

TITULO DE LA INVENCION

COMPOSICIONES Y MÉTODOS PARA MODIFICAR LA DETERMINACIÓN VEGETAL

INCORPORACIÓN DE LISTADO DE SECUENCIAS

[001] Esta solicitud reivindica el beneficio de la Solicitud Provisional de los EE.UU. con Nro. de Acta 63/278.903, presentada el 12 de noviembre de 2021, cuya divulgación se incorpora a la presente a modo de referencia en su totalidad.

INCORPORACIÓN DE LISTADO DE SECUENCIAS

[002] Se presenta electrónicamente y se incorpora a la presente a modo de referencia el listado de secuencias contenido en el archivo denominado “MONS522US_ST26.xml”, cuyo tamaño es de 265 kilobytes según la medición del sistema operativo Microsoft Windows y que fue creado el 07 de noviembre de 2022.

CAMPO DE LA INVENCION

[003] La presente divulgación está relacionada con el campo de la biotecnología agrícola, y más específicamente con métodos y composiciones para la edición genómica en plantas.

ANTECEDENTES DE LA INVENCION

[004] Las tecnologías para la edición genómica de precisión son poderosas herramientas para manipular la expresión y la función genética y tienen el potencial de mejorar rasgos agrícolas de importancia. En la técnica existe la continua necesidad de desarrollar composiciones y métodos novedosos para editar el genoma vegetal de manera efectiva y eficiente en diversos cultivos vegetales para modificar la determinación, aumentar el rendimiento, reducir el encamado y lograr otros resultados beneficiosos.

SÍNTESIS

[005] En la presente se proporcionan plantas, semillas de plantas, partes de plantas o células vegetales modificadas, que comprenden una modificación que reduce la expresión o la actividad de TFL1, o un homólogo de este, en comparación con la expresión o la actividad de TFL1 o su homólogo en una planta, semilla de planta, parte de planta o célula vegetal, por lo demás idéntica, que carece de la modificación. En algunas realizaciones, la modificación está presente en al menos un alelo de un gen *TFL1* endógeno o un homólogo de este. El gen *TFL1* puede ser un gen *TFL1b*. En realizaciones particulares, el gen *TFL1* o su homólogo codifica una proteína que tiene al menos 70%, al menos 75%, al menos 80%, al menos 85%, al menos 90%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99%, al menos 99,5%, o 100% de identidad de secuencia con SEQ ID NO:2. En otras realizaciones, la modificación es en una región no codificadora del gen *TFL1* o su homólogo, cuyos ejemplos no limitantes incluyen un promotor, un intrón, una región 5' no traducida, una región 3' no traducida, y combinaciones de cualquiera de estos. La planta puede ser, o una semilla de planta, parte de planta o célula vegetal proporcionada en la presente pueden ser de, por ejemplo, una planta leguminosa, una planta de algodón, una planta de canola, una planta de maíz, una planta de sorgo, una planta de arroz, una planta de trigo, una planta de cebada, una planta de tomate o una planta de pimiento, u otro cultivo, ornamental u otro tipo de especie de planta. En realizaciones adicionales, la planta leguminosa es una planta de soja, una planta de poroto, una planta de arveja, una planta de garbanzo, una planta de alfalfa, una planta de maní, una planta de algarrobo, una planta de lenteja, o una planta de regaliz. En realizaciones específicas, la planta leguminosa es una planta de soja. En algunas realizaciones, la planta, semilla de planta, parte de planta o célula vegetal es heterocigota para la modificación, y en otras realizaciones, la planta, semilla de planta, parte de planta o célula vegetal es homocigota para la modificación. En ciertas realizaciones, la planta, semilla de planta, parte de planta o célula vegetal se define como que comprende una primera modificación en un primer alelo del gen *TFL1* y una segunda modificación en un segundo alelo del gen *TFL1*, donde la primera modificación y la segunda modificación son diferentes entre sí.

[006] Una planta, semilla de planta, parte de planta o célula vegetal proporcionada en la presente, en ciertas realizaciones, comprende una modificación que reduce la expresión o la actividad de TFL1, o un homólogo de este, donde la modificación comprende una delección, una inserción, una

sustitución, una inversión o una combinación de estas. En algunas realizaciones, por ejemplo, la modificación está localizada a aproximadamente 100, 125, 150, 175, 200, 225, 250, 275, o 300 nucleótidos o más del extremo 3' de una secuencia seleccionada del grupo que consiste en SEQ ID NO:4, 67-77, 79 y 81. En otras realizaciones, una modificación en el promotor de al menos un alelo del gen *TFL1b* está comprendida dentro de una región genómica entre las posiciones nucleotídicas 1237 y 1570 de la secuencia de referencia SEQ ID NO:4. En algunas realizaciones, la modificación comprende una delección de al menos aproximadamente 10, al menos aproximadamente 15, al menos aproximadamente 20, al menos aproximadamente 25, al menos aproximadamente 30, al menos aproximadamente 35, al menos aproximadamente 40, al menos aproximadamente 45, al menos aproximadamente 50, al menos aproximadamente 55, al menos aproximadamente 60, al menos aproximadamente 65, al menos aproximadamente 70, al menos aproximadamente 75, al menos aproximadamente 80, al menos aproximadamente 85, al menos aproximadamente 90, al menos aproximadamente 95, al menos aproximadamente 100, al menos aproximadamente 125, o al menos aproximadamente 150 nucleótidos consecutivos. La planta, semilla de planta, parte de planta o célula vegetal también puede comprender, por ejemplo, una modificación en al menos un alelo de un promotor del gen *TFL1b*, donde la modificación se selecciona del grupo que consiste en: una delección de 30 pares de bases del nucleótido 1539 al nucleótido 1568, en comparación con la secuencia de referencia SEQ ID NO: 4; una delección de 388 pares de bases del nucleótido 1217 al nucleótido 1604, en comparación con la secuencia de referencia SEQ ID NO: 4; una delección de 112 pares de bases del nucleótido 1518 al nucleótido 1629, en comparación con la secuencia de referencia SEQ ID NO: 4; una delección de 272 pares de bases del nucleótido 951 al nucleótido 1222, en comparación con la secuencia de referencia SEQ ID NO: 4; una delección de 44 pares de bases del nucleótido 1364 al nucleótido 1407, en comparación con la secuencia de referencia SEQ ID NO: 4; una delección de 150 pares de bases del nucleótido 1367 al nucleótido 1516, en comparación con la secuencia de referencia SEQ ID NO: 4; una delección de 1053 pares de bases del nucleótido 754 al nucleótido 1806, en comparación con la secuencia de referencia SEQ ID NO: 4; una delección de 104 pares de bases del nucleótido 1216 al nucleótido 1319, en comparación con la secuencia de referencia SEQ ID NO: 4; una delección de 807 pares de bases del nucleótido 1159 al nucleótido 1965, en comparación con la secuencia de referencia SEQ ID NO: 4; una delección de 455 pares de bases del nucleótido 760 al

nucleótido 1214, en comparación con la secuencia de referencia SEQ ID NO: 4; una delección de 90 pares de bases del nucleótido 939 al nucleótido 1208, en comparación con la secuencia de referencia SEQ ID NO: 4; una inversión de 37 pares de bases, donde la secuencia desde el nucleótido 1029 al nucleótido 1065 de SEQ ID NO: 4 ha sido eliminada, invertida y reinsertada en la misma localización; una delección de 599 pares de bases del nucleótido 1066 al nucleótido 1664, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 601 pares de bases desde el nucleótido 952 hasta el nucleótido 1552, en comparación con la secuencia de referencia SEQ ID NO: 4; una delección de 132 pares de bases desde el nucleótido 1677 hasta el nucleótido 1808, en comparación con la secuencia de referencia SEQ ID NO: 4; una delección de 35 pares de bases desde el nucleótido 1524 hasta el nucleótido 1558, en comparación con la secuencia de referencia SEQ ID NO: 4; una delección de 930 pares de bases desde el nucleótido 476 hasta el nucleótido 1405, en comparación con la secuencia de referencia SEQ ID NO: 4; una delección de 195 pares de bases desde el nucleótido 1365 hasta el nucleótido 1559, en comparación con la secuencia de referencia SEQ ID NO: 4; una delección de 633 pares de bases desde el nucleótido 928 hasta el nucleótido 1560, en comparación con la secuencia de referencia SEQ ID NO: 4; una delección de 1221 pares de bases desde el nucleótido 593 hasta el nucleótido 1813, en comparación con la secuencia de referencia SEQ ID NO: 4; una delección de 5 pares de bases desde el nucleótido 1552 hasta el nucleótido 1556, en comparación con la secuencia de referencia SEQ ID NO: 4; una delección de 29 pares de bases desde el nucleótido 1537 hasta el nucleótido 1565, en comparación con la secuencia de referencia SEQ ID NO: 4; una delección de 38 pares de bases desde el nucleótido 1209 hasta el nucleótido 1246, en comparación con la secuencia de referencia SEQ ID NO: 4; una delección de 49 pares de bases desde el nucleótido 1552 hasta el nucleótido 1600, en comparación con la secuencia de referencia SEQ ID NO: 4; una delección de 49 pares de bases desde el nucleótido 1553 hasta el nucleótido 1601, en comparación con la secuencia de referencia SEQ ID NO: 4; una delección de 11 pares de bases desde el nucleótido 1368 hasta el nucleótido 1378, en comparación con la secuencia de referencia SEQ ID NO: 4; una delección de 28 pares de bases desde el nucleótido 1368 hasta el nucleótido 1395, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 41 pares de bases desde el nucleótido 1344 hasta el nucleótido 1384, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 235 pares de bases desde el nucleótido 1219 hasta el nucleótido 1453, en comparación

con la secuencia de referencia SEQ ID NO:4; una delección de 13 pares de bases desde el nucleótido 1215 hasta el nucleótido 1227, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 7 pares de bases desde el nucleótido 1370 hasta el nucleótido 1376, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 337 pares de bases desde el nucleótido 1220 hasta el nucleótido 1556, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 161 pares de bases desde el nucleótido 1216 hasta el nucleótido 1376, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 17 pares de bases desde el nucleótido 1368 hasta el nucleótido 1384, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 283 pares de bases desde el nucleótido 1366 hasta el nucleótido 1648, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 120 pares de bases desde el nucleótido 1370 hasta el nucleótido 1489, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 75 pares de bases desde el nucleótido 1541 hasta el nucleótido 1615, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 9 pares de bases desde el nucleótido 1367 hasta el nucleótido 1375, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 51 pares de bases desde el nucleótido 1551 hasta el nucleótido 1601, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 7 pares de bases desde el nucleótido 1371 hasta el nucleótido 1377, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 186 pares de bases desde el nucleótido 1368 hasta el nucleótido 1553, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 14 pares de bases desde el nucleótido 1365 hasta el nucleótido 1378, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 16 pares de bases desde el nucleótido 1365 hasta el nucleótido 1380, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 43 pares de bases desde el nucleótido 1552 hasta el nucleótido 1594, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 10 pares de bases desde el nucleótido 1367 hasta el nucleótido 1376, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 131 pares de bases desde el nucleótido 1552 hasta el nucleótido 1682, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 438 pares de bases desde el nucleótido 954 hasta el nucleótido 1391, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 26 pares de bases desde el nucleótido 1535 hasta el nucleótido 1560, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 111 pares de bases

desde el nucleótido 1534 hasta el nucleótido 1644, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 663 pares de bases desde el nucleótido 937 hasta el nucleótido 1599, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 98 pares de bases desde el nucleótido 1551 hasta el nucleótido 1648, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 34 pares de bases desde el nucleótido 1526 hasta el nucleótido 1559, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 79 pares de bases desde el nucleótido 1528 hasta el nucleótido 1606, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 61 pares de bases desde el nucleótido 1542 hasta el nucleótido 1602, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 381 pares de bases desde el nucleótido 1214 hasta el nucleótido 1594, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 187 pares de bases desde el nucleótido 1368 hasta el nucleótido 1554, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 109 pares de bases desde el nucleótido 1369 hasta el nucleótido 1477, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 5 pares de bases desde el nucleótido 1550 hasta el nucleótido 1554, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 1267 pares de bases desde el nucleótido 734 hasta el nucleótido 2000, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 190 pares de bases desde el nucleótido 1371 hasta el nucleótido 1560, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 36 pares de bases desde el nucleótido 1537 hasta el nucleótido 1572, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 46 pares de bases desde el nucleótido 1541 hasta el nucleótido 1586, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 5 pares de bases desde el nucleótido 1552 hasta el nucleótido 1556, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 955 pares de bases desde el nucleótido 669 hasta el nucleótido 1623, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 38 pares de bases desde el nucleótido 1521 hasta el nucleótido 1558, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 109 pares de bases desde el nucleótido 1369 hasta el nucleótido 1477, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 15 pares de bases desde el nucleótido 1540 hasta el nucleótido 1554, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 458 pares de bases desde el nucleótido 1217 hasta el nucleótido 1674, en comparación

con la secuencia de referencia SEQ ID NO:4; una delección de 81 pares de bases desde el nucleótido 1546 hasta el nucleótido 1626, en comparación con la secuencia de referencia SEQ ID NO:4; una delección de 89 pares de bases desde el nucleótido 1547 hasta el nucleótido 1635, en comparación con la secuencia de referencia SEQ ID NO:4; y combinaciones de cualquiera de estas. En algunas realizaciones, la planta, semilla de planta, parte de planta o célula vegetal comprende una secuencia de polinucleótidos seleccionada del grupo que consiste en SEQ ID NO:30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127 y 128. Una planta, semilla de planta, parte de planta o célula vegetal proporcionada en la presente también puede comprender, por ejemplo, una secuencia cromosómica en el gen *TFL1b* que tiene al menos 70%, al menos 75%, al menos 80%, al menos 85%, al menos 90%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99%, al menos 99,5% o 100% de identidad de secuencia con SEQ ID NO:4 en las regiones fuera de la delección, la inserción, la sustitución o la inversión.

[007] También se proporcionan en la presente plantas modificadas o las semillas, partes de planta, células de estas, que comprenden una modificación que altera el fenotipo de determinación de la planta en relación con una planta por lo demás idéntica que carece de la modificación. En ciertas realizaciones, la modificación aumenta la determinación de la planta, en comparación con la determinación de una planta por lo demás idéntica, que carece de la modificación. En algunas realizaciones, la planta modificada alcanza antes la fecha de su floración terminal, exhibe una tasa de encamado reducida, exhibe sustancialmente el mismo rendimiento o un rendimiento aumentado, exhibe menor susceptibilidad a la enfermedad fúngica, o cualquier combinación posible de estas características, en comparación con una planta por lo demás idéntica, que carece de la modificación.

[008] En ciertas realizaciones, se proporciona un polinucleótido que comprende una secuencia seleccionada del grupo que consiste en SEQ ID NO:30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127 y 128. También se proporciona un ARN guía que comprende

una secuencia de polinucleótidos seleccionada del grupo que consiste en SEQ ID NO:13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28 y 29.

[009] Asimismo se divulga en la presente un método para producir una planta que comprende un gen *TFLI* modificado, el método comprende: a) introducir una modificación en al menos un sitio diana en un gen *TFLI* endógeno o un homólogo de este de una célula vegetal; b) identificar y seleccionar una o más células vegetales del paso (a) que comprende dicha modificación en dicho gen *TFLI* o un homólogo de este; y c) regenerar al menos una planta de la al menos una o más células seleccionadas en el paso (b). En algunas realizaciones, el sitio diana está localizado en una región no codificadora de un gen *TFLI* endógeno o un homólogo de este. En otras realizaciones, la región no codificadora se selecciona del grupo que consiste en un promotor, un intrón, una región 5' no traducida, una región 3' no traducida, y combinaciones de cualquiera de estos. En realizaciones adicionales, la región no codificadora es un promotor. En aún otras realizaciones adicionales, la modificación es facilitada por la presencia de al menos una enzima modificadora del genoma específica del sitio en dicha célula vegetal. Los ejemplos no limitantes de tales enzimas incluyen una nucleasa guiada por ARN, una nucleasa con dedos de cinc, una meganucleasa, una nucleasa TALE (efectora del tipo activador de la transcripción, por sus siglas en inglés), una recombinasa, una transposasa, y combinaciones de cualquiera de estas. Los ejemplos de nucleasas guiadas por ARN incluyen una nucleasa Cas, una nucleasa Cpf1, o una variante de cualquiera de ellas. Algunas enzimas modificadoras del genoma específicas de sitio que podrían usarse de acuerdo con la divulgación crean al menos una rotura de hebra en el sitio diana. Los métodos divulgados en la presente se pueden usar, por ejemplo, para producir cualquier modificación de acuerdo con la divulgación, inclusive una sustitución, una inserción, una inversión, una delección, una duplicación, y una combinación de estas. En algunas realizaciones, la modificación es una delección y la delección comprende una región de al menos 10, al menos 15, al menos 20, al menos 25, al menos 30, al menos 35, al menos 40, al menos 45, al menos 50, al menos 55, al menos 60, al menos 65, al menos 70, al menos 75, al menos 80, al menos 85, al menos 90, al menos 95, al menos 100, al menos 125, o al menos 150 nucleótidos consecutivos.

[010] Los métodos proporcionados en la presente pueden hallar un uso, por ejemplo, en la producción de una planta que tiene un fenotipo deseado. Los ejemplos no limitantes de tal fenotipo

incluyen fecha más temprana de floración terminal, tasa de encamado reducida, sustancialmente el mismo rendimiento o rendimiento aumentado, menor susceptibilidad a la enfermedad fúngica y aumento de la determinación, o cualquier combinación posible de estos. Tal fenotipo se puede definir, en realizaciones específicas, como el que está presente en una planta cuando se la compara con otra por lo demás idéntica, que carece de una modificación que le confiere el fenotipo de acuerdo con la presente divulgación.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

[011] Los dibujos a continuación forman parte de la presente memoria descriptiva y se incluyen para demostrar adicionalmente ciertos aspectos de la presente divulgación. La divulgación se puede comprender mejor con referencia a uno o más de estos dibujos en combinación con la descripción detallada de realizaciones específicas aquí presentada.

[012] La FIG. 1 muestra las posiciones relativas de los ARNg en la región promotora de aproximadamente 2 kb corriente arriba del sitio de inicio de la transcripción (tis) del gen *GmTFL1b*.

[013] La FIG. 2 muestra delecciones homocigotas no redundantes, exclusivas. Los asteriscos (*) indican pequeñas delecciones menores a 20 pares de bases. Obsérvese que en el nombre de los alelos se omitió el prefijo “AL”.

[014] La FIG. 3 muestra los niveles de expresión relativos del gen *GmTFL1b* en la punta del brote y las hojas en plantas que comprenden varios alelos *Dt1* generados por la edición del promotor. La expresión de *GmTFL1b* en el tejido foliar fue nominal para todas las muestras ensayadas.

[015] La FIG. 4 muestra imágenes que ilustran fenotipos de racimos terminales. Las fotos se tomaron dentro de los dos días luego de la fecha de la floración terminal de plantas de tipo silvestre.

[016] La FIG. 5 muestra una comparación de a pares entre la secuencia polipeptídica de TFL1b de soja y las secuencias polipeptídicas de homólogos putativos de especies vegetales pertinentes identificadas mediante una búsqueda BLAST. Los números con coma decimal el triángulo superior derecho de la tabla muestran el porcentaje de residuos idénticos en posiciones de alineación en relación con posiciones de alineación superpuestas entre las dos secuencias. Los números enteros

en el triángulo inferior izquierdo muestran el número de las variaciones de residuos entre las dos secuencias en las posiciones de alineación.

BREVE DESCRIPCIÓN DE LAS SECUENCIAS

[017] SEQ ID NO:1 es la secuencia codificadora de polinucleótidos del gen *Glycine max TFL1b* (*GmTFL1b*).

[018] SEQ ID NO:2 es la secuencia de aminoácidos para la proteína TFL1b (codificada por SEQ ID NO:1).

[019] SEQ ID NO:3 es la secuencia de polinucleótidos para el gen *GmTFL1b*, que incluye las regiones no traducidas 5' y 3' (UTR, por sus siglas en inglés) e intrones.

[020] SEQ ID NO:4 es la secuencia de polinucleótidos de la región promotora de 2 kb corriente arriba del sitio de inicio de la transcripción (tis) del gen *GmTFL1b*. En las plantas de soja de tipo silvestre, SEQ ID NO:4 precede inmediatamente a SEQ ID NO:3.

[021] c21q1SEQ ID NO:5 es la secuencia de polinucleótidos de un andamio común compatible con el gen *Cpf1*.

[022] SEQ ID NO:6 es la secuencia de polinucleótidos para el promotor FLT del virus del mosaico de la Dalia.

[023] SEQ ID NO:7 es una secuencia de polinucleótidos que codifica una enzima endonucleasa guiada por ARN Cpf1 de *Lachnospiraceae bacterium*, con codones optimizados para arroz.

[024] SEQ ID NO:8 es la secuencia de polinucleótidos para una señal de localización nuclear de *Solanum lycopersicum*.

[025] SEQ ID NO:9 es la secuencia de polinucleótidos para un promotor de ubiquitina de *Medicago truncatula*.

[026] SEQ ID NO:10 es una secuencia de polinucleótidos que codifica una enzima endonucleasa guiada por ARN Cpf1 de *Lachnospiraceae bacterium*, con codones optimizados para maíz.

[027] SEQ ID NO:11 es la secuencia de polinucleótidos para un promotor de ARN polimerasa III (Pol) de soja.

[028] SEQ ID NO:12 es la secuencia de polinucleótidos para un promotor de soja 7SL_CR10.

[029] SEQ ID NO:13–29 son secuencias de polinucleótidos para las secuencias espaciadoras en los ARN guía (ARNg) usadas para la edición de la región promotora del gen *GmTFL1b*.

[030] SEQ ID NO:30–46 y 82–128 son secuencias de polinucleótidos para alelos del promotor *GmTFL1b* que tienen varias deleciones (y en el caso de SEQ ID NO:40, también una inversión) en comparación con SEQ ID NO:4.

[031] SEQ ID NO:47 es la secuencia de polinucleótidos de la UTR 5' del gen *GmTFL1b*.

[032] SEQ ID NO:48, 49 y 50 son secuencias de polinucleótidos para el primero, el segundo y el tercer intrón, respectivamente, del gen *GmTFL1b*.

[033] SEQ ID NO:51 es la secuencia de polinucleótidos de la UTR 3' del gen *GmTFL1b*.

[034] SEQ ID NO:52-66, 78 y 80 son las secuencias de aminoácidos de los homólogos de la proteína TFL1b hallados en *Medicago truncatula* (SEQ ID NO:52), *Cajanus cajan* (SEQ ID NO:53), *Pisum sativum* (SEQ ID NO:54), *Arabidopsis thaliana* (SEQ ID NO:55), *Brassica napus* (SEQ ID NO:56), *Gossypium hirsutum* (SEQ ID NO:57), *Capsicum annuum* (SEQ ID NO:58), *Nicotiana tabacum* (SEQ ID NO:59), *Solanum tuberosum* (SEQ ID NO:60), *Solanum lycopersicum* (SEQ ID NO:61), *Zea mays* (SEQ ID NO:62), *Oryza sativa* (SEQ ID NO:63), *Hordeum vulgare* (SEQ ID NO:64), *Triticum aestivum* (SEQ ID NO:65), *Sorghum bicolor* (SEQ ID NO:66), *Arachis hypogaea* (SEQ ID NO:78), y *Cicer arietinum* (SEQ ID NO:80), respectivamente.

