dicha molécula de ADN recombinante.

14. Una semilla transgénica, donde la semilla comprende la molécula de ADN recombinante de la reivindicación 1.

15 15. Un método de producción de un producto básico que comprende obtener una planta transgénica o parte de esta de acuerdo con la reivindicación 12 y producir el producto básico a partir de esta.

16. El método de la reivindicación 15, donde el producto básico consiste en semillas, semillas procesadas, concentrado de proteínas, aislado de proteínas, almidón, granos, partes de plantas, aceite de semillas, biomasa, harina y sémola.

20 17. Un método de expresión de una molécula de ADN pasible de transcripción que comprende obtener una planta transgénica de acuerdo con la reivindicación 12 y cultivar la planta, donde se expresa el ADN pasible de transcripción.

RESUMEN

La invención proporciona construcciones y moléculas de ADN recombinante, así como sus secuencias de nucleótidos, útiles para modular la expresión génica en las plantas. La invención también proporciona plantas transgénicas, células de plantas, partes de plantas y semillas que comprenden las moléculas de ADN recombinante enlazadas de manera operativa a las moléculas de ADN pasible de transcripción heterólogas, así como sus métodos de uso.

5