[035] SEQ ID NO:67-77, 79 y 81 son las secuencias de polinucleótidos de la región promotora de aproximadamente 2 kb corriente arriba del sitio de inicio de la transcripción (tis) de los homólogos del gen *TFL1b* hallados en *Zea mays* (SEQ ID NO:67), *Sorghum bicolor* (SEQ ID NO:68), *Oryza sativa* (SEQ ID NO:69), *Triticum aestivum* (SEQ ID NO:70), *Hordeum vulgare* (SEQ ID NO:71), *Solanum lycopersicum* (SEQ ID NO:72), *Gossypium hirsutum* (SEQ ID NO:73), *Capsicum annuum* (SEQ ID NO:74), *Brassica napus* (SEQ ID NO:75), *Arabidopsis thaliana* (SEQ ID NO:76), *Medicago truncatula* (SEQ ID NO:77), *Arachis hypogaea* (SEQ ID NO:79), y *Cicer arietinum* (SEQ ID NO:81), respectivamente.

DESCRIPCIÓN DETALLADA

[036] El hábito de crecimiento del tallo es un importante rasgo agronómico que afecta directamente las características de la planta tales como altura de la planta, momento y duración de la floración, producción de nodos, y arquitectura radicular en soja (*Glycine max*). La altura de la planta es un rasgo agronómico especialmente importante en soja y otros cultivos, dado que este puede afectar el potencial de rendimiento y la resistencia al encamado. El momento de la transición desde las fases vegetativas hacia las fases reproductivas en el meristemo apical del tallo influye sobre la altura de la planta. En la soja, el hábito de crecimiento y el momento de la diferenciación terminal de las puntas del tallo, tanto en el ápice del tallo como en las puntas de las ramas, es controlado por el locus *Dt1* (*crecimiento indeterminado 1*). El locus *Dt1* también influye sobre otros rasgos relacionados, tales como la densidad de las ramas, la densidad de las vainas del tallo, el número de nodos del tallo, la cantidad de vainas con tres semillas, y el número total de semillas. La expresión de tipo silvestre de *Dt1* especifica el crecimiento indeterminado y tiene dominancia incompleta sobre el alelo *dt1* (*crecimiento determinado 1*), lo cual causa el crecimiento determinado. La expresión del gen en el locus *Dt1*, *TERMINAL FLOWER 1b* (*GmTFL1b*), en las puntas de los tallos protege el meristemo apical de la diferenciación terminal inducida por *FT2a*, el cual es el ortólogo del *LOCUS FLOWERING T* de la soja.

[037] La mayoría de las variedades de soja cultivadas se clasifican, por ejemplo, como las que tienen uno de dos hábitos de crecimiento del tallo, indeterminado o determinado. Las plantas de soja indeterminadas y determinadas son similares en cuanto a su desarrollo durante la fase de crecimiento vegetativo, pero tienen significativas diferencias en el hábito de crecimiento del tallo en la etapa reproductiva/de floración. Las variedades de soja indeterminadas producen plantas que crecen en altura desde la punta del tallo durante varias semanas, mientras que la floración comienza simultáneamente más abajo en el tallo cuando la planta está aún en la fase de crecimiento vegetativo. Por el contrario, las variedades determinadas producen plantas que completan su crecimiento en altura en el tallo principal al comienzo de las etapas reproductivas (etapa R1) y, aproximadamente al mismo tiempo, producen todas las flores que las plantas producirán. Como resultado, las plantas determinadas tienen tan solo aproximadamente la mitad o aproximadamente dos tercios de la altura de las plantas indeterminadas.

[038] Las variedades de soja indeterminadas generalmente tienen un potencial de rendimiento más elevado en comparación con variedades determinadas, pero tienen la desventaja de ser más proclives al encamado del tallo debido a que tienen mayor peso en la parte superior. No obstante, debido a su papel esencial en el desarrollo, el locus *Dt1* fue fijado tempranamente en la domesticación de la soja, lo cual resultó en una falta general de diversidad en el locus. Esto ha inhibido los esfuerzos para desarrollar variedades con fenotipos de determinación novedosos. Asimismo, la competición entre plantas en el contexto de un cultivo de soja favorece a las plantas más altas con mayor flexibilidad, de modo que inclusive en cribajes mediante mutagénesis, la probabilidad de seleccionar plantas semi-determinadas más compactas, más cortas, con una variación en el grado de determinación, es muy poco probable.

[039] La presente divulgación representa un significativo avance en la técnica en cuanto a que proporciona alelos modificados genéticamente que confieren novedosos fenotipos intermedios entre los actuales hábitos de crecimiento indeterminado y determinado (semi-determinado) en soja y otros cultivos, como así también métodos para su producción, con lo cual se ofrecen mejoras en rasgos fundamentales que llevan al encamado reducido de los cultivos y al aumento de la productividad por planta y por lote. Los métodos y las composiciones divulgados en la presente ofrecen la oportunidad de crear diversidad que no puede seleccionarse a partir del mejoramiento vegetal convencional o de la mutagénesis aleatoria. Por consiguiente, se proporcionan en la presente métodos y composiciones para modificar la determinación en plantas que se pueden usar para lograr beneficios tales como, por ejemplo, el desarrollo de plantas semi-determinadas que ofrecen beneficios exclusivos a los productores, a pesar de los mercados que hasta la fecha consistieron predominantemente solo en variedades determinadas e indeterminadas de cultivos tales como la soja.

[040] Para producir soja y otras plantas que tienen fenotipos novedosos de hábito de crecimiento, la presente divulgación proporciona, en ciertas realizaciones, métodos y composiciones para la creación de alelos novedosos en el locus *Dt1* mediante la edición del promotor del gen *TFL1b*. Por ejemplo, se modificó la región del promotor corriente arriba del gen *TFL1b* como se divulga en la presente por la co-expresión de ocho ARN guía que tienen como blanco aproximadamente 1,2 kb de la secuencia. Se seleccionaron y evaluaron individuos editados que albergaban una serie de

deleciones de 30 hasta 1746 pares de bases. Sorprendentemente se demostró que como resultado se pudieron crear líneas *dtl* editadas que representaban una serie de alelos en el locus *Dtl* con un espectro aparente de fenotipos de determinación, clasificados dentro del rango comprendido entre determinados hasta super-indeterminados. Se generaron alelos que conferían un rango de fechas de floración terminal que se extendía desde 3 semanas antes de las fechas de la variedad indeterminada de fondo hasta casi una semana después, permitiendo por primera vez la capacidad de manipular genéticamente el grado óptimo de semi-determinación. Por lo tanto, la presente divulgación representa un avance significativo en la técnica en cuanto a que permite la producción de novedosos alelos manipulados genéticamente en soja y otros cultivos que confieren novedosos fenotipos de semi-determinación con el potencial de proteger el rendimiento al reducir el encamado y, por lo tanto, aumentar la productividad por planta.

I. Edición genómica

[041] La presente divulgación proporciona, en ciertas realizaciones, plantas, partes de plantas, células vegetales y semillas producidas mediante modificación del genoma usando integración o edición genómica específica del sitio. La edición genómica se puede usar para realizar una o más ediciones o mutaciones en un sitio diana deseado en el genoma de una planta, tal como cambiar la expresión y/o la actividad de uno o más genes, o integrar la inserción de una secuencia o transgén en una locación deseada en el genoma de una planta. Potencialmente se puede seleccionar cualquier sitio o locus en el genoma de una planta para realizar una edición genómica (o edición génica) o la integración dirigida al sitio de un transgén, un constructo o una secuencia de ADN que pueda ser transcripta. Como se usa en la presente, un “sitio diana” para edición genómica o integración dirigida al sitio se refiere a la locación de una secuencia de polinucleótidos dentro del genoma de una planta a la que se dirige una nucleasa específica del sitio y a la cual escinde para introducir una rotura de doble hebra (DSB, por sus siglas en inglés) o el corte de una sola hebra (*nick*) en la estructura del ácido nucleico de la secuencia de polinucleótidos y/o su hebra de ADN complementaria dentro del genoma de la planta. Un sitio diana puede comprender, por ejemplo, al menos 10, al menos 11, al menos 12, al menos 13, al menos 14, al menos 15, al menos 16, al menos 17, al menos 18, al menos 19, al menos 20, al menos 21, al menos 22, al menos 23, al menos 24, al menos 25, al menos 26, al menos 27, al menos 29 o al menos 30 nucleótidos consecutivos. Un

“sitio diana” para una nucleasa guiada por ARN puede comprender la secuencia de cualquier hebra complementaria de una molécula de ácido nucleico (ADN) de doble hebra o cromosoma en el sitio diana. Una nucleasa específica del sitio se puede unir a un sitio diana, tal como mediante un ARN guía no codificador (por ejemplo, entre otros, un ARN CRISPR (ARNcr) o un ARN guía simple (ARNgs) como se describirá a continuación). Un ARN guía no codificador proporcionado en la presente puede ser complementario a un sitio diana (por ejemplo, complementario a cualquier hebra de una molécula de ácido nucleico bicatenaria o cromosoma en el sitio diana). Se apreciará que es posible que no se requiera identidad o complementariedad perfecta para que un ARN guía no codificador se una a un sitio diana o lo hibride. Por ejemplo, se pueden tolerar al menos 1, al menos 2, al menos 3, al menos 4, al menos 5, al menos 6, al menos 7 o al menos 8 errores de apareamiento (o más) entre un sitio diana y un ARN no codificador. Un “sitio diana” también se refiere a la locación de una secuencia de polinucleótidos dentro del genoma de una planta a la que se dirige y la cual es escindida por cualquier otra nucleasa específica del sitio que puede no estar guiada por una molécula de ARN no codificador, tal como una nucleasa con dedos de cinc (ZFN), una nucleasa efectora del tipo activador de la transcripción (TALEN), una meganucleasa, etc., para introducir una rotura de doble hebra (DSB) o un corte en una sola hebra en la secuencia de polinucleótidos y/o su hebra de ADN complementaria. Como se usa en la presente, una “región diana” o una “región de focalización” se refiere a una secuencia de polinucleótidos o una región que está flanqueada por dos o más sitios diana. Sin ser taxativos, en algunas realizaciones, una región diana puede estar sujeta a una mutación, delección, inserción o inversión. Como se usa en la presente, “flanqueada” cuando se usa para describir una región diana de una secuencia de polinucleótidos o una molécula, se refiere a dos o más sitios diana de la secuencia de polinucleótidos o la molécula que circundan la región diana, con un sitio diana a cada lado de la región diana.

[042] Como se usa en la presente, una “técnica de edición genómica focalizada” se refiere a cualquier método, protocolo, o técnica que permita la edición precisa y/o focalizada de una locación específica en el genoma de una planta (es decir, la edición es completamente o en su mayor parte no aleatoria) usando una nucleasa específica del sitio, tal como una meganucleasa, una nucleasa con dedos de cinc (ZFN), una endonucleasa guiada por ARN (por ejemplo, el sistema

CRISPR/Cas9), una endonucleasa TALE (efectora del tipo activador de la transcripción) (TALEN), una recombinasa, o una transposasa. Como se usa en la presente, “edición” o “edición genómica” se refiere a generar una mutación, delección, inversión o sustitución dirigida de al menos 1, al menos 2, al menos 3, al menos 4, al menos 5, al menos 6, al menos 7, al menos 8, al menos 9, al menos 10, al menos 15, al menos 20, al menos 25, al menos 30, al menos 35, al menos 40, al menos 45, al menos 50, al menos 75, al menos 100, al menos 250, al menos 500, al menos 1000, al menos 2500, al menos 5000, al menos 10.000, o al menos 25.000 nucleótidos de una secuencia de ácidos nucleicos del genoma endógeno de una planta. Como se usa en la presente, “edición” o “edición genómica” también puede abarcar la inserción dirigida o la integración dirigida al sitio de al menos 1, al menos 2, al menos 3, al menos 4, al menos 5, al menos 6, al menos 7, al menos 8, al menos 9, al menos 10, al menos 15, al menos 20, al menos 25, al menos 30, al menos 35, al menos 40, al menos 45, al menos 50, al menos 75, al menos 100, al menos 250, al menos 500, al menos 750, al menos 1000, al menos 1500, al menos 2000, al menos 2500, al menos 3000, al menos 4000, al menos 5000, al menos 10.000, o al menos 25.000 nucleótidos en el genoma endógeno de una planta. Una “edición” o “edición genómica” en singular se refiere a tal mutación, delección, inversión, sustitución o inserción dirigida, mientras que “ediciones” o “ediciones genómicas” se refiere a dos o más mutaciones, delecciones, inversiones, sustituciones y/o inserciones dirigidas, donde cada “edición” se introduce mediante una técnica de edición genómica dirigida.

[043] De acuerdo con algunas realizaciones, una nucleasa específica del sitio puede ser co-administrada con una molécula plantilla donadora para servir como una plantilla para realizar una edición, mutación o inserción deseada en el genoma en el sitio diana deseado a través de la reparación de la rotura de doble hebra (/DSB) o corte creado por la nucleasa específica del sitio. De acuerdo con algunas realizaciones, una nucleasa específica del sitio se puede co-administrar con una molécula de ADN que comprende un gen marcador seleccionable o detectable.

[044] Una nucleasa específica del sitio proporcionada en la presente se puede seleccionar del grupo que consiste en una nucleasa con dedos de cinc (ZFN), una endonucleasa TALE (TALEN), una meganucleasa, una endonucleasa guiada por ARN, una recombinasa, una transposasa, o cualquier combinación de estas. Ver, por ejemplo, Khandagale *et al.* (*Plant Biotechnol Rep* 10:327-343,

2016); y Gaj *et al.* (*Trends Biotechnol.* 31(7):397-405, 2013). Las nucleasas con dedos de cinc (ZFN) son proteínas sintéticas que consisten en un dominio de unión a ADN con dedos de cinc modificado genéticamente fusionado con un dominio de escisión (o semi-dominio de escisión), el cual puede derivar de una endonucleasa de restricción (por ejemplo, *FokI*). El dominio de unión a ADN puede ser canónico (C2H2) o no canónico (por ejemplo, C3H o C4). El dominio de unión a ADN puede comprender uno o más dedos de cinc (por ejemplo, 2, 3, 4, 5, 6, 7, 8, 9 o más dedos de cinc) según el sitio diana pero generalmente puede estar compuesto por 3-4 (o más) dedos de cinc. Múltiples dedos de cinc en un dominio de unión a ADN pueden estar separados por una o más secuencias de enlace. Las ZFN pueden estar diseñadas para escindir casi cualquier tramo de ADN de doble hebra mediante la modificación del dominio de unión a ADN de los dedos de cinc. Las ZFN forman dímeros a partir de monómeros compuestos por un dominio de escisión a ADN no específico (por ejemplo, derivado de la nucleasa *FokI*) fusionado con un dominio de unión a ADN que comprende una matriz con dedos de cinc para unirse a una secuencia de ADN en el sitio diana. Los aminoácidos en las posiciones -1, +2, +3 y +6 en relación con el inicio de la hélice α del dedo de cinc, que contribuyen a la unión específica con el sitio diana, se pueden modificar y adecuar para ajustarse a secuencias diana específicas. Los otros aminoácidos pueden formar una estructura de consenso para generar ZFN con especificidades de secuencias diferentes.

[045] Los métodos y las reglas para el diseño de ZFN que se dirijan y se unan a secuencias diana específicas son conocidos en la técnica. Ver, por ejemplo, Publicaciones de Solicitud de Patente de los EE.UU. Nro. 2005/0064474, 2009/0117617 y 2012/0142062. El dominio de nucleasa *FokI* puede requerir dimerización para escindir ADN y, por lo tanto, se necesitan dos ZFN con sus regiones C terminal para unirse a hebras de ADN opuestas al sitio de escisión (separadas por 5-7 pares de bases). El monómero ZFN puede cortar el sitio diana si los dos sitios de unión a ZF son palindrómicos. Una ZFN, como se usa en la presente, se usa en sentido amplio e incluye una ZFN monomérica que puede escindir ADN de doble hebra sin la ayuda de otra ZFN. El término ZFN también se puede usar para hacer referencia a uno o ambos miembros de un par de ZFN que se modifican genéticamente para trabajar juntas en la escisión de ADN en el mismo sitio. Debido a que las especificidades de unión a ADN de los dominios con dedos de cinc se pueden volver a modificar genéticamente usando uno o varios métodos, teóricamente se pueden construir ZFN a

medida para dirigirse a casi cualquier secuencia diana (por ejemplo, en un gen o próxima a este en el genoma de una planta). Los métodos disponibles públicamente para modificar genéticamente dominios con dedos de cinc incluyen el ensamble dependiente del contexto (CoDA, por sus siglas en inglés), la modificación genética con agrupamiento oligomerizado (OPEN) y el ensamble modular.

[046] Los efectores del tipo activador de la transcripción (TALE) se pueden manipular genéticamente para que se unan a prácticamente cualquier secuencia de ADN, tal como en el locus genómico de un gen en una planta o cerca de este. TALE tiene un dominio central de unión a ADN compuesto de 13-28 monómeros repetitivos de 33-34 aminoácidos. Los aminoácidos de cada monómero están altamente conservados, excepto por los residuos aminoacídicos hipervariables en las posiciones 12 y 13. Los dos aminoácidos variables se llaman di-residuos variables repetitivos (RVD). Los pares de aminoácidos NI, NG, HD y NN de RVD preferentemente reconocen adenina, timina, citosina y guanina/adenina, respectivamente, y la modulación de los RVD pueden reconocer bases consecutivas de ADN. Esta simple relación entre una secuencia de aminoácidos y el reconocimiento del ADN ha posibilitado la manipulación genética de dominios específicos de unión a ADN mediante la selección de una combinación de segmentos repetitivos que contienen los RVD apropiados.

[047] Las TALEN son enzimas artificiales de restricción generadas por la fusión del dominio de unión a ADN TALE con un dominio nucleasa. En algunos aspectos, la nucleasa se selecciona de un grupo que consiste en *PvuII*, *MutH*, *TevI*, *FokI*, *AlwI*, *MlyI*, *SbfI*, *SdaI*, *StsI*, *CleDORF*, *Clo051* y *Pept071*. Cuando cada miembro de un par TALEN se une a los sitios en el ADN que flanquean un sitio diana, los monómeros *FokI* se dimerizan y provocan la ruptura del ADN de doble hebra en el sitio diana. El término TALEN, como se usa en la presente, es amplio e incluye una TALEN monomérica que puede escindir ADN de doble hebra sin la ayuda de otra TALEN. El término TALEN también se refiere a uno o ambos miembros de un par de TALEN que trabajan juntos para escindir ADN en el mismo sitio.

[048] Además del dominio de escisión *FokI* de tipo silvestre, se han diseñado variantes del dominio de escisión *FokI* con mutaciones para mejorar la especificidad de la escisión y la actividad de escisión. El dominio *FokI* funciona como un dímero que requiere dos constructos con dominios

de unión a ADN exclusivos para sitios en el genoma diana con orientación y espaciado apropiados. Tanto el número de residuos aminoacídicos entre el dominio TALEN de unión a ADN y el dominio de escisión *FokI* y el número de bases entre los dos sitios de unión TALEN individuales son parámetros para lograr altos niveles de actividad. Los dominios de escisión *PvuII*, *MutH* y *TevI* son alternativas útiles a las variantes *FokI* y *FokI* para usar con los TALE. *PvuII* funciona como un dominio de escisión altamente específico cuando está acoplado a un TALE (ver Yank *et al.*, *PLoS One* 8:e82539, 2013). *MutH* es capaz de introducir cortes específicos de hebra en el ADN (ver Gabsalilow *et al.*, *Nucleic Acids Research*. 41:e83, 2013). *TevI* introduce roturas de doble hebra en el ADN en sitios diana (ver Beurdeley *et al.*, *Nature Communications* 4:1762, 2013).

[049] La relación entre secuencia de aminoácidos y reconocimiento de ADN del dominio de unión TALE posibilita el diseño de proteínas. Se pueden emplear programas de software tales como DNAWorks para diseñar constructos TALE. Los expertos en la técnica conocen otros métodos para diseñar constructos TALE. Ver Doyle *et al.* (*Nucleic Acids Research* 40:W117-122, 2012); Cermak *et al.* (*Nucleic Acids Research* 39:e82, 2011); y tale-nt.cac.cornell.edu/about. En otro aspecto, una TALEN proporcionada en la presente es capaz de generar una DSB selectiva.

[050] Una nucleasa específica del sitio puede ser una meganucleasa. Las meganucleasas, que se identifican comúnmente en los microbios, tales como la familia LAGLIDADG de endonucleasas direccionadas, son enzimas exclusivas con elevada actividad y secuencias largas de reconocimiento (> 14 bp) que resultan en la digestión específica del sitio de un ADN diana. Las versiones genéticamente manipuladas de meganucleasas de origen natural generalmente tienen secuencias de reconocimiento de ADN extendidas (por ejemplo, 14 a 40 pares de bases). La manipulación genética de meganucleasas puede ser un desafío mayor que el de ZFN y TALEN debido a que las funciones de reconocimiento y escisión del ADN están entrelazadas en un solo dominio. Se han utilizado métodos de mutagénesis y cribaje de alto rendimiento para crear novedosas variantes de meganucleasas que reconocen secuencias exclusivas y poseen actividad nucleasa mejorada.

[051] Una nucleasa específica de sitio puede ser una nucleasa guiada por ARN. De acuerdo con algunas realizaciones, una endonucleasa guiada por ARN se puede seleccionar del grupo que consiste en Cas1, Cas1B, Cas2, 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, Cpf1, CasX, CasY, y los homólogos o versiones modificadas de cualquiera de ellas, como así también Argonauta (ejemplos no limitantes de proteínas Argonauta incluyen Argonauta *Thermus thermophilus* (TtAgo), Argonauta *Pyrococcus furiosus* (PfAgo), Argonauta *Natronobacterium gregoryi* (NgAgo), y homólogos o versiones modificadas de cualquiera de ellas). De acuerdo con algunas realizaciones, una endonucleasa guiada por ARN es una enzima Cas9 o Cpf1. La nucleasa guiada por ARN se puede administrar como una proteína con o sin un ARN guía, o el ARN guía puede formar un complejo con la enzima nucleasa guiada por ARN y administrarse como una ribonucleoproteína (RNP).

[052] En el caso de endonucleasas guiadas por ARN, se puede proporcionar además una molécula de ARN guía para dirigir la endonucleasa hacia un sitio diana en el genoma de la planta mediante apareamiento de bases o hibridación para provocar una DSB o un corte en el sitio diana o cerca de este. El ARN guía se puede transformar o introducir en una célula o tejido vegetal como una molécula de ARNg, o como una molécula, un constructo o vector de ADN recombinante que comprende una secuencia de ADN que se puede transcribir que codifica al ARN guía enlazado operativamente con un promotor. Como se entiende en la técnica, un ARN guía puede comprender, por ejemplo, un ARN CRISPR (ARNcr), un ARN guía monocatenario (ARNgs), o cualquier otra molécula de ARN que pueda guiar o dirigir una endonucleasa hacia un sitio diana específico en el genoma. Una proteína prototípica asociada a CRISPR, Cas9 de *S. pyogenes*, se une naturalmente a dos ARN, un ARN guía CRISPR (ARNcr) y un ARN CRISPR transactivante (ARNcr transactivante), para ensamblar una ribonucleoproteína CRISPR (RNPcr). Un “ARN guía monocatenario” (o “ARNgs”) es una molécula de ARN que comprende un ARNcr covalentemente enlazado con un ARNcr transactivante mediante una secuencia enlazadora, la cual puede estar expresada como una sola molécula o transcrito de ARN. El ARN guía comprende una secuencia de localización o guía (también denominada en la presente como “secuencia espaciadora”) que es idéntica o complementaria con un sitio diana dentro del genoma de la planta, tal como en un gen o próximo a este. El ARN guía es generalmente una molécula de ARN no codificante que no

codifica una proteína. La secuencia guía del ARN guía puede tener una longitud de al menos 10 nucleótidos, tal como 12-40 nucleótidos, 12-30 nucleótidos, 12-20 nucleótidos, 12-35 nucleótidos, 12-30 nucleótidos *[sic]*, 15-30 nucleótidos, 17-30 nucleótidos o 17-25 nucleótidos de longitud, o aproximadamente 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 o más nucleótidos de longitud. La secuencia guía puede ser al menos 95%, al menos 96%, al menos 97%, al menos 99% o 100% idéntica o complementaria con al menos 10, al menos 11, al menos 12, al menos 13, al menos 14, al menos 15, al menos 16, al menos 17, al menos 18, al menos 19, al menos 20, al menos 21, al menos 22, al menos 23, al menos 24, al menos 25 o más nucleótidos consecutivos de una secuencia de ADN en el sitio diana genómico.

[053] Como se mencionó anteriormente, un gen diana para edición genómica puede ser cualquiera de los genes del tipo *TERMINAL FLOWER 1* descritos en la presente para supresión, inclusive el gen *TERMINAL FLOWER 1b* (*GmTFL1b*) de la soja. En el caso de mutaciones para la reducción de la expresión (*knock-down*) del gen *TFL1* mediante edición genómica, se puede localizar una endonucleasa guiada por ARN hacia una secuencia corriente arriba o corriente abajo, tal como un promotor y/o una secuencia potenciadora, o un intrón, una secuencia 5'UTR y/o 3'UTR del gen *TFL1* para mutar uno o más promotores y/o secuencias potenciadoras del gen *TFL1* para afectar o reducir su nivel de expresión. Para reducir la expresión del gen *GmTFL1b* en soja se puede emplear un ARN guía, el cual comprende una secuencia guía que es al menos 90%, al menos 95%, al menos 96%, al menos 97%, al menos 99% o 100% idéntica o complementaria con al menos 10, al menos 11, al menos 12, al menos 13, al menos 14, al menos 15, al menos 16, al menos 17, al menos 18, al menos 19, al menos 20, al menos 21, al menos 22, al menos 23, al menos 24, al menos 25 o más nucleótidos consecutivos de las SEQ ID NO: 3 o 4 o una secuencia complementaria de estas (por ejemplo, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 o más nucleótidos consecutivos de las SEQ ID NO: 3 o 4 o una secuencia complementaria de estas), aunque se pueden producir empalmes alternativos y diferentes límites exón/intrón. Como se usa en la presente, el término “consecutivo” con referencia a una secuencia de polinucleótidos o proteínas significa sin deleciones ni huecos en la secuencia.

[054] Como se usa en la presente, con respecto a una secuencia dada, un “complemento”, una “secuencia complementaria” y un “complemento inverso” se usan de modo intercambiable. Las

tres expresiones se refieren a secuencias inversamente complementarias de una secuencia de nucleótidos, es decir, a una secuencia complementaria a una secuencia dada con los nucleótidos en orden invertido.

[055] Como se usa en la presente, el término “antisentido” se refiere a secuencias de ADN o ARN que son complementarias a una secuencia de ADN o ARN específica. Las moléculas de ARN antisentido son ácidos nucleicos monocatenarios que se pueden combinar con una hebra o secuencia sentido de ARN o ARNm para formar dúplex debido a la complementariedad de las secuencias. La expresión “hebra antisentido” se refiere a una hebra de ácido nucleico que es complementaria a la hebra “sentido”. La “hebra sentido” de un gen o locus es la hebra de ADN o ARN que tiene la misma secuencia que una molécula de ARN transcripta del gen o locus (con excepción de uracilo en ARN y timina en ADN).

[056] Un motivo adyacente protoespaciador (PAM) puede estar presente en el genoma inmediatamente adyacente y corriente arriba hacia el extremo 5' de la secuencia del sitio genómico diana complementaria a la secuencia de localización del ARN guía - es decir, inmediatamente corriente abajo (3') hacia la hebra sentido (+) del sitio genómico diana (en relación con la secuencia de localización del ARN guía) como se conoce en la técnica. Ver, por ejemplo, Wu *et al.* (*Quant Biol.* 2(2):59-70, 2014). La secuencia genómica PAM en la hebra sentido (+) adyacente al sitio diana (relativo a la secuencia de localización del ARN guía) puede comprender 5'-NGG-3. Sin embargo, la secuencia correspondiente del ARN guía (es decir, inmediatamente corriente abajo (3') a la secuencia de localización del ARN guía), en general, puede no ser complementaria de la secuencia PAM genómica.

[057] En algunas realizaciones, una nucleasa específica del sitio es una recombinasa. Los ejemplos no taxativos de recombinasas que se pueden emplear incluyen una serina recombinasa unida a un motivo de reconocimiento de ADN, una tirosina recombinasa unida a un motivo de reconocimiento de ADN, o cualquier enzima recombinasa conocida en la técnica unida a un motivo de reconocimiento de ADN. En ciertas realizaciones, la nucleasa específica de sitio es una recombinasa o transposasa, la cual puede ser una transposasa o recombinasa de ADN unida o fusionada con un dominio de unión a ADN. Los ejemplos no taxativos de recombinasas incluyen una tirosina recombinasa seleccionada del grupo que consiste en una recombinasa Cre, una

recombinasa *Gin*, una recombinasa *Flp*, y una recombinasa *Tnp1* unida a un motivo de reconocimiento de ADN proporcionado en la presente. En un aspecto de la presente divulgación, una recombinasa *Cre* o una recombinasa *Gin* proporcionada en la presente está ligada a un dominio de unión al ADN con dedos de cinc, un dominio de unión al ADN TALE, o una nucleasa Cas9. En otro aspecto, una serina recombinasa seleccionada del grupo que consiste en una integrasa PhiC31, una integrasa R4 y una integrasa TP-901 puede estar unida a un motivo de reconocimiento de ADN proporcionado en la presente. En inclusive otro aspecto, una ADN transposasa seleccionada del grupo que consiste en un TALE-piggyBac y un TALE-mutador se pueden unir a un dominio de unión al ADN proporcionado en la presente.

[058] Diversas nucleasas específicas del sitio, tales como recombinasas, nucleasas con dedos de cinc (ZFN), meganucleasas y TALEN no están guiadas por ARN y, en cambio, se basan en la estructura de sus proteínas para determinar su sitio diana para provocar la DSB o el corte, o se fusionan, se ligan o unen a un dominio o motivo de proteína de unión al ADN. La estructura de las proteínas de la nucleasa específica del sitio (o el dominio de unión al ADN fusionado/unido/ligado) puede dirigir la nucleasa específica del sitio hacia el sitio diana. De acuerdo con muchas de estas realizaciones, nucleasas específicas del sitio no guiadas por ARN, tales como recombinasas, nucleasas con dedos de cinc (ZFN), meganucleasas y TALEN, se pueden diseñar, modificar genéticamente y construir de acuerdo con métodos conocidos para localizar y unirse con un sitio diana en el locus genómico o cerca de este de un gen endógeno de una planta para crear una DSB o un corte en dicho locus genómico. La DSB o el corte creado mediante la nucleasa específica del sitio no guiada por ARN puede llevar a reducir la expresión génica mediante la reparación de la DSB o el corte, lo cual puede resultar en una mutación o inserción de una secuencia en el sitio de la DSB o el corte a través de mecanismos de reparación celular. Tales mecanismos de reparación celular pueden estar guiados por una molécula molde donadora.

[059] Como se usa en la presente, una “molécula donadora”, “plantilla donadora”, o “molécula de plantilla donadora” (colectivamente una “plantilla donadora”), la cual puede ser un polinucleótido recombinante, una plantilla o secuencia donadora de ADN o ARN, se define como una molécula de ácido nucleico que tiene una plantilla o secuencia de ácido nucleico homóloga (por ejemplo, secuencia de homología) y/o una secuencia de inserción para inserción o recombinación dirigida

al sitio o de localización en el genoma de una célula vegetal mediante la reparación de un corte o una DSB en el genoma de una célula vegetal. Una plantilla donadora puede ser una molécula separada de ADN que comprende una o más secuencias homólogas y/o una secuencia de inserción para integración localizada, o una plantilla donadora puede ser una porción de una secuencia (es decir, una región de plantilla donadora) de una molécula de ADN que además comprende uno o más casetes de expresión, genes/transgenes, y/o secuencias de ADN que se pueden transcribir. Por ejemplo, una “plantilla donadora” se puede usar para la integración dirigida al sitio de un transgén o constructo, o como una plantilla para introducir una mutación, tal como una inserción, delección, sustitución, etc., en un sitio diana dentro del genoma de una planta. Una técnica de edición genómica dirigida proporcionada en la presente puede comprender el uso de una o más, dos o más, tres o más, cuatro o más, o cinco o más moléculas o plantillas donadoras. Una plantilla donadora proporcionada en la presente puede comprender al menos uno, al menos dos, al menos tres, al menos cuatro, al menos cinco, al menos seis, al menos siete, al menos ocho, al menos nueve, o al menos diez genes o transgenes y/o secuencias de ADN que se pueden transcribir. Alternativamente, una plantilla donadora puede no comprender ningún gen, transgén ni secuencias de ADN que se pueden transcribir.

[060] Sin ser taxativos, un gen/transgén o secuencia de ADN que se puede transcribir de una plantilla donadora puede incluir, por ejemplo, un gen de resistencia a insecticidas, un gen de tolerancia a herbicidas, un gen para uso eficiente de nitrógeno, un gen para uso eficiente de agua, un gen potenciador del rendimiento, un gen para la calidad nutricional, un gen de unión a ADN, un gen marcador seleccionable, un ARNi o un constructo de supresión, un gen de la enzima para modificación genómica específica del sitio, un ARN guía simple de un sistema CRISPR/Cas9, un casete de expresión basado en geminivirus. De acuerdo con otras realizaciones, una secuencia de inserción de una plantilla donadora puede comprender una secuencia codificadora de proteínas o una secuencia de ADN que se puede transcribir que codifica una molécula de ARN no codificadora, la cual puede estar dirigida a un gen endógeno para suprimirlo. Una plantilla donadora puede comprender un promotor operativamente enlazado con una secuencia codificadora, un gen o una secuencia de ADN que se puede transcribir, tal como un promotor constitutivo, un promotor que funciona específicamente en un tejido o preferentemente en un

tejido, un promotor de etapas de desarrollo, o un promotor inducible. Una plantilla donadora puede comprender un líder, potenciador, promotor, sitio de inicio de la transcripción, 5'-UTR, uno o más exones, uno o más intrones, un sitio de terminación de la transcripción, una región o secuencia, 3'-UTR, y/o una señal de poliadenilación, que pueden estar operativamente enlazados con una secuencia codificadora, un gen (o transgén) o una secuencia de ADN que se puede transcribir que codifican un ARN no codificador, un ARN guía, un ARNm y/o proteína. Una plantilla donadora puede ser una molécula o plásmido de ADN o ARN monocatenario o bicatenario.

[061] Una “secuencia de inserción” de una plantilla donadora es una secuencia diseñada para la inserción localizada en el genoma de una célula vegetal, la cual puede tener cualquier longitud adecuada. Por ejemplo, la secuencia de inserción de una plantilla donadora puede tener una longitud de entre 2 y 50.000, entre 2 y 10.000, entre 2 y 5000, entre 2 y 1000, entre 2 y 500, entre 2 y 250, entre 2 y 100, entre 2 y 50, entre 2 y 30, entre 15 y 50, entre 15 y 100, entre 15 y 500, entre 15 y 1000, entre 15 y 5000, entre 18 y 30, entre 18 y 26, entre 20 y 26, entre 20 y 50, entre 20 y 100, entre 20 y 250, entre 20 y 500, entre 20 y 1000, entre 20 y 5000, entre 20 y 10.000, entre 50 y 250, entre 50 y 500, entre 50 y 1000, entre 50 y 5000, entre 50 y 10.000, entre 100 y 250, entre 100 y 500, entre 100 y 1000, entre 100 y 5000, entre 100 y 10.000, entre 250 y 500, entre 250 y 1000, entre 250 y 5000, o entre 250 y 10.000 nucleótidos o pares de bases. Una plantilla donadora también puede tener al menos una secuencia de homología o brazo de homología, tal como dos brazos de homología, para dirigir la integración de una secuencia de inserción o mutación en un sitio diana dentro del genoma de una planta mediante recombinación homóloga, donde la secuencia de homología o los brazos de homología son idénticos o complementarios, o tienen un porcentaje de identidad o un porcentaje de complementariedad, con una secuencia en el sitio diana, o cerca de este, dentro del genoma de la planta. Cuando una plantilla donadora comprende uno o más brazos de homología y una secuencia de inserción, el o los brazos de homología flanquearán o circundarán la secuencia de inserción de la plantilla donadora. Cada brazo de homología puede ser al menos 70%, al menos 75%, al menos 80%, al menos 85%, al menos 90%, al menos 95%, al menos 96%, al menos 97%, al menos 99% o 100% idéntico o complementario con al menos 20, al menos 25, al menos 30, al menos 35, al menos 40, al menos 45, al menos 50, al menos 60, al menos 70, al menos 80, al menos 90, al menos 100, al menos 150,

al menos 200, al menos 250, al menos 500, al menos 1000, al menos 2500, o al menos 5000 nucleótidos consecutivos de una secuencia de ADN diana dentro del genoma de una planta.

[062] Con la presente divulgación se puede emplear cualquier método conocido en la técnica para integración dirigida al sitio. En presencia de una molécula de plantilla donadora con una secuencia de inserción, se puede reparar la DSB o el corte mediante recombinación homóloga entre el o los brazos de homología de la plantilla donadora y el genoma de la planta, o mediante la unión de extremos no homólogos (NHEJ, por sus siglas en inglés), lo cual resulta en la integración dirigida al sitio de la secuencia de inserción en el genoma de la planta para crear un evento de inserción localizada en el sitio de la DSB o el corte. Por consiguiente, se puede lograr la inserción o integración específica del sitio de un transgén, una secuencia de ADN que se puede transcribir, un constructo o secuencia si el transgén, la secuencia de ADN que se puede transcribir, el constructo o la secuencia está localizado en la secuencia de inserción de la plantilla donadora

[063] Para suprimir un gen *TFLI* de acuerdo con realizaciones de la presente divulgación se puede emplear cualquier método conocido en la técnica para la supresión de un gen diana, inclusive expresión de ARN antisentido, ARN bicatenarios (ARNbc) o secuencias de ARN invertidas repetidas, o mediante co-supresión o ARN de interferencia (ARNi) a través de la expresión de ARN pequeños de interferencia (ARNpi), ARN de horquilla corta (ARhc), ARNpi transactivante (ARNta), o micro ARN (miARN). Asimismo, se pueden emplear moléculas de ARN sentido y/o antisentido que se dirijan a las secuencias o regiones genómicas codificadoras y/o no codificadoras dentro del gen *TFLI* o cerca de este para provocar el silenciamiento del gen. Por consiguiente, se puede usar cualquiera de estos métodos para la supresión localizada de un gen endógeno *TFLI* de una manera específica en un tejido o preferida en un tejido. Ver, por ejemplo, las Publicaciones de Solicitud de Patente de los EE.UU. Nro. 2009/0070898, 2011/0296555 y 2011/0035839.

[064] También se puede emplear la introducción de una DSB o corte para introducir mutaciones localizadas en el genoma de una planta. De acuerdo con este enfoque, se pueden introducir mutaciones, tales como deleciones, inserciones, inversiones y/o sustituciones en un sitio diana mediante la reparación imperfecta de la DSB o el corte para producir el bloqueo o la reducción de la expresión de un gen. Tales mutaciones se pueden generar mediante la reparación imperfecta del locus localizado inclusive sin usar una molécula de plantilla donadora. El bloqueo de la expresión

de un gen (“*knock-out*”) se puede lograr induciendo una DSB o un corte en el locus endógeno del gen o cerca de este que resulte en la falta de expresión de la proteína o en la expresión de una proteína no funcional, mientras que la reducción de la expresión de un gen (“*knock-down*”) se puede lograr de una manera similar induciendo una DSB o un corte en el locus endógeno del gen o cerca de este que es reparado de manera imperfecta en un sitio que no afecta la secuencia codificadora del gen en una manera que eliminaría la función de la proteína codificadora. Por ejemplo, el sitio de la DSB o el corte dentro del locus endógeno puede estar en la región corriente arriba o 5’ del gen (por ejemplo, un promotor y/o una secuencia potenciadora) para afectar o reducir su nivel de expresión.

[065] De modo similar, tales mutaciones dirigidas de bloqueo o reducción de expresión de un gen se pueden generar con una molécula de plantilla donadora para dirigir una mutación particular o deseada en el sitio diana o cerca de este mediante la reparación de la DSB o el corte. La molécula de la plantilla donadora puede comprender una secuencia homóloga con o sin una secuencia de inserción y comprender una o más mutaciones, tal como una o más delecciones, inserciones, inversiones y/o sustituciones, en relación con la secuencia genómica localizada en el sitio de la DSB o el corte, o cerca de este. Por ejemplo, las mutaciones dirigidas de bloqueo o reducción de expresión de un gen se pueden lograr con la sustitución, inserción, delección o inversión de al menos una porción del gen, tal como con la introducción de un desplazamiento de marco o un codón de parada prematuro en la secuencia codificadora del gen o con la interrupción de una secuencia promotora o la secuencia de otro elemento regulador no codificador del gen. También se puede introducir la delección de una porción de un gen mediante la generación de DSB o cortes en dos sitios diana, provocando así una delección de la región diana intercalada flanqueada por los sitios diana.

[066] En un aspecto, la presente divulgación proporciona una planta de soja modificada, una parte de planta de esta, que comprende un alelo mutante del gen *TFLI*, donde el alelo mutante comprende al menos una modificación genómica que involucra al menos 10, al menos 20, al menos 30, al menos 40, al menos 50, al menos 60, al menos 70, al menos 75, al menos 80, al menos 85, al menos 90, al menos 95, al menos 100, al menos 110, al menos 125, al menos 150, al menos 175 o al menos 200 nucleótidos consecutivos de la región promotora del gen endógeno *TFLI*. La

secuencia promotora del gen *TFL1b* de soja comprende la secuencia SEQ ID NO:4, la cual es una secuencia de polinucleótidos de 2 kb corriente arriba del sitio de inicio de la transcripción en el gen *TFL1b*. La modificación genómica puede ser la delección de una región que comprende al menos 10, al menos 20, al menos 30, al menos 40, al menos 50, al menos 60, al menos 70, al menos 75, al menos 80, al menos 85, al menos 90, al menos 95, al menos 100, al menos 110, al menos 125, al menos 150, al menos 175 o al menos 200 nucleótidos consecutivos de la secuencia de SEQ ID NO:4. Tal delección en SEQ ID NO:4 puede incluir una región que abarca: desde el nucleótido 1539 hasta el nucleótido 1568; desde el nucleótido 1217 hasta el nucleótido 1604; desde el nucleótido 1518 hasta el nucleótido 1629; desde el nucleótido 951 hasta el nucleótido 1222; desde el nucleótido 1364 hasta el nucleótido 1407; desde el nucleótido 1367 hasta el nucleótido 1516; desde el nucleótido 754 hasta el nucleótido 1806; desde el nucleótido 1216 hasta el nucleótido 1319; desde el nucleótido 1159 hasta el nucleótido 1965; desde el nucleótido 760 hasta el nucleótido 1214; desde el nucleótido 939 hasta el nucleótido 1028; desde el nucleótido 1066 hasta el nucleótido 1664; desde el nucleótido 952 hasta el nucleótido 1552; desde el nucleótido 1677 hasta el nucleótido 1808; desde el nucleótido 1524 hasta el nucleótido 1558; desde el nucleótido 476 hasta el nucleótido 1405; desde el nucleótido 1365 hasta el nucleótido 1559; desde el nucleótido 928 hasta el nucleótido 1560; o desde el nucleótido 593 hasta el nucleótido 1813 de SEQ ID NO:4.

[067] En un aspecto, la modificación genómica puede ser la inversión de una región de al menos 10, al menos 20, al menos 30, al menos 40, al menos 50, al menos 60, al menos 70, al menos 75, al menos 80, al menos 85, al menos 90, al menos 95, al menos 100, al menos 110, al menos 125, al menos 150, al menos 175 o al menos 200 nucleótidos consecutivos dentro de la secuencia de SEQ ID NO:4. Tal inversión puede comprender una región desde el nucleótido 1029 hasta el nucleótido 1065 de SEQ ID NO:4 que ha sido removida, invertida y reinsertada en la misma localización en SEQ ID NO:4. En un aspecto, un alelo mutante del gen *TFL1* puede comprender dos o más modificaciones en la región promotora del gen endógeno *TFL1*. En la presente se divulgan ejemplos de tales alelos mutantes del gen *TFL1b* de soja e incluyen, por ejemplo, un alelo que comprende dos delecciones en la secuencia SEQ ID NO:4; donde la primera delección abarca una región desde el nucleótido 952 hasta el nucleótido 1552 de SEQ ID NO:4 y la segunda delección

abarca desde el nucleótido 1677 hasta el nucleótido 1808 de SEQ ID NO:4; y un alelo que comprende dos deleciones y una inversión dentro de la secuencia de SEQ ID NO:4, donde la primera deleción abarca una región desde el nucleótido 939 hasta el nucleótido 1028 de SEQ ID NO:4 y la segunda deleción abarca desde el nucleótido 1066 hasta el nucleótido 1664 de SEQ ID NO:4, y la inversión comprende una región que abarca desde el nucleótido 1029 hasta el nucleótido 1065 de SEQ ID NO:4 que se ha eliminado, invertido y reinsertado en la misma locación en SEQ ID NO:4.

[068] Se pueden realizar otras modificaciones localizadas en la región promotora para generar alelos novedosos en el gen *TFL1b* de soja y sus homólogos. SEQ ID NO:67-77, 79 y 81 respectivamente representan una secuencia de polinucleótidos de aproximadamente 2 kb corriente arriba del sitio de inicio de la transcripción en los homólogos del gen *TFL1b* hallados en *Zea mays* (SEQ ID NO:67), *Sorghum bicolor* (SEQ ID NO:68), *Oryza sativa* (SEQ ID NO:69), *Triticum aestivum* (SEQ ID NO:70), *Hordeum vulgare* (SEQ ID NO:71), *Solanum lycopersicum* (SEQ ID NO:72), *Gossypium hirsutum* (SEQ ID NO:73), *Capsicum annuum* (SEQ ID NO:74), *Brassica napus* (SEQ ID NO:75), *Arabidopsis thaliana* (SEQ ID NO:76), *Medicago truncatula* (SEQ ID NO:77), *Arachis hypogaea* (SEQ ID NO:79), y *Cicer arietinum* (SEQ ID NO:81). Por ejemplo, se pueden localizar uno o más sitios de modificación a aproximadamente 200 nucleótidos o a más de 200 nucleótidos del extremo 3' de las secuencias de SEQ ID NO:4, 67-77, 79 y 81. En un aspecto, se pueden realizar una o más modificaciones dentro de la región de ADN que abarca desde la posición nucleotídica 1237 hasta la posición nucleotídica 1570 de SEQ ID NO:4 para generar un alelo novedoso en el gen *TFL1b* de soja.

[069] En un aspecto adicional, la presente divulgación proporciona una planta de soja modificada, o una parte de planta de esta, que comprende un alelo mutante del gen *TFL1*, donde el alelo mutante comprende una o más secuencias de unión, donde las secuencias de unión son al menos 30, al menos 60, al menos 100 nucleótidos en el sitio de unión. Como se usa en la presente, una “unión” o “sitio de unión” es el punto de conexión entre la secuencia de nucleótidos en el sitio de una deleción, inserción, sustitución o inversión. En el caso de una deleción, la unión es el punto de conexión en el sitio de la deleción de las secuencias que previamente flanqueaban la deleción. Por ejemplo, en el caso de la deleción de 30 pares de bases desde el nucleótido 1539 hasta el

nucleótido 1568, en comparación con la secuencia de referencia SEQ ID NO:4 descrita en la presente, la unión sería entre el nucleótido 1538 y el nucleótido 1569. En el caso de una inserción, sustitución o inversión, la unión es el punto de conexión entre la secuencia insertada, invertida o sustituida y las secuencias de ADN flanqueantes. En el caso de una inserción, sustitución o inversión, una unión se halla en el extremo 5' de la inserción, sustitución o inversión, y otra unión se halla en el extremo 3' de la inserción, sustitución o inversión. Una “secuencia de unión” se refiere a una secuencia de ADN de cualquier longitud que abarca una unión. Una secuencia de unión puede comprender al menos 10 nucleótidos, al menos 15 nucleótidos, al menos 20 nucleótidos, al menos 25 nucleótidos, al menos 30 nucleótidos, al menos 40 nucleótidos, al menos 50 nucleótidos, al menos 60 nucleótidos, al menos 70 nucleótidos, al menos 80 nucleótidos, al menos 90 nucleótidos, al menos 100 nucleótidos, al menos 200 nucleótidos, al menos 300 nucleótidos o más.

II. Constructos para edición genómica

[070] Se proporcionan constructos y vectores de ADN recombinante que comprenden una secuencia de polinucleótidos que codifica una nucleasa específica del sitio, tal como una nucleasa con dedos de cinc (ZFN), una meganucleasa, una endonucleasa guiada por ARN, una endonucleasa TALE (TALEN), una recombinasa, o una transposasa, donde la secuencia codificadora está operativamente enlazada con un promotor que se puede expresar en una planta. En el caso de endonucleasas guiadas por ARN, además se proporcionan constructos y vectores de ADN recombinante que comprenden una secuencia de polinucleótidos que codifica una ARN guía, donde el ARN guía comprende una secuencia guía con la longitud suficiente que tiene un porcentaje de identidad o complementariedad con un sitio diana dentro del genoma de una planta, tal como un gen localizado *TFL1* o cerca de este. Una secuencia de polinucleótidos de un constructo y vector de ADN recombinante que codifica una nucleasa específica del sitio o un ARN guía puede estar operativamente enlazada con un promotor que se puede expresar en una planta, tal como un promotor inducible, un promotor constitutivo, un promotor que funciona específicamente en un tejido, etc.

[071] Como se usa en la presente, un “gen” se refiere a una secuencia de ácidos nucleicos que forma una unidad genética y funcional y codifica para una o más moléculas polipeptídicas y/o de

ARN relacionadas con la secuencia. Un gen generalmente contiene una región codificadora operativamente enlazada con las secuencias reguladoras apropiadas que regulan la expresión de un producto génico (por ejemplo, un polipéptido o un ARN funcional). Un gen puede tener varios elementos en la secuencia, inclusive, entre otros, un promotor, una región no traducida (UTR), exones, intrones, y otras secuencias reguladoras corriente arriba o corriente abajo.

[072] Como se usa en la presente, un “alelo” se refiere a una secuencia de ácidos nucleicos alternativa de un gen o en un locus en particular (por ejemplo, una secuencia de ácidos nucleicos de un gen o locus que es diferente a otros alelos para el mismo gen o locus). Dicho alelo se puede considerar i) de tipo silvestre o ii) mutante si una o más mutaciones o ediciones están presentes en la secuencia de ácidos nucleicos del alelo mutante en relación con el alelo de tipo silvestre. Un alelo mutante o editado para un gen puede tener una actividad o nivel de expresión reducido o eliminado para el gen relativo al alelo de tipo silvestre. Por ejemplo, un alelo mutante o editado para el gen *TFL1* puede tener una delección en la región promotora corriente arriba del gen endógeno *TFL1*. En el caso de los organismos diploides tales como la soja, un alelo puede darse en un cromosoma, y un segundo alelo puede darse en el mismo locus en un segundo cromosoma homólogo. Si un alelo en un locus en un cromosoma de una planta es un alelo mutante o editado y el otro alelo correspondiente en el cromosoma homólogo de la planta es de tipo silvestre, entonces se dice que la planta es heterocigota para el alelo mutante o editado. No obstante, si ambos alelos en un locus son alelos mutantes o editados, entonces se dice que la planta es homocigota para los alelos mutantes o editados. Una planta homocigota para alelos mutantes o editados en un locus puede comprender el mismo alelo mutante o editado o alelos mutantes o editados diferentes si es heteroalélica o bialélica.

[073] Como se usa en la presente, un “gen de tipo silvestre” o “alelo de tipo silvestre” se refiere a un gen o alelo que tiene una secuencia o genotipo que es el más común en una especie particular de planta, u otra secuencia o genotipo que tiene solo variaciones, polimorfismos, u otras mutaciones silenciosas naturales en relación con la secuencia o genotipo del tipo más común que no tiene impacto significativo sobre la expresión y la actividad del gen o el alelo. Por cierto, un gen o alelo de “tipo silvestre” no contiene variaciones, polimorfismos ni otro tipo de mutación que

afecte de manera sustancial la función, actividad, expresión o consecuencia fenotípica normal del gen o el alelo en relación con la secuencia o el genotipo más común.

[074] En general, el término “variante” se refiere a moléculas con algunas diferencias, generadas sintética o naturalmente, en sus secuencias de aminoácidos o de nucleótidos en comparación con polinucleótidos o polipéptidos de referencia (nativos), respectivamente. Estas diferencias incluyen sustituciones, inserciones, deleciones o cualquier combinación deseada de tales cambios en una secuencia de polinucleótidos o aminoácidos nativa.

[075] Como se usa en la presente, el término “expresión” se refiere a la biosíntesis de un producto génico, y típicamente la transcripción y/o traducción de una secuencia de nucleótidos, tal como un gen endógeno, un gen heterólogo, un transgén o una secuencia codificadora de proteínas y/o ARN, en una célula, un tejido, un órgano u organismo, tal como una planta, parte de una planta o célula vegetal, tejido u órgano.

[076] El término “recombinante” con referencia a una molécula de polinucleótido (ADN o ARN), proteína, constructo, vector, etc., se refiere a una molécula o secuencia de proteínas o polinucleótidos hecha por el hombre y que no se encuentra normalmente en la naturaleza, y/o está presente en un contexto en el cual esta no se halla normalmente en la naturaleza, inclusive una molécula de polinucleótido (ADN o ARN), proteína, constructo, etc., que comprende una combinación de dos o más secuencias de proteínas o polinucleótidos que no se darían naturalmente juntas de la misma manera sin la intervención humana, tal como una molécula de polinucleótido, proteína, constructo, etc., que comprende al menos dos secuencias de proteínas o polinucleótidos que están operativamente enlazadas pero que son heterólogas entre sí. Por ejemplo, el término “recombinante” se puede referir a cualquier combinación de dos o más secuencias de ADN o proteínas en la misma molécula (por ejemplo, un plásmido, constructo, vector, cromosoma, proteína, etc.) donde tal combinación está hecha por el hombre y no se halla normalmente en la naturaleza. Como se usa en esta definición, la frase “no se halla normalmente en la naturaleza” significa que no se halla en la naturaleza sin la introducción del ser humano. Una molécula de proteína o polinucleótido, constructo, etc. recombinante puede comprender una o más secuencias de polinucleótidos o proteínas que están i) separadas de otras secuencias de proteínas o polinucleótidos que existen próximas entre sí en la naturaleza, y/o ii) adyacentes (o contiguas) a

otras secuencias de proteínas o polinucleótidos que no están naturalmente próximas entre sí. Tal molécula de proteína o polinucleótido, constructo, etc. recombinante, también se puede referir a una molécula o secuencia de proteínas o polinucleótidos que ha sido manipulada genéticamente y/o construida fuera de una célula. Por ejemplo, una molécula de ADN recombinante puede comprender cualquier plásmido, vector, etc. manipulado genéticamente o hecho por la mano del hombre, y puede incluir una molécula de ADN lineal o circular. Tales plásmidos, vectores, etc. pueden contener diversos elementos de mantenimiento que incluyen un origen de replicación procariota y marcador seleccionable, como así también uno o más transgenes o casetes de expresión tal vez en adición a un gen marcador seleccionable vegetal, etc. La expresión “operativamente enlazada” se refiere a un enlace funcional entre un promotor u otro elemento regulador y una secuencia de ADN que se puede transcribir o una secuencia codificadora de un gen (o transgén), de forma tal que el promotor, etc., opera o funciona para iniciar, asistir, afectar, provocar, y/o promover la transcripción y la expresión de la secuencia de ADN que se puede transcribir o la secuencia codificadora, al menos en ciertas células, tejidos, etapas de desarrollo, y/o condiciones.

[077] Se considera que la referencia en esta solicitud a “molécula aislada de ADN”, o un “polinucleótido aislado”, o una frase o término equivalente significa que la molécula de ADN o el polinucleótido es uno que está presente solo o en combinación con otras composiciones, pero no dentro de su ambiente natural. Por ejemplo, los elementos de los ácidos nucleicos tales como secuencia codificadora, secuencia de intrones, secuencia líder no traducida, secuencia promotora, secuencia de terminación transcripcional, y similares, que se encuentran naturalmente dentro del ADN del genoma de un organismo no se consideran “aislados” siempre que el elemento esté dentro del genoma del organismo y en la ubicación dentro del genoma en la cual este se halla naturalmente. No obstante, cada uno de estos elementos, y subpartes de estos elementos, estarían “aislados” dentro del alcance de esta divulgación siempre que el elemento no esté dentro del genoma del organismo y en la ubicación dentro del genoma en la cual este se halla naturalmente. Igualmente, una secuencia de nucleótidos que codifica una proteína o cualquier variante de origen natural de esa proteína sería una secuencia de nucleótidos aislada siempre que la secuencia de nucleótidos no esté dentro del ADN del organismo en el cual la secuencia que codifica la proteína

se halla naturalmente. Una secuencia de nucleótidos sintética que codifique la secuencia de aminoácidos de la proteína de origen natural se consideraría aislada a los fines de esta divulgación. A los fines de esta divulgación, cualquier secuencia de nucleótidos transgénica, es decir, la secuencia de nucleótidos del ADN insertada en el genoma de las células de una planta o bacteria, o presente en un vector extracromosómico, se consideraría una secuencia de nucleótidos aislada si esta está presente dentro de la estructura plasmídica o similar usada para transformar las células, dentro del genoma de la planta o bacteria, o está presente en cantidades detectables en tejidos, progenie, muestras biológicas o productos primarios derivados de la planta o bacteria.

[078] Tal como se entiende comúnmente en la técnica, el término “promotor” generalmente puede referirse a una secuencia de ADN que contiene un sitio de unión a ARN polimerasa, un sitio de inicio de la transcripción y/o una caja TATA y asiste o promueve la transcripción y expresión de una secuencia de polinucleótidos y/o gen (o transgén) asociados que se pueden transcribir. Un promotor se puede producir, variar o derivar de forma sintética a partir de una secuencia promotora conocida o de origen natural o de otra secuencia promotora. Un promotor también puede incluir un promotor quimérico que comprende una combinación de dos o más secuencias heterólogas. Un promotor de la presente divulgación puede así incluir variantes o fragmentos de secuencias promotoras que son similares en su composición pero no idénticas a otra u otras secuencias conocidas o proporcionadas en la presente. Un promotor proporcionado en la presente, o una variante o fragmento de este, puede comprender un “promotor mínimo” que proporciona un nivel basal de transcripción y comprende una caja TATA o una secuencia de ADN equivalente para reconocimiento y unión del complejo de la ARN polimerasa II para el inicio de la transcripción. Un promotor se puede clasificar de acuerdo con una variedad de criterios relacionados con el patrón de expresión de una secuencia o gen (inclusive un transgén) codificadores o que se pueden transcribir, asociados, operativamente enlazados con el promotor, tal como constitutivos, de desarrollo, que funcionan específicamente en un tejido, inducibles, etc. Los promotores que impulsan la expresión en todos o casi todos los tejidos de la planta se denominan promotores “constitutivos”. Los promotores que impulsan la expresión durante ciertos períodos o etapas del desarrollo se denominan promotores “de desarrollo”. Los promotores que impulsan la expresión aumentada en ciertos tejidos de la planta en relación con otros tejidos de la planta se denominan

promotores “potenciadores para el tejido” o “que funcionan preferentemente en un tejido”. Por lo tanto, un promotor “que funciona preferentemente en un tejido” provoca una expresión relativamente superior o preferencial en uno o más tejidos específicos de la planta, pero con niveles menores de expresión en otro u otros tejidos de la planta. Los promotores que se expresan dentro de uno o más tejidos específicos de la planta, con escasa o nula expresión en otros tejidos de la planta, se denominan promotores “que funcionan específicamente en un tejido”. Un promotor “inducible” es un promotor que inicia la transcripción en respuesta a un estímulo del medioambiente tal como frío, sequía o luz, u otros estímulos, tales como lesiones o aplicación química. Un promotor también se puede clasificar en términos de su origen, tal como, heterólogo, homólogo, quimérico, sintético, etc.

[079] Como se usa en la presente, un “promotor que se puede expresar en una planta” se refiere a un promotor que puede iniciar, asistir, afectar, provocar y/o promover la transcripción y expresión de su secuencia de ADN asociada, secuencia codificadora o gen que se pueden transcribir en una célula o un tejido vegetal.

[080] El término “heterólogo” con referencia a un promotor u otra secuencia reguladora en relación con una secuencia de polinucleótidos asociada (por ejemplo, una secuencia de ADN o una secuencia codificadora o gen que se pueden transcribir) es un promotor o una secuencia reguladora que no está operativamente enlazada con dicha secuencia de polinucleótidos asociada en la naturaleza sin la intervención humana -por ejemplo, el promotor o la secuencia reguladora tiene un origen diferente en relación con la secuencia de polinucleótidos asociada y/o el promotor o la secuencia reguladora no aparecen naturalmente en la especie de planta que se ha de transformar con el promotor o la secuencia reguladora.

[081] Como se usa en la presente, un “gen endógeno” o un “locus endógeno” se refiere a un gen o locus y a su localización natural y original en el cromosoma. Como se usa en la presente, el “gen endógeno *TFL1*” se refiere al locus génico de *TFL1* en su localización original en el cromosoma.

[082] Como se usa en la presente, en el contexto de un gen codificador de proteínas, un “exón” se refiere a un segmento de una molécula de ADN o ARN que contiene información que codifica para una secuencia de proteínas o polipéptidos.

[083] Como se usa en la presente, un “intrón” de un gen se refiere a un segmento de una molécula de ADN o ARN, el cual no contiene información que codifica para una proteína o un polipéptido, y el cual se transcribe primero en una secuencia de ARN pero luego se elimina de la molécula de ARN maduro.

[084] Como se usa en la presente, una “región no traducida (UTR)” de un gen se refiere a un segmento de una secuencia o molécula de ARN (por ejemplo, una molécula de ARNm) expresada a partir de un gen (o transgén), pero con la exclusión de las secuencias del exón e intrón de la molécula de ARN. Una “región no traducida (UTR)” también se refiere a un segmento o secuencia de ADN que codifica tal segmento UTR de una molécula de ARN. Una región no traducida puede ser una región 5'-UTR o 3'-UTR dependiendo de si está localizada en el extremo 5' o 3' de una molécula o secuencia de ADN o ARN en relación con una región codificadora de la molécula o secuencia de ADN o ARN (es decir, corriente arriba (5') o corriente abajo (3') de las secuencias del exón e intrón, respectivamente).

[085] Como se usa en la presente, una “secuencia de terminación de la transcripción” se refiere a una secuencia de ácidos nucleicos que contiene una señal que activa la liberación de la molécula transcripta de ARN recién sintetizada de un complejo de ARN polimerasa y marca el fin de la transcripción de un gen o locus.

[086] Como se usa en la presente, un “homólogo” u “homólogos” significa una proteína en un grupo de proteínas que realizan la misma función biológica, por ejemplo, proteínas que pertenecen a la misma familia de proteínas del tipo TFL1 y que proporcionan una rasgo común aumentado en plantas modificadas de esta divulgación. Los homólogos son expresados por genes homólogos. Con referencia a los genes homólogos, los homólogos incluyen los ortólogos, por ejemplo, genes expresados en especies diferentes que evolucionaron de genes ancestrales comunes por especiación y codifican proteínas que retienen la misma función, pero no incluyen los parálogos, es decir, genes que están relacionados por duplicación pero que evolucionaron para codificar proteínas con funciones diferentes. Los genes homólogos incluyen los alelos de origen natural y las variantes creadas artificialmente.

[087] Las expresiones “identidad porcentual”, “% de identidad” o “por ciento idéntica” como se usan en la presente con referencia a dos o más secuencias de proteínas o nucleótidos se calcula (i)

comparando dos secuencias óptimamente alineadas (nucleótido o proteína) en una ventana de comparación, (ii) determinando el número de posiciones en las cuales la base de ácido nucleico (en el caso de secuencias de nucleótidos) o residuos aminoacídicos (para las proteínas) idénticos aparecen en ambas secuencias para obtener el número de posiciones coincidentes, (iii) dividiendo el número de posiciones coincidentes por el número total de posiciones en la ventana de comparación, y luego (iv) multiplicando este cociente por 100 para obtener la identidad porcentual. Si la “identidad porcentual” se calcula con relación a una secuencia de referencia sin especificar una ventana de comparación en particular, entonces la identidad porcentual se determina dividiendo el número de posiciones coincidentes en la región de alineamiento por la longitud total de la secuencia de referencia. Por consiguiente, a los fines de la presente solicitud, cuando dos secuencias (de consulta y sujeto) están óptimamente alineadas (con huecos permitidos en su alineamiento), la “identidad porcentual” de la secuencia de consulta es igual al número de posiciones idénticas entre las dos secuencias dividido por el número total de posiciones en la secuencia de consulta en toda su longitud (o una ventana de comparación), el cual luego se multiplica por 100. Cuando el porcentaje de identidad de secuencia se usa con referencia a proteínas se reconoce que las posiciones de residuos que no son idénticas a menudo difieren en sustituciones conservativas de aminoácidos, donde los residuos aminoacídicos se sustituyen por otros residuos aminoacídicos con propiedades químicas similares (por ejemplo, carga o hidrofobicidad) y por lo tanto no cambian las propiedades funcionales de la molécula. Cuando las secuencias difieren en sustituciones conservativas, el porcentaje de identidad de secuencia se puede ajustar hacia arriba para corregir la naturaleza conservativa de la sustitución. Se dice que las secuencias que difieren en tales sustituciones conservativas tienen “similitud de secuencia” o “similitud”. Las secuencias que tienen una identidad porcentual con una secuencia base pueden exhibir la actividad de la secuencia base.

[088] La degeneración del código genético proporciona la posibilidad de sustituir al menos una base de la secuencia codificadora de la proteína de un gen con una base diferente sin provocar cambios en la secuencia de aminoácidos del polipéptido producido a partir del gen. Cuando están óptimamente alineadas, las proteínas homólogas o sus correspondientes secuencias de nucleótidos, generalmente tienen al menos aproximadamente 70%, al menos aproximadamente 75%, al menos

aproximadamente 80%, al menos aproximadamente 85%, al menos aproximadamente 90%, al menos aproximadamente 92%, al menos aproximadamente 93%, al menos aproximadamente 94%, al menos aproximadamente 95%, al menos aproximadamente 96%, al menos aproximadamente 97%, al menos aproximadamente 98%, al menos aproximadamente 99%, o inclusive al menos aproximadamente 99,5% de identidad en toda la longitud de una proteína o su correspondiente secuencia de nucleótidos identificada como asociada con impartir un fenotipo de determinación modificado cuando se expresa en células vegetales. De acuerdo con realizaciones de la presente invención, el gen *TFL1* o su homólogo codifica una proteína que tiene al menos 70%, al menos 75%, al menos 80%, al menos 85%, al menos 90%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99%, al menos 99,5%, o 100% de identidad de secuencia con SEQ ID NO:2. Los ejemplos de homólogos de la proteína TFL1b de soja (SEQ ID NO:2) incluyen, entre otros, las secuencias de SEQ ID NO:52-66, 78 y 80.

[089] Los homólogos se infieren a partir de la similitud de la secuencia, comparando secuencias de proteínas, por ejemplo, manualmente o con el uso de una herramienta basada en la computación. Para obtener un alineamiento óptimo de secuencias para calcular su identidad porcentual, se conocen en la técnica diversos algoritmos y programas para alineamiento de secuencias múltiple o de a pares, tales como ClustalW o Basic Local Alignment Search Tool® (BLAST), etc., que se pueden usar para comparar la identidad de secuencia o la similitud entre dos o más secuencias de nucleótidos o proteínas. También se puede usar BLAST, por ejemplo, para buscar secuencias de proteínas de un organismo base en una base de datos de secuencias de proteínas de varios organismos, y hallar secuencias similares. El valor de Expectativa generado (valor E) se puede emplear para medir el nivel de similitud de secuencias. El hecho de que una proteína obtenga el valor E más bajo para un organismo en particular no implica que necesariamente sea un ortólogo o sea el único ortólogo, se emplea una consulta recíproca para filtrar las secuencias obtenidas con valores E significativos para la identificación de ortólogos. La consulta recíproca conlleva la búsqueda de aciertos significativos frente a una base de datos de secuencias de proteínas del organismo base. Un acierto se puede identificar como un ortólogo cuando el mejor acierto de la consulta recíproca es la propia proteína o un parálogo de la proteína de consulta. Donde los

ortólogos del proceso de consulta recíproca luego son diferenciados de los parálogos entre todos los homólogos, lo cual permite la inferencia de los genes con equivalencia funcional.

[090] La expresión “complementariedad porcentual” o como se usa en la presente con referencia a dos secuencia de nucleótidos, es similar al concepto de identidad porcentual pero se refiere al porcentaje de nucleótidos de una secuencia de consulta que forman pares de bases o hibridan nucleótidos de una secuencia sujeto de manera óptima cuando las secuencias de consulta y sujeto están dispuestas linealmente y sus bases óptimamente apareadas sin estructuras de plegamiento secundarias, tales como bucles, tallos u horquillas. Dicha complementariedad porcentual puede ser entre dos hebras de ADN, dos hebras de ARN, o una hebra de ADN y una hebra de ARN. La “complementariedad porcentual” se calcula (i) apareando bases o hibridando de manera óptima las dos secuencia de nucleótidos en una disposición lineal y completamente extendida (es decir, sin plegamientos ni estructuras secundarias) en una ventana de comparación, (ii) determinando el número de posiciones donde se aparean pares de bases entre las dos secuencias en la ventana de comparación para obtener el número de posiciones complementarias, (iii) dividiendo el número de posiciones complementarias por el número total de posiciones en la ventana de comparación, y (iv) multiplicando este cociente por 100 para obtener el porcentaje de complementariedad de las dos secuencias. Se puede determinar el apareamiento óptimo de bases de dos secuencias sobre la base de los apareamientos conocidos de bases de nucleótidos, tales como G-C, A-T y A-U, a través de enlaces de hidrógeno. Si la “complementariedad porcentual” se calcula con relación a una secuencia de referencia sin especificar una ventana de comparación en particular, entonces la complementariedad porcentual se determina dividiendo el número de posiciones complementarias entre las dos secuencias lineales por la longitud total de la secuencia de referencia. Por lo tanto, a los fines de la presente divulgación, cuando el apareamiento de bases de dos secuencias (consulta y sujeto) es óptimo (con admisión de falta de coincidencias o falta de apareamiento de bases de nucleótidos, pero sin admisión de plegamientos ni estructuras secundarias), la “complementariedad porcentual” de la secuencia de consulta es igual al número de posiciones de pares de bases apareadas entre las dos secuencias dividido por el número total de posiciones en la secuencia de consulta en toda su longitud (o por el número de posiciones en la secuencia consulta en una ventana de comparación), el cual luego se multiplica por 100.

[091] Como se usa en la presente, un “fragmento” de un polinucleótido se refiere a una secuencia que comprende al menos aproximadamente 50, al menos aproximadamente 75, al menos aproximadamente 95, al menos aproximadamente 100, al menos aproximadamente 125, al menos aproximadamente 175, al menos aproximadamente 200, al menos aproximadamente 225, al menos aproximadamente 250, al menos aproximadamente 275, al menos aproximadamente 300, al menos aproximadamente 500, al menos aproximadamente 600, al menos aproximadamente 700, al menos aproximadamente 750, al menos aproximadamente 800, al menos aproximadamente 900 o al menos aproximadamente 1000 nucleótidos contiguos, o más, de una proteína o molécula de ADN como se divulga en la presente. Los métodos para producir tales fragmentos a partir de una molécula promotora de inicio son bien conocidos en la técnica. Los fragmentos de una proteína o molécula de ADN pueden exhibir la actividad de la proteína o molécula de ADN de la cual derivan.

[092] De acuerdo con otro aspecto, la presente divulgación proporciona métodos para alterar un fenotipo, tal como aumento del rendimiento, modificación de la determinación, o reducción del encamado en una planta que comprenden: (a) modificar el genoma de una célula vegetal mediante: (i) identificación de un gen endógeno de la planta correspondiente al gen *TFL1* de soja, tal como el gen *GmTFL1b* descrito en la presente, y sus homólogos, y (ii) modificación de la secuencia promotora del gen endógeno en la célula vegetal mediante mutagénesis dirigida para modificar el nivel de expresión del gen endógeno; y (b) regeneración o desarrollo de una planta a partir de la célula vegetal. Diversos genes y proteínas *TFL1* de diferentes especies vegetales se pueden identificar y considerar homólogos u ortólogos de *TFL1* para usar en la presente divulgación si estos tienen una secuencia de proteínas y/o ácido nucleico similar y comparten aminoácidos y/o uno o más dominios estructurales conservados con al menos un gen o proteína *TFL1* conocida.

[093] Se puede usar un transgén marcador seleccionable vegetal en un vector o constructo de transformación de la presente divulgación para asistir en la selección de células o tejido transformado debido a la presencia de un agente de selección, tal como un antibiótico o herbicida, donde el transgén marcador seleccionable vegetal proporciona tolerancia o resistencia al agente de selección. Por lo tanto, el agente de selección puede tener un sesgo que favorezca la supervivencia, el desarrollo, el crecimiento, la proliferación, etc. de células transformadas que expresan el gen marcador seleccionable vegetal, tal como para aumentar la proporción de células o tejidos

transformados en la planta R₀. Los genes marcadores seleccionables vegetales comúnmente utilizados incluyen, por ejemplo, aquellos que confieren tolerancia o resistencia a antibióticos, tales como kanamicina y paromomicina (*nptII*), higromicina B (*aph IV*), estreptomicina o espectinomomicina (*aadA*) y gentamicina (*aac3* y *aacC4*), o aquellos que confieren tolerancia o resistencia a herbicidas tales como glufosinato (*bar* o *pat*), dicamba (DMO) y glifosato (*proA* o EPSPS). También se pueden emplear genes marcadores detectables vegetales, que proporcionan la capacidad de detectar visualmente transformantes, tales como luciferasa o proteína fluorescente verde (GFP), o un gen que expresa un gen *uidA* (GUS) o *glucuronidasa beta* para los cuales se conocen diversos sustratos cromogénicos. La transformación vegetal también se puede llevar a cabo en ausencia de selección durante uno o más pasos o etapas de cultivo, desarrollo o regeneración de explantes, tejidos, plantas y/o partes de plantas transformadas.

III. Métodos de transformación

[094] Se proporcionan métodos y composiciones para la transformación de una célula vegetal, tejido o explante con una molécula o constructo de ADN recombinante que codifica una o más moléculas requeridas para edición genómica localizada (por ejemplo, uno o más ARN guía y/o nucleasas dirigidas al sitio). Los métodos adecuados para la transformación de células vegetales hospederas incluyen virtualmente cualquier método mediante el cual se puede introducir ADN o ARN en una célula (por ejemplo, donde un constructo de ADN recombinante es establemente integrado en un cromosoma de una planta o donde un constructo de ADN recombinante o un ARN es transitoriamente proporcionado a una célula vegetal) y son ampliamente conocidos en la técnica. Dos métodos efectivos para transformación celular son transformación mediada por bacterias, tal como mediada por *Agrobacterium* o transformación mediada por *Rhizobium*, y transformación mediada por bombardeo de microproyectiles o partículas. Los métodos para bombardeo con microproyectiles se ilustran, por ejemplo, en las Patentes de los EE.UU. nro. 5.550.318; 5.538.880; 6.160.208 y 6.399.861. Los métodos de transformación mediada por *Agrobacterium* se describen, por ejemplo, en la Patente de los EE.UU. nro. 5.591.616. Otros métodos de transformación vegetal, tales como microinyección, electroporación, infiltración por vacío, presión, ultrasonido, agitación con fibras de carburo de silicio, transformación mediada por PEG, etc., también son conocidos en la técnica.

[095] La transformación de material vegetal se practica en cultivo tisular en un medio nutriente, por ejemplo, una mezcla de nutrientes que permita a las células crecer *in vitro*. Las células receptoras diana incluyen, entre otras, células meristemáticas, células de puntas de brotes, de hipocótilo, de callos, de embriones maduros e inmaduros, y células gaméticas tales como de microsporas y pollen. El callo se puede iniciar a partir de fuentes tisulares que incluyen, entre otras, embriones inmaduros o maduros, hipocótilos, meristemos apicales de plántulas, microsporas y similares. Las células que contienen núcleo transgénico se cultivan hasta obtener plantas transgénicas. De acuerdo con los métodos de la presente se puede emplear cualquier técnica o método adecuado para la transformación de una célula vegetal que se conozca en la técnica. Para la transformación, generalmente se introduce ADN solo en un pequeño porcentaje de las células vegetales diana en cualquier experimento de transformación. Se emplean genes marcadores para proporcionar un sistema eficiente para identificar aquellas células que se transforman de manera estable al recibir e integrar una molécula de ADN recombinante en sus genomas.

[096] Como se usan en la presente, los términos “regeneración” y “regenerando” se refieren al proceso de cultivar y desarrollar una planta a partir de una o más células vegetales a través de uno o más pasos de cultivo. Las células, los explantes, los tejidos transformados o editados que contienen una edición o inserción de secuencia de ADN se pueden cultivar, desarrollar o regenerar en plantas transgénicas en cultivos, cepellón o suelo, de acuerdo con métodos conocidos en la técnica. Por lo tanto, ciertas realizaciones de la divulgación están relacionadas con métodos y constructos para regenerar una planta a partir de una célula con ADN genómico modificado obtenida mediante edición genómica. La planta regenerada luego se puede usar para propagar plantas adicionales.

[097] De acuerdo con un aspecto de la presente divulgación, plantas regeneradas o una planta de progenie, parte de planta o semilla de esta se puede cribar o seleccionar sobre la base de un marcador, rasgo o fenotipo producido por la edición o mutación, o mediante integración dirigida al sitio de una secuencia, transgén de inserción, etc., en la planta desarrollada o regenerada, o una planta de progenie, parte de planta o semilla de esta. Si una mutación, edición, rasgo o fenotipo dado es recesivo, pueden ser necesarias una o más generaciones o cruzamientos (por ejemplo, autofecundación) a partir de la planta inicial R_0 para producir una planta homocigota para la

edición o mutación de modo que el rasgo o fenotipo pueda ser observado. Las plantas de progenie, tales como las plantas cultivadas de la semilla R₁ o en generaciones siguientes, se pueden ensayar para determinar su cigosidad usando cualquier ensayo de cigosidad, tal como usando un ensayo de polimorfismo de nucleótido único (SNP), secuenciación de ADN, amplificación térmica, o reacción en cadena de la polimerasa (PCR), y/o inmunotransferencia (Southern blotting) que posibilitan la distinción entre plantas heterocigotas, homocigotas y de tipo silvestre.

[098] Se proporcionan métodos y técnicas para cribado, y/o identificación de células o plantas, etc., para determinar la presencia de ediciones o transgenes localizados, y seleccionar células o plantas que comprenden ediciones o transgenes localizados, los cuales pueden estar basados en uno o más fenotipos o rasgos, o en la presencia o ausencia de un marcador molecular o secuencia de polinucleótidos o proteínas en las células o plantas. Como se usa en la presente, una “técnica molecular” se refiere a cualquier método conocido en los campos de la biología molecular, bioquímica, genética, biología vegetal, o biofísica que suponga el uso, la manipulación o el análisis de un ácido nucleico, una proteína o un lípido. Sin resultar taxativos, las técnicas moleculares útiles para detectar la presencia de una secuencia modificada en un genoma incluyen cribado fenotípico; tecnologías de marcación molecular tales como análisis SNP mediante tecnología TaqMan® o Illumina/Infinium; inmunotransferencia (Southern blot); PCR; ensayos por inmunoadsorción ligado a enzimas (ELISA); y secuenciación (por ejemplo, Sanger, Illumina®, 454, Pac-Bio, Ion Torrent™). En un aspecto, un método de detección proporcionado en la presente comprende cribado fenotípico. En otro aspecto, un método de detección proporcionado en la presente comprende análisis SNP. En un aspecto adicional, un método de detección proporcionado en la presente comprende Southern blot. En un aspecto adicional, un método de detección proporcionado en la presente comprende PCR. En un aspecto, un método de detección proporcionado en la presente comprende ELISA. En un aspecto adicional, un método de detección proporcionado en la presente comprende determinar la secuencia de un ácido nucleico o una proteína. Sin ser taxativos, los ácidos nucleicos se pueden detectar usando hibridación. La hibridación entre ácidos nucleicos se expone en detalle en Sambrook *et al.* (1989, *Molecular Cloning: A Laboratory Manual*, 2da. Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).

[099] Los ácidos nucleicos se pueden aislar usando técnicas de rutina en la técnica. Por ejemplo, se pueden aislar ácidos nucleicos usando cualquier método, inclusive, entre otros, tecnología de ácido nucleico recombinante, y/o PCR. Las técnicas generales de PCR se describen, por ejemplo, en *PCR Primer: A Laboratory Manual*, Dieffenbach & Dveksler, Eds., Cold Spring Harbor Laboratory Press, 1995. Las técnicas de ácidos nucleicos recombinantes incluyen, por ejemplo, digestión de enzimas de restricción y ligamiento, las cuales se pueden usar para aislar un ácido nucleico. Los ácidos nucleicos aislados también se pueden sintetizar químicamente, tanto como una sola molécula de ácido nucleico como una serie de oligonucleótidos.

[0100] La detección (por ejemplo, de un producto de amplificación, de un complejo de hibridación, de un polipéptido) se puede lograr usando marcadores detectables que pueden estar unidos o asociados con un anticuerpo o sonda de hibridación. Se considera que el término “marcador” abarca el uso de marcadores directos como así también marcadores indirectos. Los marcadores detectables incluyen enzimas, grupos prostéticos, materiales fluorescentes, materiales luminiscentes, materiales bioluminiscentes, y materiales radiactivos. El cribado y selección de plantas modificadas (por ejemplo, editadas) o células vegetales puede realizarse a través de cualquier metodología conocida por los expertos en la técnica de la biología molecular. Los ejemplos de metodologías de cribado y selección incluyen, entre otros, análisis Southern, amplificación por PCR para la detección de un polinucleótido, Northern blots (inmunotransferencia), protección RNasa, extensión de cebador, amplificación por PCR con transcriptasa inversa (RT-PCR) para detectar transcritos de ARN, secuenciación Sanger, tecnologías de secuenciación de próxima generación (por ejemplo, Illumina®, PacBio®, Ion Torrent™, etc.), ensayos enzimáticos para detectar actividad de enzimas o ribozimas de polipéptidos y polinucleótidos, y electroforesis en gel de proteínas, Western blots (inmunotransferencia), inmunoprecipitación, e inmunoensayos enlazados a enzimas para detectar polipéptidos. También se pueden usar otras técnicas tales como hibridación *in situ*, tinción de enzimas, e inmunotinción para detectar la presencia o expresión de polipéptidos y/o polinucleótidos. En la técnica se conocen métodos para realizar todas las técnicas referidas.

[0101] Como se usa en la presente, el término “polipéptido” se refiere a una cadena de al menos dos aminoácidos enlazados de manera covalente. Los polipéptidos se pueden codificar mediante

polinucleótidos proporcionados en la presente. Un ejemplo de un polipéptido es una proteína. Las proteínas proporcionadas en la presente se pueden codificar mediante las moléculas de ácido nucleico proporcionadas en la presente. Los polipéptidos se pueden purificar a partir de fuentes naturales (por ejemplo, una muestra biológica) aplicando métodos conocidos tales como intercambio iónico DEAE, filtración en gel y cromatografía de hidroxapatita. Un polipéptido también se puede purificar, por ejemplo, expresando un ácido nucleico en un vector de expresión. Además, se puede obtener un polipéptido purificado mediante síntesis química. El grado de pureza de un polipéptido se puede medir usando cualquier método apropiado, por ejemplo, cromatografía en columna, electroforesis en gel de poliacrilamida, o análisis HPLC (cromatografía líquida de alto rendimiento, por sus siglas en inglés).

[0102] Los polipéptidos se pueden detectar usando anticuerpos. Las técnicas para detectar polipéptidos usando anticuerpos incluyen ensayos o inmunoabsorción ligados a enzimas (ELISA), Western blots, inmunoprecipitaciones e inmunofluorescencia. Un anticuerpo proporcionado en la presente puede ser un anticuerpo policlonal o un anticuerpo monoclonal. Se puede generar un anticuerpo con afinidad de unión específica para un polipéptido proporcionado en la presente usando métodos ampliamente conocidos en la técnica. Un anticuerpo proporcionado en la presente se puede unir a un soporte sólido tal como una placa de microtitulación usando métodos conocidos en la técnica.

[0103] Una planta que se puede transformar con un vector de transformación o molécula de ADN recombinante que comprende un ARN guía puede incluir una variedad de plantas florales o angiospermas, las cuales se pueden definir ulteriormente como que incluyen especies de plantas dicotiledóneas (dicot) o especies de plantas monocotiledóneas (monocot). Una planta dicotiledónea podrían ser miembros de la familia *Fabaceae* (tal como legumbres), girasol (*Helianthus annuus*), cártamo (*Carthamus tinctorius*), sésamo (*Sesamum spp.*), tabaco (*Nicotiana tabacum*), patata (*Solanum tuberosum*), algodón (*Gossypium barbadense*, *Gossypium hirsutum*), batata (*Ipomoea batatas*), mandioca (*Manihot esculenta*), café (*Coffea spp.*), té (*Camellia spp.*), árboles frutales, tales como manzano (*Malus spp.*), *Prunus spp.*, tales como ciruelo, damasco, duraznero, cerezo, etc., peral (*Pyrus spp.*), higo (*Ficus carica*), etc., árboles de cítricos (*Citrus spp.*), cacao (*Theobroma cacao*), palta (*Persea americana*), olivo (*Olea europaea*), almendro

(*Prunus amygdalus*), nogal (*Juglans* spp.), frutilla (*Fragaria* spp.), melón (*Citrullus lanatus*), pimiento (*Capsicum* spp.), remolacha (*Beta vulgaris*), vid (*Vitis*, *Muscadinia*), tomate (*Lycopersicon esculentum*, *Solanum lycopersicum*), pepino (*Cucumis sativus*), y miembros de la familia *Brassicaceae*, tales como berro de thale (*Arabidopsis thaliana*) y *Brassica* sp. (por ejemplo, *B. napus*, *B. rapa*, *B. juncea*), particularmente aquellas especies de *Brassica* útiles como fuente de aceite de semillas. Las legumbres y las plantas leguminosas incluyen arvejas (*Pisum sativum*) alfalfa (*Medicago sativa*), trébol barril (*Medicago truncatula*), guandú (*Cajanus cajan*) guar (*Cyamopsis tetragonoloba*), algarrobo (*Ceratonia siliqua*), fenogreco (*Trigonella foenum-graecum*), soja (*Glycine max*), poroto común (*Phaseolus vulgaris*), caupí (*Vigna unguiculata*), frijol mungo (*Vigna radiata*), frijol de Lima (*Phaseolus lunatus*), haba (*Vicia faba*), lenteja (*Lens culinaris* o *Lens esculenta*), maní (*Arachis hypogaea*), regaliz (*Glycyrrhiza glabra*), y garbanzo (*Cicer arietinum*). Una planta monocotiledónea podría ser palma oleaginosa (*Elaeis* spp.), coco (*Cocos* spp.), banana (*Musa* spp.), y cereales tales como maíz (*Zea mays*), cebada (*Hordeum vulgare*), sorgo (*Sorghum bicolor*), arroz (*Oryza sativa*), y trigo (*Triticum aestivum*). Debido a que la presente divulgación se puede aplicar a un amplio rango de especies vegetales, la presente divulgación se aplica además a otras estructuras botánicas análogas a vainas de plantas leguminosas, tales como cápsulas, silículas, frutas, nueces, tubérculos, etc.

IV. Plantas modificadas genómicamente

[0104] Como se usa en la presente, “modificado” en el contexto de una planta, semilla de una planta, parte de una planta, célula vegetal y/o genoma vegetal, se refiere a una planta, semilla de planta, parte de planta, célula vegetal, y/o genoma vegetal que comprende una modificación genética en el nivel de expresión y/o secuencia endógena de uno o más genes de interés relativos a una planta de tipo silvestre o de control, semilla de planta, parte de planta, célula vegetal, y/o genoma vegetal. Por cierto, el término “modificado” se puede referir además a una planta, a una semilla de una planta, a una célula vegetal y/o al genoma vegetal que tiene una o más delecciones que afectan la expresión de un gen *TFLI* endógeno introducidas a través de mutagénesis química, inserción o excisión de transposones, o cualquier otra técnica de mutagénesis conocida, o introducidas a través de edición genómica. En un aspecto, una planta, semilla de planta, parte de planta, célula vegetal y/o genoma vegetal modificados pueden comprender uno o más transgenes.

Por lo tanto, para mayor claridad, una planta, semilla de planta, parte de planta, célula vegetal y/o genoma vegetal modificados incluyen una planta, semilla de planta, parte de planta, célula vegetal y/o genoma vegetal mutados, editados y/o transgénicos que tienen un nivel de expresión, patrón de expresión y/o secuencia modificados de un gen *TFL1* en relación con una planta, semilla de planta, parte de planta, célula vegetal y/o genoma vegetal de tipo silvestre o de control.

[0105] Las plantas, partes de plantas, semillas, etc., pueden haber sido sometidas a mutagénesis, edición genómica o integración dirigida al sitio, transformación genética, o una combinación de estas técnicas. Tales plantas, semillas de plantas, partes de plantas y células vegetales “modificadas” incluyen plantas, semillas de plantas, partes de plantas y células vegetales que son la progenie o derivan de plantas, semillas de plantas, partes de plantas y células vegetales “modificadas” que retienen el cambio molecular (por ejemplo, cambio en el nivel y/o la actividad de expresión) del gen *TFL1*. Una semilla modificada proporcionada en la presente puede dar lugar a una planta modificada proporcionada en la presente. Una planta, semilla de planta, parte de planta, célula vegetal o genoma vegetal modificado proporcionado en la presente puede comprender un constructo o vector de ADN recombinante o una edición genómica como se proporciona en la presente. Un “producto de planta modificada” puede ser cualquier producto elaborado a partir de una planta, parte de planta, célula vegetal o cromosoma vegetal modificado proporcionado en la presente, o cualquier porción o componente de estos.

[0106] Las plantas modificadas además se pueden retrocruzar o cruzar con otras plantas para producir semillas y progenie de plantas modificadas. También se puede preparar una planta modificada al cruzar una primera planta que comprende una secuencia o un constructo o una edición de ADN (por ejemplo, una delección genómica) con una segunda planta que carece de la secuencia, el constructo o la edición de ADN. Por ejemplo, se puede introducir una secuencia o inversión de ADN en una primera línea de plantas que es susceptible de transformación o edición, la cual luego se puede cruzar con una segunda línea de plantas para introgresar la secuencia o la edición de ADN (por ejemplo, delección) en la segunda línea de plantas. Además, la progenie de estos cruzamientos se pueden retrocruzar para obtener la línea deseada múltiples veces, tal como a través de 6 a 8 generaciones o retrocruzamientos, para producir una planta de progenie que tenga sustancialmente el mismo genotipo que la línea parental original, excepto por la introducción de

la secuencia o edición de ADN. Una planta, célula vegetal o semilla proporcionada en la presente puede ser una planta, célula vegetal o semilla híbrida. Como se usa en la presente, un “híbrido” se crea mediante el cruzamiento de dos plantas de diferentes variedades, líneas, cruzamientos endogámicos, o especies, de forma tal que la progenie comprende material genético de cada parental. Los expertos en la técnica reconocen que también se pueden crear híbridos de orden más alto.

[0107] Una planta, parte de planta, célula vegetal o semilla modificada proporcionada en la presente puede ser de una variedad de elite o una línea de elite. Una “variedad de elite” o un “cultivar de elite” se refiere a una variedad que ha resultado del mejoramiento y la selección para un desempeño agronómico superior.

[0108] Como se usa en la presente, la expresión “planta de control” (o asimismo una semilla de planta, parte de planta, célula vegetal, y/ o genoma vegetal de “control”) se refiere a una planta (o semilla de planta, parte de planta, célula vegetal y/o genoma vegetal) que se usa para comparar con una planta modificada (o semilla de planta, parte de planta, célula vegetal y/o genoma vegetal modificado) y tiene la misma o similar caracterización genética (por ejemplo, las mismas líneas parentales, cruzamiento híbrido, línea endogámica, probadores, etc.) que la planta modificada (o semilla de planta, parte de planta, célula vegetal y/o genoma vegetal), excepto por una o más ediciones genómicas (por ejemplo, una delección) que afecta al gen *TFLI*. Por ejemplo, una planta de control puede ser una línea endogámica que es igual a la línea endogámica usada para producir la planta modificada, o una planta de control puede ser el producto del mismo cruzamiento híbrido de líneas parentales endogámicas que la planta modificada, excepto por la ausencia en la planta de control de cualquier evento o edición genómica que afecta al gen *TFLI*. De modo similar, una “planta de control sin modificar” se refiere a una planta que comparte una caracterización genética sustancialmente similar o esencialmente idéntica a la de la planta modificada, pero sin la única o las varias modificaciones genéticas en el genoma (por ejemplo, mutación o edición) de la planta modificada. A los fines de la comparación con una planta, semilla de planta, parte de planta, célula vegetal y/o genoma vegetal modificados, una “planta de tipo silvestre” (o de igual manera una semilla de planta, parte de planta, célula vegetal y/o genoma vegetal del “tipo silvestre”) se refiere a una planta, semilla de planta, parte de planta, célula vegetal y/o genoma vegetal de control no

transgénicos cuyos genomas no han sido editados. Como se usa en la presente, una planta, semilla de planta, parte de planta, célula vegetal y/o genoma vegetal “de control” también puede ser una planta, semilla de planta, parte de planta, célula vegetal y/o genoma vegetal que tiene caracterización genética similar (pero no igual o idéntica) a la de una planta, semilla de planta, parte de planta, célula vegetal y/o genoma vegetal modificado, si se la considera suficientemente similar para la comparación de las características o rasgos que se han de analizar.

[0109] Como se usan en la presente, las expresiones “suprimir”, “supresión”, “inhibir” “inhibición”, “inhibiendo”, “bloqueo”, “reducción” y “regulación descendente” se refieren a un descenso, reducción o eliminación del nivel de expresión de un ARNm y/o proteína codificados por un gen diana en una planta, célula vegetal, o tejido vegetal en una o más etapas del desarrollo vegetal, en comparación con el nivel de dicho ARNm y/o proteína diana en una planta, célula vegetal o tejido vegetal de control o de tipo silvestre en las mismas etapas del desarrollo vegetal. De acuerdo con algunas realizaciones, se proporciona una planta modificada que tiene un nivel de expresión del gen *TFLI* que está reducido en al menos un tejido de la planta en al menos 5%, al menos 10%, al menos 20%, al menos 25%, al menos 30%, al menos 40%, al menos 50%, al menos 60%, al menos 70%, al menos 75%, al menos 80%, al menos 90% o 100% en comparación con una planta de control. De acuerdo con realizaciones adicionales, se proporciona una planta modificada que tiene un nivel de expresión del gen *TFLI* que está reducido en al menos un tejido de la planta en 5%-20%, 5%-25%, 5%-30%, 5%-40%, 5%-50%, 5%-60%, 5%-70%, 5%-75%, 5%-80%, 5%-90%, 5%-100%, 75%-100%, 50%-100%, 50%-90%, 50%-75%, 25%- 75%, 30%-80% o 10%-75%, en comparación con una planta de control.

[0110] De acuerdo con algunas realizaciones, se proporciona una planta modificada que tiene un nivel de ARNm de *TFLI* que está reducido en al menos un tejido de la planta en al menos 5%, al menos 10%, al menos 20%, al menos 25%, al menos 30%, al menos 40%, al menos 50%, al menos 60%, al menos 70%, al menos 75%, al menos 80%, al menos 90% o 100% en comparación con una planta de control. De acuerdo con realizaciones, se proporciona una planta modificada que tiene un nivel de expresión de ARNm de *TFLI* que está reducido en al menos un tejido de la planta en 5%-20%, 5%-25%, 5%- 30%, 5%-40%, 5%-50%, 5%-60%, 5%-70%, 5%-75%, 5%-80%, 5%-90%, 5%-100%, 75%-100%, 50%-100%, 50%-90%, 50%-75%, 25%-75%, 30%-80% o 10%-

75%, en comparación con una planta de control. De acuerdo con algunas realizaciones, se proporciona una planta modificada que tiene un nivel de expresión de proteína *TFL1* que está reducido en al menos un tejido de la planta en al menos 5%, al menos 10%, al menos 20%, al menos 25%, al menos 30%, al menos 40%, al menos 50%, al menos 60%, al menos 70%, al menos 75%, al menos 80%, al menos 90% o 100% en comparación con una planta de control. De acuerdo con algunas realizaciones, se proporciona una planta modificada que tiene un nivel de expresión de proteína TFL1 que está reducido en al menos un tejido de la planta en 5%-20%, 5%-25%, 5%-30%, 5%-40%, 5%-50%, 5%-60%, 5%-70%, 5%-75%, 5%-80%, 5%-90%, 5%-100%, 75%-100%, 50%-100%, 50%-90%, 50%-75%, 25%-75%, 30%-80% o 10%-75%, en comparación con una planta de control.

[0111] La presente divulgación está relacionada con una planta con características económicamente importantes mejoradas, inclusive, entre otras, aumento del rendimiento, aumento de la determinación, tiempo reducido para alcanzar la fecha de floración terminal, tasa de encamado reducida, y susceptibilidad reducida a la enfermedad fúngica. Más específicamente con respecto al rendimiento, la presente divulgación está relacionada con una planta modificada que comprende una mutación o edición genómica según se describe en la presente, donde la planta tiene rendimiento aumentado en comparación con una planta de control. Muchas plantas de esta divulgación exhiben componentes de aumento del rendimiento o rasgos de rendimiento mejorados en comparación con una planta de control. El rendimiento se puede definir como el producto medible de valor económico de un cultivo. El rendimiento se puede definir en cuanto a cantidad y/o calidad. Por ejemplo, el rendimiento de la soja incluye vainas por planta, vainas por acre, semillas por planta, semillas por vaina, peso por semilla, peso por vaina, vainas por nudo, número de nudos, y el número de entrenudos por planta. El rendimiento puede depender directamente de diversos factores, por ejemplo, el número y el tamaño de los órganos, la arquitectura vegetal (tal como el número de ramas, biomasa de la planta, etc.), momento y duración de la floración, período de llenado de granos. La arquitectura y el desarrollo radicular, la eficacia de la fotosíntesis, la absorción de nutrientes, la tolerancia al estrés, el vigor temprano, la senescencia retardada y fenotipos funcionales de senescencia foliar retardada (*stay-green*) pueden ser factores que

determinan el rendimiento. Por lo tanto, la optimización de los factores mencionados puede contribuir con el aumento del rendimiento de los cultivos.

[0112] Las plantas modificadas que comprenden o derivan de células vegetales que están transformadas con un ADN recombinante de esta divulgación se pueden mejorar adicionalmente con rasgos apilados, por ejemplo, una planta de cultivo modificada que tiene un rasgo mejorado que resulta en la expresión del ADN divulgado en la presente en combinación con uno o más genes de interés agronómico que proporcionan un rasgo agronómico beneficioso (tal como rasgos de resistencia a herbicidas y/o plaguicidas) a las plantas de cultivo. Por ejemplo, los rasgos conferidos por los constructos de ADN recombinante de la presente divulgación se pueden apilar con otros rasgos de interés agronómico, tales como un rasgo que proporciona resistencia a insectos tales como el uso de un gen de *Bacillus thuringiensis* para proporcionar resistencia frente a lepidópteros, coleópteros, homópteros, hemiópteros y otros insectos, o rasgos de calidad mejorados tales como valor nutricional mejorado. Las moléculas y los métodos para impartir resistencia frente a insectos, nemátodos y virus se divulgan en las Patentes de los EE.UU. nro. 5.250.515; 5.880.275; 6.506.599; 5.986.175; y en la Publicación de Solicitud de Patente de los EE.UU. nro. 2003/0150017 A1.

[0113] Los herbicidas respecto de los cuales se ha demostrado la tolerancia de plantas transgénicas y a los que se les pueden aplicar los métodos y composiciones de la presente divulgación incluyen, entre otros, glifosato, dicamba, glufosinato, sulfonilurea, bromoxinilo, norflurazón, ácido 2,4-D (2,4-diclorofenoxi) acético, ariloxifenoxi-propionatos, inhibidores de p-hidroxifenil piruvato dioxigenasa (HPPD) y herbicidas inhibidores de protoporfirinógeno oxidasa (PPO). Las moléculas de polinucleótidos que codifican las proteínas que participan en la tolerancia a herbicidas conocidas en la técnica incluyen, entre otras, una molécula de polinucleótido que codifica la 5-enolpiruvilshikimato-3-fosfato sintasa (EPSPS) divulgada en las Patentes de los EE.UU. nro. 5.094.945, 5.627.061, 5.633.435 y 6.040.497 para impartir tolerancia a glifosato; moléculas de polinucleótido que codifican una glifosato óxidorreductasa (GOX) divulgadas en la Patente de los EE.UU. nro. 5.463.175 y una glifosato-N-acetil transferasa (GAT) divulgada en la Publicación de Patente de los EE.UU. nro. 2003/0083480 A1 también para impartir tolerancia a glifosato; dicamba monooxigenasa divulgada en la Publicación de Solicitud Patente de los EE.UU. nro. 2003/0135879 A1 para impartir tolerancia a dicamba; una molécula de polinucleótido que codifica la bromoxinil

nitrilasa (Bxn) divulgada en la Patente de los EE.UU. nro. 4.810.648 para impartir tolerancia a bromoxinilo; una molécula de polinucleótido que codifica la fitoeno desaturasa (crtl) descrita en Misawa *et al.* (*Plant J.* 4:833-840, 1993) y en Misawa *et al.* (*Plant J.* 6:481-489, 1994) para tolerancia a norflurazón; una molécula de polinucleótido que codifica una acetohidroxiácido sintasa (AHAS, también conocida como ALS) descrita en Sathasivan *et al.* (*Nucl. Acids Res.* 18:2188-2193, 1990) para impartir tolerancia a herbicidas sulfonilurea; moléculas de polinucleótidos conocidas como genes *bar* divulgados en DeBlock *et al.* (*EMBO J.* 6:2513-2519, 1987) para impartir tolerancia a glufosinato y bialafos; moléculas de polinucleótidos divulgadas en la Publicación de Solicitud de Patente de los EE.UU. 2003/010609 A1 para impartir tolerancia al ácido N-amino metil fosfónico; moléculas de polinucleótidos divulgadas en la Patente de los EE.UU. nro. 6.107.549 para impartir resistencia al herbicida piridina; moléculas y métodos para impartir tolerancia a múltiples herbicidas tales como glifosato, atrazina, inhibidores ALS, herbicidas isoxaflutol y glufosinato se divulgan en la Patente de los EE.UU. nro. 6.376.754 y la Publicación de Solicitud de Patente de los EE.UU. 2002/0112260.

[0114] También se pueden emplear con la presente divulgación elementos genéticos, métodos y transgenes que confieren resistencia a enfermedades fúngicas (Patentes de los EE.UU. nro. 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, 6.506.962). Las enfermedades de la soja provocadas por hongos incluyen, entre otras, *Phakopsora pachyrhizi*, *Phakopsora meibomia* (roya de la soja asiática), *Colletotrichum truncatum*, *Colletotrichum dematium* var. *truncatum*, *Glomerella glycines* (antracnosis de la soja), *Phytophthora sojae* (*Phytophthora* podredumbre de tallo y raíz), *Sclerotinia sclerotiorum* (*Sclerotinia* stem rot), *Fusarium solani* f. sp. *glycines* (síndrome de muerte súbita), *Fusarium* spp. (*Fusarium* podredumbre de la raíz), *Macrophomina phaseolina* (podredumbre carbonosa), *Septoria glycines*, (septoriosis), *Pythium aphanidermatum*, *Pythium debaryanum*, *Pythium irregulare*, *Pythium ultimum*, *Pythium myriotylum*, *Pythium torulosum* (deterioro de la semilla por *Pythium*), *Diaporthe phaseolorum* var. *sojae* (añublo de la vaina), *Phomopsis longicola* (añublo del tallo), *Phomopsis* spp. (deterioro de la semilla por *Phomopsis*), *Peronospora manshurica* (Downy mildiu), *Rhizoctonia solani* (podredumbre de raíz y tallo por *Rhizoctonia*, añublo aéreo por *Rhizoctonia*), *Phialophora gregata* (podredumbre marrón del tallo), *Diaporthe phaseolorum*

var. *caulivora* (cancro del tallo), *Cercospora kikuchii* (tinción púrpura de la semilla), *Alternaria* sp. (añublo temprano), *Cercospora sojina* (mancha foliar ojo de rana), *Sclerotium rolfsii* (añublo sureño), *Arkoola nigra* (añublo negro foliar), *Thielaviopsis basicola*, (podredumbre negra de la raíz), *Choanephora infundibulifera*, *Choanephora trispora* (añublo foliar por *Choanephora*), *Leptosphaerulina trifolii* (mancha foliar por *Leptosphaerulina*), *Mycoleptodiscus terrestris* (podredumbre de la raíz por *Mycoleptodiscus*), *Neocosmospora vasinfecta* (podredumbre del tallo por *Neocosmospora*), *Phyllosticta sojicola* (podredumbre foliar por *Phyllosticta*), *Pyrenochaeta glycines* (mancha foliar por *Pyrenochaeta*), *Cylindrocladium crotalariae* (podredumbre roja de la corona), *Dactuliochaeta glycines* (mancha roja foliar), *Spaceloma glycines* (Scab), *Stemphylium botryosum* (añublo foliar por *Stemphylium*), *Corynespora cassiicola* (añublo temprano), *Nematospora coryli* (mancha de levadura), y *Phymatotrichum omnivorum* (podredumbre de la raíz del algodón).

V. Definiciones

[0115] Las definiciones que siguen se proporcionan para definir y clarificar el significado de estos términos con referencia a las realizaciones pertinentes de la presente divulgación tal como se usan en la presente y para guiar a aquellos con capacitación normal en la técnica en la comprensión de la presente divulgación. A menos que se indique lo contrario, los términos y expresiones se han de entender de acuerdo con su significado y uso convencionales en la técnica correspondiente, en particular en el campo de la biología molecular y la transformación vegetal.

[0116] Cuando se introducen los elementos de la presente divulgación o una o más de sus realizaciones, se considera que los artículos “un”, “uno”, “una”, “el”, “la” y “dicho/a” significan que existe uno o más de los elementos.

[0117] La expresión “y/o”, cuando se usa en una lista de dos o más ítems, significa cualquiera de los ítems, cualquier combinación de los ítems, o todos los ítems con los cuales está asociada la expresión.

[0118] Las expresiones “que comprende”, “que incluye” y “que tiene” se consideran inclusivas y significan que pueden existir elementos adicionales diferentes de los elementos enumerados. Por ejemplo, cualquier método que “comprende”, “tiene” o “incluye” uno o más pasos no se limita a poseer solo tales uno o más pasos y también puede cubrir otros pasos no enumerados. De modo

similar, cualquier composición o dispositivo que “comprende”, “tiene” o “incluye” una o más características no se limita a poseer solo tales una o más características y también puede cubrir otras características no enumeradas.

[0119] Como se usa en la presente, una “planta” incluye una planta completa, explante, parte de planta, plántula, o plantones en cualquier etapa de regeneración o desarrollo.

[0120] Como se usa en la presente, una “parte de planta” se puede referir a cualquier órgano o tejido intacto de una planta, tal como meristemo, órgano/estructura del brote (por ejemplo, hoja, tallo o nudo), raíz, flor u órgano/estructura floral (por ejemplo, bráctea, sépalo, pétalo, estambre, carpelo, antera y óvulo), semilla, embrión, endosperma, recubrimiento de semilla, fruto, el ovario maduro, propágulo u otros tejidos vegetales (por ejemplo, tejido vascular, tejido dérmico, tejido molido, y similares), o cualquier porción de esta. Las partes de plantas de la presente divulgación pueden ser viables, no viables, regenerables y/o no regenerables. Un “propágulo” puede incluir cualquier parte de planta que puede desarrollarse hasta una planta entera.

[0121] Un “embrión” es una parte de una semilla de una planta que consiste en tejidos precursores (por ejemplo, tejido meristemático) que puede desarrollarse en la totalidad o una parte de una planta adulta. Un “embrión” además puede incluir una porción de un embrión de una planta.

[0122] Un “meristemo” o “tejido meristemático” comprende células no diferenciadas o células meristemáticas, que son capaces de diferenciarse para producir uno o más tipos de partes, tejidos o estructuras de una planta, tales como la totalidad o parte de un brote, tallo, raíz, hoja, semilla, etc.

[0123] Como se usa en la presente, “hábito de crecimiento determinado” se refiere al cese del crecimiento vegetativo después de que el tallo principal termina en un agrupamiento de vainas maduras. Determinadas variedades de soja comienzan la floración cuando se han desarrollado todos o la mayoría de los nudos en el tallo principal. Estas usualmente tienen racimos alargados que pueden tener varios centímetros de longitud y pueden tener un gran número de flores. Como se usa en la presente, “hábito de crecimiento indeterminado” se refiere al desarrollo de hojas y flores simultáneamente a través de una porción de sus períodos reproductivos, con una a tres vainas en el ápice terminal. Las variedades de soja indeterminadas, cuando crecen en su latitud de

adaptación, florecen cuando se han desarrollado aproximadamente la mitad de los nudos en el tallo principal. Tienen racimos cortos con pocas flores, y su nudo terminal tiene solo una pocas flores. Las variedades de soja “semi-determinadas” también florecen cuando se han desarrollado aproximadamente la mitad de los nudos en el tallo principal, pero el desarrollo del nudo y la floración en el tallo principal se detienen más abruptamente que en las variedades indeterminadas. Sus racimos son cortos y tienen pocas flores, a excepción del racimo terminal, el cual puede tener varias veces más flores que aquellos a menor altura en la planta.

[0124] Como se usa en la presente, “encamado” se refiere a la inclinación de los tallos hasta cerca del nivel del suelo en plantas de cultivos de granos. El encamado se clasifica en una escala de 1 a 9. Generalmente, un puntaje de 1 indica plantas erectas. Un puntaje de 5 indica plantas que están inclinadas en un ángulo de 45 grados en relación con el suelo y un puntaje de 9 indica plantas que están tendidas en el suelo. La Tabla 4 muestra la escala de encamado usada en la presente para evaluar las plantas de soja de control y las modificadas.

[0125] Como se usa en la presente, “fecha de floración terminal” se refiere a la fecha en la cual existe una flor abierta en el nudo terminal del tallo principal en el 50% de las plantas en un segmento de una parcela uniforme.

[0126] Como se usa en la presente, la “fase vegetativa” del desarrollo de una planta es el período de crecimiento entre la germinación y la floración. Las etapas en la fase vegetativa de la soja son las siguientes: VE (emergencia), (etapa cotiledónea), V1 (primera hoja trifoliada), V2 (segunda hoja trifoliada), V3 (tercera hoja trifoliada), V(n) (enésima hoja trifoliada) y V6 (inminente inicio de la floración). Como se usa en la presente, la “fase reproductiva” del desarrollo de una planta es el período entre la floración y el final de la cosecha. Las etapas en la fase reproductiva de la soja son las siguientes: R₁ (inicio de la floración, primera flor), R₂ (floración completa, flores en los 2 nudos superiores), R₃ (inicio de formación de vainas, vainas de 4,76 mm (3/16”) en los 4 nudos superiores), R₄ (vainas completamente desarrolladas, vainas de 19,05 mm (3/4”) en los 4 nudos superiores), R₅ (semilla de 3,17 mm (1/8”) en los 4 nudos superiores), R₆ (semillas completamente desarrolladas en los 4 nudos superiores), R₇ (inicio de la maduración, una vaina madura), y R₈ (madurez completa, 95% de las vainas en la planta han alcanzado el color de madurez). Las etapas vegetativas y reproductivas de la soja son ampliamente conocidas por los expertos en la técnica y

se pueden hallar numerosas publicaciones que describen estas etapas, por ejemplo, en Internet, tales como la publicación de la North Dakota State University A-1174, de junio de 1999, revisada y reimpressa en agosto de 2004.

[0127] Todos los métodos descritos en la presente se pueden llevar a cabo en cualquier orden adecuado a menos que en la presente se indique lo contrario o esté claramente en contradicción con el contexto. El uso de cualquiera de los ejemplos, o del lenguaje a modo de ejemplo (por ejemplo, “tal como”) proporcionado con respecto a ciertas realizaciones en la presente, tiene el mero propósito de iluminar mejor la presente divulgación y no impone una limitación al alcance de la presente divulgación a menos que se indique lo contrario.

[0128] Habiendo descrito la presente divulgación en detalle, se observará que son posibles modificaciones, variaciones y realizaciones equivalentes sin apartarse del espíritu ni del alcance de la presente divulgación tal como se define ulteriormente en las reivindicaciones adjuntas. Asimismo, se deberá apreciar que todos los ejemplos en la presente divulgación inclusive los que siguen a continuación se proporcionan como ejemplos no limitantes.

EJEMPLOS

Ejemplo 1. Diseño de constructos para edición de genes.

[0129] El empleo de edición génica para realizar mutagénesis dirigida de loci de determinación ofrece la oportunidad de crear alelos que confieren un rango de fenotipos de determinación, lo cual a su vez resulta en diversidad en la arquitectura y el hábito de crecimiento de las plantas de soja. La expresión endógena específica del alelo dominante del locus *Dt1* de soja, el gen *GmTFL1b*, en las puntas de los tallos protege el meristemo apical de la diferenciación terminal, lo cual lleva a la indeterminación. El nivel de determinación se puede modular al variar el nivel de expresión del gen *GmTFL1b* mediante la modificación de la región promotora del gen *GmTFL1b*. Se proporciona la secuencia codificadora del gen *GmTFL1b* como SEQ ID NO:1, la secuencia de aminoácidos para la proteína TFL1b se proporciona como SEQ ID NO:2, y la secuencia de ADNc para el gen *GmTFL1b*, que incluye 5' UTR y 3'UTR, se proporciona como SEQ ID NO:3.

[0130] Se localizó la región promotora corriente arriba del gen *GmTFL1b* para mutagénesis usando edición génica. Se diseñaron cuatro constructos de edición génica para transformación

vegetal con múltiples ARN guía (ARNg) para apuntar a varias localizaciones dentro de una región promotora de 2 kb (SEQ ID NO:4) corriente arriba del sitio de inicio de la transcripción (tis) del gen *GmTFL1b*, como se ilustra en la FIG. 1.

[0131] Se diseñó cada uno de los constructos de vectores de transformación vegetal para producir roturas de doble hebra (DSB) en múltiples localizaciones con el direccionamiento de los ARNg. Es posible que se produzcan pequeñas deleciones en los sitios DSB, dado que son deleciones de grandes segmentos entre sitios DSB.

[0132] En este ejemplo, cada uno de los constructos para edición genómica generalmente contenían dos o tres casetes o regiones funcionales pertinentes para la edición génica y la creación de las DSB en la región promotora del gen *GmTFL1b*: expresión de una proteína Cpf1, expresión de dos a cuatro ARNg dirigidos a la región promotora del gen *GmTFL1b* y, opcionalmente, expresión de otros cuatro ARNg adicionales dirigidos a la región promotora del gen *GmTFL1b*. Cada ARNg contiene un andamio común compatible con el gen Cpf1 (SEQ ID NO:5), y una exclusiva secuencia espaciadora/de localización complementaria al sitio diana pretendido según se enumera en la Tabla 1. En la Tabla 1 se enumeran las secuencias de ADN que codifican los espaciadores de ARNg y sus sitios diana pretendidos.

Tabla 1. ARN guías a modo de ejemplo usados para editar la región promotora del gen endógeno *GmTFL1b*.

Constructo	Espaciador ARN guía	Sitio diana	Secuencia espaciadora	SEQ ID NO
pM552	g+1774A	pos. 1774 a pos. 1796	TGACCAGCGGAGAGAGAAAGAGA	13
	g-1570C	pos. 1548 a pos. 1570	TCTCCCATCAGCTCGATCAGGTC	14
	g-1389A	pos. 1367 a pos. 1389	GTATACTACTAGTACCATGGACG	15
	g-1237G	pos. 1215 a pos. 1237	TTGGAAGATTCCTGATGAAGAGG	16
	g+1083A	pos. 1083 a pos. 1105	CCAATATTGGAACCTTGGTCTGCA	17
	g+936A	pos. 936 a pos. 958	AGCACGGGATGTGCATGTACACA	18
	g+749C	pos. 749 a pos. 771	ACAGTGAAACGATACCAACCTGC	19
	g-672T	pos. 650 a pos. 672	TTGGTCTGAAACTGTTAGTTGCT	20
pM206	g-1570C	pos. 1548 a pos. 1570	TCTCCCATCAGCTCGATCAGGTC	21

Constructo	Espaciador ARN guía	Sitio diana	Secuencia espaciadora	SEQ ID NO
	g-1389A	pos. 1367 a pos. 1389	GTATACTACTAGTACCATGGACG	22
	g-1237G	pos. 1215 a pos. 1237	TTGGAAGATTCCTGATGAAGAGG	23
pM205	g-1570C	pos. 1548 a pos. 1570	TCTCCCATCAGCTCGATCAGGTC	24
	g-1237G	pos. 1215 a pos. 1237	TTGGAAGATTCCTGATGAAGAGG	25
pM207	g-1389A	pos. 1367 a pos. 1389	GTATACTACTAGTACCATGGACG	26
	g-1237G	pos. 1215 a pos. 1237	TTGGAAGATTCCTGATGAAGAGG	27
pM204	g-1570C	pos. 1548 a pos. 1570	TCTCCCATCAGCTCGATCAGGTC	28
	g-1389A	pos. 1367 a pos. 1389	GTATACTACTAGTACCATGGACG	29

[0133] El casete de expresión de Cpf1 del constructo de edición pM552 comprendía un promotor FLT del virus mosaico de la Dalia (SEQ ID NO:6) operativamente enlazado con una secuencia codificadora de una enzima endonucleasa guiada por ARN Cpf1 de *Lachnospiraceae bacterium* (SEQ ID NO:7) que tenía codones optimizados para arroz, flanqueada en ambos lados por una copia de una señal de localización nuclear (SEQ ID NO:8). Ver, por ejemplo, Gao *et al.*, (*Nature Biotechnol.* 35(8):789-792, 2017). El casete de expresión de Cpf1 en los otros cuatro constructos mostrados en la Tabla 1 comprendían un promotor ubiquitina de *Medicago truncatula* (SEQ ID NO:9) operativamente enlazado con una secuencia con codones optimizados para maíz que codificaba una enzima endonucleasa guiada por ARN Cpf1 de *Lachnospiraceae bacterium* (SEQ ID NO:10), flanqueada en cada lado por una copia de una señal de localización nuclear (SEQ ID NO:8).

[0134] Un tipo de casete de expresión ARNg, presente en todos los constructos, comprendía una secuencia que codificaba de dos a cuatro ARNg operativamente enlazada con un promotor de ARN polimerasa III (Pol3) de soja (SEQ ID NO:11). Secuencias espaciadoras como las que se mencionan en la Tabla 1 se dirigían a sitios de rotura alternativos en la región promotora de *GmTFL1b*. Un tipo adicional de casete de expresión ARNg, que está presente solo en pM552, comprendía una secuencia que codificaba otros cuatro ARN guías operativamente enlazada con

un promotor 7SL_CR10 de soja (SEQ ID NO:12). Secuencias espaciadoras como las que se mencionan en la Tabla 1 se dirigen a sitios de rotura alternativos en la región promotora de *GmTFL1b*.

[0135] Específicamente, constructos para edición génica para transformación vegetal serán diseñados para que comprendan ARNg dirigidos a una región de ADN que abarca desde la posición nucleotídica 1237 hasta la posición nucleotídica 1570 de la secuencia de referencia SEQ ID NO:4.

Ejemplo 2. Confirmación de ediciones en la región promotora de *GmTFL1b* y cigosidad de plantas con edición genómica.

[0136] Se transformó una línea de soja de tipo silvestre endogámica vía transformación mediada por *Agrobacterium* con el vector pM552 descrito en el Ejemplo 1 precedente. El tejido vegetal transformado se cultivó para producir plantas maduras R₀. Las plantas R₀ que tenían una o más ediciones genómicas exclusivas se autocruzaron para producir plantas R₁. Las plantas R₁ que eran homocigotas para alelos que comprendían secuencias promotoras de *GmTFL1b* editadas y carecían de secuencias de ADN-T se autocruzaron para producir plantas R₂.

[0137] Para determinar las ediciones realizadas en la región promotora de *GmTFL1b*, se usó la técnica de secuenciación de amplicones para producir secuencias mutantes para la región promotora de 2 kb para compararla con la secuencia de tipo silvestre. La secuenciación de amplicones supone la generación de uno o más productos de PCR exclusivos en toda la región genómica de interés para secuenciación de próxima generación. Los datos de las secuencias para cada muestra luego se mapean respecto de una secuencia de referencia para identificar diferencias en las secuencias de consenso. Las plantas con delecciones exclusivas dentro del rango de 30 a 1221 pares de bases (pb) de longitud fueron seleccionadas para proporcionar diversa cobertura de las mutaciones en la región promotora apuntada. Plantas R₁ individuales producidas mediante autofecundación de plantas R₀ que tenían una o más de las ediciones fueron ensayadas para determinar la naturaleza de las ediciones y la cigosidad del mutante o alelo editado y se describen en la Tabla 2. Todas las plantas editadas descritas en la Tabla 2 se produjeron usando el vector de transformación del constructo pM552. En la Tabla 2, “Evento R₀” es el identificador de la planta

R₀; “Nombre del alelo” es el identificador para un alelo exclusivo; y Nulo (WT (tipo silvestre por sus siglas en inglés)) corresponde a la planta de soja endogámica sin editar. “Lesión causal” indica las coordenadas de inicio y finalización de un segmento eliminado del promotor, con el extremo 5’ de la secuencia promotora de 2 kb (SEQ ID NO:4) como el punto inicial. “Cigosidad R₁” indica la cigosidad de la planta R₁, como homocigota, heterocigota, no determinada (ND), o letal (para una edición que dio por resultado un fenotipo letal). El alelo *AL430a* contiene tres modificaciones al gen endógeno *GmTFL1b*: dos delecciones (un segmento que comienza en la posición 939 y termina en la posición 1028 de SEQ ID NO:4; y un segmento que comienza en la posición 1066 y termina en la posición 1664 de SEQ ID NO:4) y una mutación por inversión (un segmento que comienza en la posición 1065 y termina en la posición 1029 de SEQ ID NO:4 invertido y reinsertado en la región eliminada).

Tabla 2. Plantas editadas producidas por el constructo de edición pM552, con delección segmentaria del promotor *GmTFL1b*.

Evento R ₀	Nombre del alelo	Delección causal (posiciones nucleotídicas de inicio y terminación de SEQ ID NO:4)	Cigosidad R ₁	SEQ ID NO
S376	Nulo (WT)	N/A	sin ediciones	N/A
S376	<i>AL376a</i>	pos. 1539 a pos. 1568	homocigota	30
S376	<i>AL376b</i>	pos. 1217 a pos. 1604	homocigota	31
S391	<i>AL391a</i>	pos. 1518 a pos. 1629	homocigota	32
S391	<i>AL391b</i>	pos. 951 a pos. 1222	homocigota	33
S398	<i>AL398</i>	pos. 1364 a pos. 1407	homocigota	34
S410	<i>AL410a</i>	pos. 1367 a pos. 1516	ND	35
S411	<i>AL411a</i>	pos. 754 a pos. 1806	homocigota	36
S411	<i>AL411b</i>	pos. 1216 a pos. 1319	homocigota	37
S411	<i>AL411c</i>	pos. 1159 a pos. 1965	heterocigota	38
S419	<i>AL419</i>	pos. 760 a pos. 1214	ND	39

S430	<i>AL430a</i>	pos. 939 a pos. 1028; INV - pos. 1065 a pos. 1029; pos. 1066 a pos. 1664	homocigota	40
S430	<i>AL430b</i>	pos. 952 a pos. 1552; pos. 1677 a pos. 1808	homocigota	41
S437	<i>AL437</i>	pos. 1524 a pos. 1558	homocigota	42
S447	<i>AL447a</i>	pos. 476 a pos. 1405	homocigota	43
S447	<i>AL447b</i>	pos. 1365 a pos. 1559	homocigota	44
S455	<i>AL455</i>	pos. 928 a pos. 1560	homocigota	45
S461	<i>AL461</i>	pos. 346 a pos. 2092	letal	N/A
S466	<i>AL466</i>	pos. 593 a pos. 1813	homocigota	46

*La lesión causal enumerada para cada alelo representa la delección más grande en el alelo, la cual se considera que conduce al fenotipo de determinación alterado observado en las plantas modificadas. En el caso de algunos alelos, también estuvieron presentes pequeñas delecciones adicionales.

[0138] Las ediciones homocigotas en la Tabla 2 precedente se ilustran adicionalmente en la FIG. 2, alineadas con los espaciadores ARNg ilustrados en la FIG. 1.

[0139] De modo similar, las plantas de soja del grupo de madurez 3,5 se usaron para producir plantas editadas empleando los vectores de transformación de los constructos pM205, pM206 y pM207. Plantas R₁ individuales producidas mediante autofecundación de plantas R₀ que tenían una o más de las ediciones fueron ensayadas para determinar la naturaleza de las ediciones y la cigosidad del mutante o alelo editado y se describen en la Tabla 3.

Tabla 3. Plantas de soja editadas del grupo de madurez 3,5 producidas mediante los constructos pM205, pM206 o pM207, con delección segmentaria del promotor *GmTFL1b*.

Constructo de edición pM205				
Evento R ₀	Nombre del alelo	Delección causal (posiciones nucleotídicas)	Cigosidad R ₁	SEQ ID NO

		de inicio y terminación de SEQ ID NO:4)		
S722316	<i>AL316a</i>	pos. 1552 a pos. 1556	homocigota	82
S722316	<i>AL316b</i>	pos. 1537 a pos. 1565	homocigota	83
S722336	<i>AL336a</i>	pos. 1209 a pos. 1246	homocigota	84
S722336	<i>AL336b</i>	pos. 1209 a pos. 1246 pos. 1552 a pos. 1600	heterocigota	85
S722336	<i>AL336c</i>	pos. 1553 a pos. 1601	heterocigota	86
Construto de edición pM207				
Evento R₀	Nombre del alelo	Delección causal (posiciones nucleotídicas de inicio y terminación de SEQ ID NO:4)	Cigosidad R₁	SEQ ID NO
S722544	<i>AL544</i>	pos. 1368 a pos. 1378	homocigota	87
S722589	<i>AL589a</i>	pos. 1368 a pos. 1395	homocigota	88
S722589	<i>AL589b</i>	pos. 1368 a pos. 1378	homocigota	89
S722590	<i>AL590</i>	pos. 1344 a pos. 1384	homocigota	90
S722601	<i>AL601</i>	pos. 1219 a pos. 1453	homocigota	91
S722676	<i>AL676</i>	pos. 1215 a pos. 1227; pos. 1370 a pos. 1376	homocigota	92
S722702	<i>AL702a</i>	pos. 1216 a pos. 1376	homocigota	94
S722702	<i>AL702b</i>	pos. 1368 a pos. 1384	homocigota	95
S722727	<i>AL727a</i>	pos. 1366 a pos. 1648	homocigota	96
S722727	<i>AL727b</i>	pos. 1370 a pos. 1489	heterocigota	97
Construto de edición pM206				
Evento R₀	Nombre del alelo	Delección causal (posiciones nucleotídicas	Cigosidad R₁	SEQ ID NO

		de inicio y terminación de SEQ ID NO:4)		
S716698	<i>AL698</i>	pos. 1220 a pos. 1556	homocigota	93
S716734	<i>AL734</i>	pos. 1370 a pos. 1376; pos. 1541 a pos. 1615	homocigota	98
S716766	<i>AL766a</i>	pos. 1367 a pos. 1375; pos. 1551 a pos. 1601	homocigota	99
S716766	<i>AL766b</i>	pos. 1371 a pos. 1377	homocigota	100
S716787	<i>AL787</i>	pos. 1368 a pos. 1553	homocigota	101
S716797	<i>AL797a</i>	pos. 1365 a pos. 1378	homocigota	102
S716797	<i>AL797b</i>	pos. 1365 a pos. 1380; pos. 1552 a pos. 1594	homocigota	103
S716800	<i>AL800a</i>	pos. 1367 a pos. 1376; pos. 1552 a pos. 1682	homocigota	104
S716800	<i>AL800b</i>	pos. 1367 a pos. 1376	homocigota	105
S716800	<i>AL800c</i>	pos. 954 a pos. 1391	homocigota	106

[0140] De modo similar, las plantas de soja del grupo de madurez 5,9 se usaron para producir plantas editadas empleando los vectores de transformación de los constructos pM204 y pM205. Plantas R₁ individuales producidas mediante autofecundación de plantas R₀ que tenían una o más de las ediciones fueron ensayadas para determinar la naturaleza de las ediciones y la cigosidad del mutante o alelo editado y se describen en la Tabla 4.

Tabla 4. Plantas editadas del grupo de madurez 5,9 producidas mediante los constructos pM204 o pM205, con delección segmentaria del promotor *GmTFL1b*.

Constructo de edición pM205				
Evento R ₀	Nombre del alelo	Delección causal (posiciones nucleotídicas de	Cigosidad R ₁	SEQ ID NO

		inicio y terminación de SEQ ID NO:4)		
S832020	<i>AL020</i>	pos. 1535 a pos. 1560	homocigota	107
S832025	<i>AL025a</i>	pos. 1534 a pos. 1644	homocigota	108
S832025	<i>AL025b</i>	pos. 937 a pos. 1599	homocigota	109
S832025	<i>AL025c</i>	pos. 1551 a pos. 1648	homocigota	110
S832049	<i>AL049a</i>	pos. 1526 a pos. 1559	homocigota	111
S832049	<i>AL049b</i>	pos. 1528 a pos. 1606	homocigota	112
S832126	<i>AL126a</i>	pos. 1542 a pos. 1602	homocigota	113
S832126	<i>AL126b</i>	pos. 1214 a pos. 1594	heterocigota	114
S737665	<i>AL665a</i>	pos. 1541 a pos. 1586	homocigota	121
S737665	<i>AL665b</i>	pos. 1552 a pos. 1556	homocigota	122
S737665	<i>AL665c</i>	pos. 669 a pos. 1623	homocigota	123
S737746	<i>AL746a</i>	pos. 1521 a pos. 1558	homocigota	124
S737746	<i>AL746b</i>	pos. 1369 a pos. 1477; pos. 1540 a pos. 1554	homocigota	125
S831991	<i>AL991a</i>	pos. 1217 a pos. 1674	homocigota	126
S831991	<i>AL991b</i>	pos. 1546 a pos. 1626	homocigota	127
S831991	<i>AL991c</i>	pos. 1547 a pos. 1635	homocigota	128
Constructo de edición pM204				
Evento R₀	Nombre del alelo	Deleción causal (posiciones nucleotídicas de inicio y terminación de SEQ ID NO:4)	Cigosidad R₁	SEQ ID NO
S737331	<i>AL331a</i>	pos. 1368 a pos. 1554	homocigota	115
S737331	<i>AL331b</i>	pos. 1369 a pos. 1477; pos. 1550 a pos. 1554	homocigota	116
S737331	<i>AL331c</i>	pos. 734 a pos. 2000	homocigota	117
S737398	<i>AL398a</i>	pos. 1371 a pos. 1560	homocigota	118

S737398	<i>AL398b</i>	pos. 1370 a pos. 1376	homocigota	119
S737398	<i>AL398c</i>	pos. 1370 a pos. 1376; pos. 1537 a pos. 1572	homocigota	120

Ejemplo 3. Niveles de expresión del gen *GmTFL1b* en plantas de soja modificadas.

[0141] Plantas R₂ que tenían alelos homocigotos de mutantes del promotor *GmTFL1b* (alelos *Dt1*) y plantas de control de tipo silvestre se seleccionaron y cultivaron en un ambiente controlado para muestreo a fin de evaluar diferencias en la expresión del gen diana. Las plantas se muestrearon en la etapa VC, que se define como el período posterior a la etapa de emergencia pero antes de la etapa V1, cuando los cotiledones y las unifoliadas están completamente expandidos. Se recolectaron tejidos de hojas unifoliadas y ápices sobre hielo seco para analizar la expresión génica mediante un ensayo TaqMan. Se midieron diez réplicas biológicas por entrada de alelo.

[0142] La Tabla 5 y la FIG. 3 muestran los niveles de expresión relativos del gen *GmTFL1b* en tejidos foliares y ápices en la etapa VC, de plantas de soja editadas con diversos alelos *Dt1* homocigotos. Las plantas editadas se produjeron usando el constructo de edición pM552. Como se muestra en la FIG. 3, la expresión de *GmTFL1b* en el tejido foliar fue nominal para todas las muestras ensayadas.

Tabla 5. Niveles de expresión relativos del gen *GmTFL1b* en tejidos apicales en la etapa VC en plantas de soja editadas con diversos alelos *Dt1* homocigotos.

Evento R ₀	Nombre del alelo	Lesión causal (posiciones nucleotídicas de inicio y terminación de SEQ ID NO:4)	Nivel de expresión relativo	
			Media	Error STD
S376	Nulo (WT)	N/A	1532,01	180,12
S376	<i>AL376a</i>	pos. 1539 a pos. 1568	785,74	117,73
S376	<i>AL376b</i>	pos. 1217 a pos. 1604	761,69	129,95
S391	<i>AL391a</i>	pos. 1518 a pos. 1629	533,60	78,67
S391	<i>AL391b</i>	pos. 951 a pos. 1222	nd	Nd
S398	<i>AL398</i>	pos. 1364 a pos. 1407	1728,11	313,12

Evento R ₀	Nombre del alelo	Lesión causal (posiciones nucleotídicas de inicio y terminación de SEQ ID NO:4)	Nivel de expresión relativo	
			Media	Error STD
S410	<i>AL410a</i>	pos. 1367 a pos. 1516	nd	Nd
S411	<i>AL411a</i>	pos. 754 a pos. 1806	665,83	114,64
S411	<i>AL411b</i>	pos. 1216 a pos. 1319	195,73	22,74
S411	<i>AL411c</i>	pos. 1159 a pos. 1965	nd	Nd
S419	<i>AL419</i>	pos. 760 a pos. 1214	nd	Nd
S430	<i>AL430a</i>	pos. 939 a pos. 1028; INV - pos. 1065 a pos. 1029 pos. 1066 a pos. 1664	138,85	29,66
S430	<i>AL430b</i>	pos. 952 a pos. 1552; pos. 1677 a pos. 1808	673,69	96,28
S437	<i>AL437</i>	pos. 1524 a pos. 1558	909,13	151,91
S447	<i>AL447a</i>	pos. 476 a pos. 1405	946,29	111,52
S447	<i>AL447b</i>	pos. 1365 a pos. 1559	2.087,50	288,51
S455	<i>AL455</i>	pos. 928 a pos. 1560	873,15	108,95
S461	<i>AL461</i>	pos. 346 a pos. 2092	nd	Nd
S466	<i>AL466</i>	pos. 593 a pos. 1813	650,28	80,61

[0143] Dado que el promotor había sido mutado mediante edición genómica, la mayor parte de los alelos *Dt1* llevaron a un nivel reducido de expresión de *GmTFL1b*, en comparación con plantas del tipo silvestre que tenían el promotor endógeno. Sin embargo, los alelos mutantes *AL398* y *AL447b* exhibieron niveles aumentados de expresión génica. La expresión de TFL1 se redujo en gran medida en todas las muestras a excepción de dos de las muestras ensayadas.

Ejemplo 4. Evaluación de los fenotipos de fecha de floración terminal y fecha de madurez en plantas que comprenden alelos editados.

[0144] Para estudiar los fenotipos de las plantas editadas, plantas R₂ con alelos homocigotos de mutantes del promotor *GmTFL1b* (alelos *Dt1*) y plantas de control de tipo silvestre fueron

seleccionadas y cultivadas en invernadero y en el campo con prácticas agronómicas convencionales.

[0145] Todas las plantas editadas mostraron una emergencia tan buena o mejor que las de tipo silvestre, inclusive las plantas casi determinadas con el alelo *AL430a*. Todas las plantas editadas iniciaron la floración como el control indeterminado de tipo silvestre. No se observaron cambios detectables en el inicio de la floración en las plantas editadas, lo cual indica que los rasgos de hábito de crecimiento del tallo controlados por el locus *Dt1* no interactuaron con los genes que controlan el inicio de la floración.

[0146] Como se observó en los invernaderos, las plantas con alelos *Dt1* editados desplegaron un rango de fenotipos de determinación que abarcaron desde casi determinadas hasta más indeterminadas que el tipo silvestre (indeterminación extrema), con distintas variedades de comportamiento semi-determinado, evidenciado en primer lugar por diferencias en las fechas de floración terminal como se muestra en la FIG. 4. La fecha de floración terminal se define como la fecha en la cual existe una flor abierta en el nudo terminal del tallo principal en el 50% de las plantas en un segmento de una parcela uniforme. En esta evaluación, las parcelas consistieron en 2 filas de 152,4 cm (5') de longitud con una separación entre ellas de 76,2 cm (30"); aproximadamente 80 semillas sembradas por parcela; las parcelas tenían una separación entre ellas de 45,72 cm (1,5').

[0147] El encamado en las parcelas de soja se evaluó a través de observaciones múltiples comenzando en la etapa de desarrollo R₆ usando una escala de calificación para determinar la porción de cada parcela que exhibía encamado de tallos. Las observaciones continuaron hasta la senescencia en las postrimerías de la etapa R₇ del desarrollo. Para calificar las parcelas respecto del encamado se usó la escala en la Tabla 6.

Tabla 6. Determinación del puntaje de encamado en parcelas de soja.

Puntaje	Fenotipo
1	90 a 100% de las plantas están erectas.
2	80 a 89% de las plantas están erectas.
3	70 a 79% de las plantas están erectas.
4	60 a 69% de las plantas están erectas.
5	50 a 59% de las plantas están erectas.
6	40 a 49% de las plantas están erectas.
7	30 a 39% de las plantas están erectas.
8	20 a 29% de las plantas están erectas.
9	0 a 19% de las plantas están erectas.

[0148] En las parcelas de observación se determinó la fecha de floración terminal, la fecha de madurez, y la calificación del encamado de las plantas de tipo silvestre y las plantas con alelos homocigotos editados. La fecha de floración terminal se calculó como el número de días calendario desde el 1° de enero del año del ensayo. La fecha de madurez se definió como la fecha en la cual 95% de las vainas en una parcela habían alcanzado el color de la madurez. Cuando no se observó la fecha de madurez con precisión se proporciona un signo “>” o “<” para indicar que la fecha de madurez es anterior o posterior a la fecha mencionada. Las fechas de floración terminal y las fechas de madurez se determinaron mediante observación secuencial de las parcelas experimentales hasta que todas las parcelas hubieron alcanzado la etapa de desarrollo. La puntuación de las etapas de desarrollo se realizó a nivel de las parcelas sobre un segmento interno predeterminado de 50,8 cm (20”) de la parcela que era uniforme y no estaba dañado (que contenía aproximadamente 15 plantas consecutivas). Se consideró que una parcela había alcanzado la etapa de desarrollo en cuestión cuando $\geq 50\%$ de las plantas en el segmento medido de la parcela estaban en esa etapa. Los cambios en la determinación se pueden cuantificar en término de días desde la plantación hasta la floración

terminal (DOTF). En el campo, con una fecha de plantación dada que es la misma para todas las plantas del ensayo, la fecha de floración terminal y otras características fenológicas se pueden registrar según la fecha del año calendario para un cálculo simplificado de las diferencias surgidas del control. Las observaciones de las plantas en el campo indicaron que los alelos editados de la región promotora del gen *GmTFL1b* resultaron en rangos diversos de fechas de floración terminal que se extienden desde 22 días antes hasta 4 días después de las fechas de las plantas indeterminadas de tipo silvestre, como se muestra en la Tabla 7. Las plantas editadas se produjeron usando el constructo de edición pM552.

Tabla 7. Valores de fecha de floración terminal (TF), fecha de madurez (R₈), y calificación de encamado.

Evento R ₀	Alelo	TF (nro. de días desde el 1 de enero)	R ₈	Calificación del encamado (1-9)	Nivel de expresión relativo (media)
S376	Nulo (WT)	214	265	6	1.532,01
S398	<i>AL398</i>	218	>267	9	1.728,11
S391	<i>AL391b</i>	216	267	6	ND
S447	<i>AL447b</i>	216	267	7	2.087,5
S447	<i>AL447a</i>	212	265	7	946,29
S455	<i>AL455</i>	211	265	3	873,15
S411	<i>AL411a</i>	207	265	3	665,83
S391	<i>AL391a</i>	203	265	1	533,6
S411	<i>AL411b</i>	199	<265	1	195,73
S430	<i>AL430a</i>	192	<265	1	138,85

[0149] Como se muestra en la Tabla 7, los cambios en las fechas de floración terminal no fueron acompañados por retardo alguno en la fecha de madurez R₈ en la mayoría de las plantas editadas, excepto por las plantas que tenían los alelos *AL398*, *AL391b* y *AL447b*, todas las cuales tuvieron una fecha de floración terminal más tardía que las plantas sin editar de tipo silvestre, y por lo tanto más indeterminadas que las plantas de tipo silvestre.

[0150] Es de esperar que las plantas con determinación más fuerte, o con fecha de floración terminal más temprana produzcan características arquitectónicas que hagan a las plantas más resistentes al encamado. Por cierto, como se muestra en la Tabla 7, existe una clara tendencia hacia la reducción del encamado en plantas con fecha de floración terminal más temprana. Se produjeron a través de la edición genómica en la región promotora del gen *GmTFL1b* plantas con un rango de fechas de floración terminal, correspondientes a niveles diversos de semi-determinación lo cual llevó a la reducción o eliminación del riesgo de encamado. El encamado es responsable de la pérdida de valor en la cosecha, particularmente con variedades de soja de elite con mayor altura. La reducción del encamado tiene el potencial de aumentar el valor y el rendimiento de la cosecha de soja. La reducción del encamado o el aumento de la capacidad de la planta de permanecer erecta también puede mejorar el acceso al campo en todos los puntos después del cierre del dosel, brindando beneficios tales como fertilización eficaz y control de plagas.

[0151] El nivel de expresión relativo medio del gen *GmTFL1b* se copió de la Tabla 5 a la Tabla 7. Dado que el gen endógeno *GmTFL1b* en las sojas indeterminadas es el alelo dominante para el rasgo de indeterminación, la reducción en el nivel de expresión del gen *GmTFL1b* llevaría a la reducción del nivel de indeterminación (o un aumento de la determinación). Como se muestra en la Tabla 7, los niveles de expresión relativos medios tienden a corresponder a la fecha de floración terminal más temprana, o a una determinación más fuerte. Esta correspondencia indica que el nivel de expresión del gen *GmTFL1b* se puede usar como predictor de la fecha de floración terminal o el nivel de determinación de plantas editadas.

[0152] Las diferencias en el nivel de expresión relativo no solo se correlacionan con mutantes de pérdida de función sino también con ganancia de función en la indeterminación. Los mutantes de ganancia de función exhiben indeterminación extrema en comparación con plantas de tipo silvestre, como así también expresión aumentada consistente con la creación de alelos hipermórficos a través de la edición genómica de la región promotora del gen *GmTFL1b*.

Ejemplo 5. Estimación de rendimiento potencial derivado del número de vainas por planta.

[0153] Como se muestra en el Ejemplo 4, las plantas de soja editadas con un fenotipo de fecha temprana de floración terminal tendieron a tener características mejoradas de encamado y posibilidad de mantenerse erguida. No obstante, una fecha temprana de floración terminal también

podría resultar en pérdida de rendimiento debido a la terminación prematura del tallo principal y las ramas, reduciendo el número de nudos productivos en la planta en tal medida que cualquier mejora en la eficacia de los nudos lograda por la determinación temprana de la dominancia apical y la completa dedicación de la fuente al desarrollo reproductivo no son capaces de compensar. Cabe destacar que se puede sostener una fecha de floración terminal significativamente más temprana que la del tipo silvestre con retención de la productividad determinada del nudo apical sin cambiar sustancialmente la altura de la planta (y probablemente el número total de nudos). Esto significa que es posible no reducir el rendimiento de vainas por planta por la introducción del rasgo de semi-determinación, al tiempo que se aprovechan los beneficios del aumento en las posibilidades de la planta de mantenerse erecta.

[0154] El número de vainas por planta se determinó al nivel de parcelas sobre un segmento interno predeterminado de 50,8 cm (20”) de la parcela que era uniforme y no estaba dañado (que contenía aproximadamente 15 plantas consecutivas). En las etapas R6 o más adelante, se cortaron todas las plantas del segmento predeterminado de cada parcela y se empacaron para asegurar que no se perdiera ninguna vaina. Se contaron y registraron sistemáticamente todas las vainas de cada paquete. El número de vainas por entrada (alelo) se informó como la media de todas las plantas contadas por entrada. Las plantas que comprendían los alelos editados *AL391a* y *AL411a*, que demostraron características óptimas de encamado (ver Tabla 7), y plantas indeterminadas de tipo silvestre se evaluaron para determinar el número de vainas por planta y los resultados preliminares se muestran en la Tabla 8 a continuación.

Tabla 8. Valores del número de vainas por planta para plantas de tipo silvestre y plantas editadas con óptimas características de encamado.

ID evento R₀	Alelo	Media vainas/planta	Error STD
S376	Nulo (WT)	45,6	18,7
S391	<i>AL391a</i>	64,5	16,1
S411	<i>AL411a</i>	59,4	20

[0155] Las plantas editadas se produjeron usando el constructo de edición pM552. El número de vainas por planta es comparable entre los 3 grupos de plantas, dentro del margen de error. Las plantas editadas semi-determinadas demostraron aumento en la erectabilidad sin pérdida de vainas. De hecho, los recuentos medios de vainas para plantas editadas semi-determinadas son mayores que para las plantas indeterminadas de tipo silvestre, lo que sugiere que la introducción de semi-determinación puede resultar en el aumento de la productividad junto con la reducción de la pérdida de rendimiento a través de la reducción del encamado.

[0156] Dado que el locus *Dt1* se fijó tempranamente durante la domesticación de la soja, estos alelos editados se pueden recapitular en cualquier germoplasma de soja indeterminado, confiriendo semi-determinación los rasgos resultantes en cualquier línea cultivada. Se puede crear un valor significativo mediante el aumento de la erectabilidad y de la productividad en acres de soja, particularmente en los ambientes de alto rendimiento. Además, el tipo de planta más compacto que resulta de los rasgos semi-determinados puede facilitar el cultivo en filas angostas e inclusive en densidades más altas. Con estas configuraciones agronómicas intensivas, el tipo de planta erecta, compacta, permitirá mejorar el flujo de aire debajo del dosel, con lo cual se reducirá la humedad y con esto, aumentará el potencial para la reducción de la susceptibilidad a enfermedades fúngicas, inclusive reducción de la susceptibilidad al moho blanco y al síndrome de muerte súbita (SDS) ambos con importancia económica.

Ejemplo 6. Edición génica de homólogos de TLF1.

[0157] Se seleccionó un subconjunto de especies de plantas con relevancia agrícola para la comparación de secuencias sobre la base del valor total cosechado y la diversidad, incluyendo miembros de las familias *Solanaceae*, *Brassicaceae* y *Leguminae*, más algodón y cereales monocotiledóneos. Para identificar el acierto de la secuencia polipeptídica de cada especie con la identidad más alta se usó el programa BLASTP del NCBI (Centro Nacional de Información Biotecnológica de los EE.UU., por sus siglas en inglés) con la secuencia polipeptídica TFL1b (*Dt1*) de soja como la secuencia de consulta y la base de datos de la colección de secuencias no redundantes como el conjunto de búsqueda. Las secuencias identificadas se alinearon mediante el algoritmo tipo CLUSTAL en CLC Workbench (parámetros por defecto para alineamiento muy preciso). Las comparaciones de a pares se proporcionan en forma de tabla (FIG. 5).

[0158] Las secuencias de polinucleótidos de la región promotora de 2 kb corriente arriba del sitio de inicio de la transcripción (tis) de los homólogos del gen *TFL1b* hallados en *Zea mays*, *Sorghum bicolor*, *Oryza sativa*, *Triticum aestivum*, *Hordeum vulgare*, *Solanum lycopersicum*, *Gossypium hirsutum*, *Capsicum annuum*, *Brassica napus*, *Arabidopsis thaliana*, *Medicago truncatula*, *Arachis hypogaea* y *Cicer arietinum* se proporcionan en la presente como SEQ ID NO:67-77, 79 y 81, respectivamente. Para crear plantas que tengan fenotipos de determinación modificados en estas especies de plantas adicionales o en otras especies de plantas, se puede usar un procedimiento similar al descrito anteriormente en el Ejemplo 1.

[0159] Siguiendo un proceso similar al descrito anteriormente en el Ejemplo 1, se pueden diseñar aproximadamente diez ARN guía para edición en una separación aproximadamente igual en la región promotora de aproximadamente 2 kb del homólogo *TFL1b* (ver ilustración en FIG. 1), siguiendo las restricciones comunes del diseño de ARNg. La región promotora basal de aproximadamente 200 nucleótidos localizados en el extremo 3' de las secuencias de SEQ ID NO:67-77, 79 y 81 se puede evitar para minimizar las posibilidades de que el promotor deje de ser completamente funcional. Luego del diseño de los ARNg, se pueden seguir pasos en línea con los descritos en el Ejemplo 1 para crear constructos de edición para producir plantas editadas, las cuales ulteriormente se pueden someter a caracterización molecular y fenotípica para seleccionar el nivel deseado de expresión relativa y modulación de la determinación.

Ejemplo 7. Evaluación de características fenotípicas adicionales de plantas que comprenden alelos *Dt1* editados.

[0160] Se evaluaron las características fenotípicas adicionales de plantas de soja que comprendían alelos *Dt1* editados. Todos los efectos sobre rasgos descritos en el ejemplo se informan como delta del valor medio para una entrada dada versus las plantas indeterminadas de tipo silvestre (plantas de soja no modificadas del grupo de madurez 3,5).

[0161] Las parcelas contenían 4 filas que tenían 304,8 cm (10') de longitud separadas entre sí por un espacio de 50,8 cm (20"). Cada rango de parcelas estaba separado por una calle de 91,4 cm (3'). Se sembró un total de 288 semillas por parcela. Cada entrada del ensayo se replicó en 12 parcelas por experimento, con mapeo aleatorizado en cada conjunto de replicación de entrada de ensayo que incluía suficientes parcelas que contenían plantas para comparación de tipo silvestre.

Para cada parcela en el campo se recolectaron mediciones de rasgos que se describen a continuación. Los valores de rasgos obtenidos reportados son la delta de las medias, la diferencia entre la media de ensayo y la media de comparación, para el total de 12 réplicas. En este diseño experimental un valor $P < 0,2$ se considera con significación estadística. La puntuación de las etapas de desarrollo se realizó a nivel de las parcelas sobre un segmento interno predeterminado de 50,8 cm (20 pulgadas) de la parcela que era uniforme y no estaba dañado (que contenía aproximadamente 15 plantas consecutivas). Se consideró que una parcela había alcanzado una etapa de desarrollo específica cuando $\geq 50\%$ de las plantas en el segmento medido de la parcela estaban en esa etapa. En el caso de comparación cuantitativa a gran escala, la determinación de la etapa se informa en términos del número de días desde la plantación hasta la etapa deseada.

[0162] Se evaluaron en la etapa de desarrollo R_8 los rasgos de desarrollo y morfología vegetal, inclusive la fecha de madurez completa, número de ramas por planta, altura de la planta, y encamado de la planta (posibilidad de mantenerse erecta). La madurez completa se determinó sobre la base del número de días desde la fecha de plantación hasta la fecha en la cual el 95% de las vainas en una planta habían alcanzado su color de vaina madura. Se registró la altura de la planta como la medición directa de la longitud del principal tallo lineal. Se evaluó el encamado de la planta como se describe en el Ejemplo 4 precedente. Los resultados se brindan en la tabla 9 a continuación.

Tabla 9. Datos de campo del año 2 sobre el desarrollo y la morfología vegetal en la etapa de crecimiento R_8 de plantas que comprenden alelos *Dt1* homocigotos editados.

Alelo	Madurez completa (media WT = 123 Días)		Nro. de ramas por planta (media WT = 2,1)		Altura de la planta (media WT = 96,3 cm (37,9 pulgadas))		Encamado de la planta (media WT = 5,8)	
	Δ media (Días)	Valor P	Δ media (Conteo)	Valor P	Δ media (cm (pulgadas))	Valor P	Δ media (Calificación)	Valor P
<i>AL391a</i>	-0,9	0,08	0,2	0,55	-7,79 (-19,79)	0,00	-4,7	0,00
<i>AL430a</i>	0,0	0,94	4,3	0,00	-29,47	0,00	-4,4	0,00

					(-74,85)			
<i>AL411b</i>	-0,4	0,41	0,4	0,08	-20,51 (-52,10)	0,00	-4,3	0,00
<i>AL376a</i>	-0,5	0,39	0,0	0,87	-6,65 (-16,89)	0,00	-4,2	0,00
<i>AL437</i>	-1,1	0,07	-0,8	0,00	-6,52 (-16,56)	0,00	-3,8	0,00
<i>AL376b</i>	-0,7	0,33	0,1	0,74	-8,22 (-20,88)	0,00	-3,8	0,00
<i>AL411a</i>	-0,6	0,23	0,0	0,88	-5,06 (-12,85)	0,00	-3,7	0,00
<i>AL430b</i>	-0,5	0,38	-0,2	0,31	-3,71 (-9,42)	0,03	-2,9	0,00
<i>AL466</i>	-0,8	0,18	-0,4	0,16	-4,25 (-10,79)	0,02	-2,7	0,00
<i>AL455</i>	-0,3	0,49	-0,2	0,37	-2,10 (-5,33)	0,12	-2,0	0,00
<i>AL447b</i>	-0,7	0,13	0,1	0,65	-2,97 (-7,54)	0,04	-1,4	0,01
<i>AL391b</i>	-0,4	0,42	-0,1	0,60	0,09 (0,23)	0,94	-1,0	0,07
<i>AL447a</i>	1,0	0,05	0,3	0,12	1,05 (2,67)	0,40	1,6	0,00
<i>AL398</i>	3,0	0,00	0,1	0,67	3,18 (8,08)	0,03	1,9	0,00

***La información para todos los alelos está ordenada por valores de encamado de las plantas, del más bajo al más alto.**

[0163] Como se muestra en la Tabla 7 del Ejemplo 4 precedente, las plantas que portan alelos indeterminados extremos (que exhiben niveles de expresión aumentados del gen *GmTFL1b*), tal como *AL398*, tuvieron terminación retardada junto con madurez significativamente retardada. No obstante, como se muestra en la Tabla 9 precedente, la terminación retardada en plantas que comprendían alelos *AL398* fue acompañada por un número modestamente superior de ramas, altura de la planta aumentada, y aumento del encamado. Todas las plantas que portaban alelos semi-determinados maduraron dentro de 1 día de las plantas indeterminadas de tipo silvestre, con

excepción de las plantas que portaban alelos *AL437*, que maduraron más de un día más temprano que las plantas indeterminadas de tipo silvestre. Las plantas que contenían alelos semi-determinados tuvieron significativas reducciones en la altura de la planta, dentro del rango de 5,08 cm (2 pulgadas) a 76,2 cm (30 pulgadas), en comparación con las plantas indeterminadas de tipo silvestre. Las plantas que contenían alelos semi-determinados tuvieron mayor erectabilidad, exhibiendo calificaciones de encamado de 1-4,7 puntos (10% - 47%) inferiores a los de las plantas indeterminadas de tipo silvestre. Esto se debió, en parte, a la reducción de la altura de la planta. A pesar de las reducciones en la altura y el encamado, la mayoría de los alelos semi-determinados permitió a las plantas portar un número de ramas por planta que no fue estadísticamente diferente al de las plantas indeterminadas de tipo silvestre. Solo las plantas que portaban el alelo *AL437* tuvieron un número de ramas sustancialmente reducido.

[0164] También se evaluaron componentes de rendimiento secundario que incluyeron número de nudos por planta, número de vainas por planta, y vainas por nudo. Los componentes de rendimiento secundario se determinaron a nivel de las parcelas sobre un segmento interno predeterminado de 50,8 cm (20 pulgadas) de la parcela que era uniforme y no estaba dañado (que contenía aproximadamente 15 plantas consecutivas). En la etapa R_8 se cortaron todas las plantas del segmento predeterminado de cada parcela y se empacaron para asegurar que no se perdiera ninguna vaina. Se contó y registró sistemáticamente el número de vainas y el número de nudos de cada planta. Se registró el número de plantas y se combinaron las vainas de una parcela dada para un análisis de componentes de rendimiento primario. Se calculó el número de vainas por nudo a partir de los valores contados. Los resultados se muestran en la Tabla 10 a continuación.

Tabla 10. Datos de campo del año 2 de componentes de rendimiento secundario en la etapa de crecimiento R_8 para plantas que comprenden alelos homocigotos *Dt1* editados.

Alelo	Nudos por planta (media WT = 34,8)		Vainas por nudo (media WT = 2,0)		Vainas por planta (media WT = 68,0)	
	Δ media (Conteo)	Valor P	Δ media (Conteo)	Valor P	Δ media (Conteo)	Valor P
<i>AL391a</i>	1,9	0,42	0,30	0,05	14,1	0,01
<i>AL430a</i>	-1,5	0,48	-0,65	0,00	-22,8	0,00

<i>AL411b</i>	-2,2	0,22	-0,14	0,18	-8,9	0,03
<i>AL376a</i>	-0,9	0,71	0,30	0,06	8,5	0,14
<i>AL437</i>	-0,7	0,75	0,04	0,75	0,4	0,94
<i>AL376b</i>	-2,4	0,52	0,86	0,00	23,0	0,01
<i>AL411a</i>	-0,3	0,86	0,32	0,01	10,4	0,01
<i>AL430b</i>	-2,8	0,18	0,12	0,30	-1,0	0,84
<i>AL466</i>	-3,3	0,13	0,13	0,29	-0,6	0,90
<i>AL455</i>	-1,0	0,61	0,32	0,02	9,3	0,03
<i>AL447b</i>	0,7	0,68	-0,05	0,57	0,2	0,96
<i>AL391b</i>	0,8	0,67	-0,11	0,29	-1,1	0,80
<i>AL447a</i>	3,8	0,03	-0,29	0,01	-3,1	0,42
<i>AL398</i>	2,0	0,25	0,03	0,73	4,8	0,23

***La información para todos los alelos está ordenada por valores de encamado de las plantas, del más bajo al más alto.**

[0165] Las plantas que portaban alelos semi-determinados que tuvieron terminación significativamente temprana en el campo también exhibieron una erectabilidad significativamente mejorada en el campo sin pérdida en el número de ramas y sin cambios en la madurez mayores a un día. Estas características sugieren que los alelos semi-determinados pueden mejorar la erectabilidad de las plantas sin disminución de la productividad. No obstante, la terminación más temprana de las puntas del tallo también podría resultar en pérdida de rendimiento al reducir el número de nudos productivos en la planta en tal medida que cualquier mejora en la eficacia de los nudos lograda por la determinación temprana de la dominancia apical y la completa dedicación de la fuente al desarrollo reproductivo no son capaces de compensar. Aunque las plantas que portan los alelos semi-determinados sí exhiben algún grado de reducción en la altura de la planta, según el alelo, estas reducciones generalmente no están asociadas con cambios significativos en los nudos productivos por planta. De manera consistente con la retención de nudos, la mayoría de los alelos semi-determinados confirieron números de vainas por planta que mantuvieron una paridad o fueron significativamente superiores a los números de vaina por planta contados para las plantas indeterminadas de tipo silvestre. La cantidad de vainas por planta en la mayoría de las plantas semi-determinadas ensayadas estuvo comprendida dentro del rango de equivalente hasta superior

en 23 vainas a las plantas indeterminadas de tipo silvestre. El número de vainas aumentado en las plantas semi-determinadas estuvo asociado con el mejoramiento de la eficacia de los nudos, indicada por números significativamente superiores de vainas por nudo. Las excepciones fueron los alelos *AL430a* y *AL411b*, cuyas mediciones de rasgos de vainas fueron significativamente negativas y que fueron el resultado de una extrema reducción en la altura de la planta debido a la magnitud de la terminación temprana y la pérdida casi de la función de *Dt1* como se mostró en la Tabla 8 anteriormente.

[0166] También se evaluaron los componentes de rendimiento primario, inclusive semillas por planta, peso de mil semillas, y rendimiento estimado del grano. Se determinaron los componentes de rendimiento primario a nivel de parcelas trillando las vainas recolectadas de todas las plantas en una parcela dada y normalizando por el número de plantas medido. Con el trillado se obtuvo el peso total de semillas y el número de semillas. El peso de mil semillas por parcela se calculó dividiendo el peso total de las semillas por el número de semillas y multiplicando por 1000. La estimación del rendimiento del grano es una función de las semillas estimadas por área unitaria multiplicadas por el peso de una sola semilla multiplicado por el factor de conversión 45,375, obteniéndose un estimado tan solo a efectos comparativos, en términos de un bushel por acre de 60 libras. Los resultados se muestran en la Tabla 11 a continuación.

Tabla 11. Datos de campo del año 2 de componentes de rendimiento primario en la etapa de crecimiento R₈ para plantas que comprendían alelos homocigotos *Dt1* editados.

Alelo	Semillas por planta (media WT = 158,9)		Peso de mil semillas (media WT = 170,66 g (6.02 onzas))		Estimación del rendimiento del grano (media WT = 6.704,9 kg/ha (99,7 bu/ac))	
	Δ media (Conteo)	Valor P	Δ media (g/1000 (oz /1000))	Valor P	Δ media (kg/ha (bu/ac))	Valor P
<i>AL391a</i>	12,6	0,33	-15,02 (-0,53)	0,00	18,83 (0,28)	0,97
<i>AL430a</i>	-64,6	0,00	1,13 (0,04)	0,78	-2.822,53 (-41,97)	0,00

<i>AL411b</i>	-38,3	0,00	-2,83 (-0,10)	0,34	-1.304,67 (-19,40)	0,00
<i>AL376a</i>	10,7	0,46	-16,73 (-0,59)	0,00	-0,6725 (-0,01)	1,00
<i>AL437</i>	-12,9	0,27	-28,07 (-0,99)	0,00	-1.262,30 (-18,77)	0,01
<i>AL376b</i>	26,5	0,19	-17,58 (-0,62)	0,01	-92,81 (-1,38)	0,91
<i>AL411a</i>	15,4	0,12	-9,64 (-0,34)	0,00	510,44 (7,59)	0,23
<i>AL430b</i>	-15,4	0,17	-9,35 (-0,33)	0,01	-333,56 (-4,96)	0,49
<i>AL466</i>	-8,6	0,46	-12,19 (-0,43)	0,00	-761,95 (-11,33)	0,13
<i>AL455</i>	14,3	0,18	-9,64 (-0,34)	0,00	266,99 (3,97)	0,55
<i>AL447b</i>	-1,6	0,87	-7,09 (-0,25)	0,01	-156,69 (-2,33)	0,69
<i>AL391b</i>	-0,2	0,98	4,25 (0,15)	0,17	219,24 (3,26)	0,60
<i>AL447a</i>	-6,3	0,50	3,68 (0,13)	0,21	157,37 (2,34)	0,69
<i>AL398</i>	2,2	0,82	-22,68 (-0,80)	0,00	-608,62 (-9,05)	0,13

***La información para todos los alelos está ordenada por valores de encamado de las plantas, del más bajo al más alto.**

[0167] Los componentes de rendimiento secundario observados para la mayoría de los alelos semi-determinados descritos en la presente estaban en paridad o eran significativamente positivos en comparación con las plantas indeterminadas de tipo silvestre. Estas características indican que mediante la edición se pueden lograr mejoras significativas en la erectabilidad de la soja sin perder rendimiento. Los alelos con componentes de rendimiento secundario positivos también exhibieron en muchos casos aumento de semillas por planta en el campo. La mayoría de los alelos semi-

determinados mostró alguna reducción en el peso de las semillas en estos experimentos. No obstante, a pesar de la reducción en el peso de las semillas, todas las plantas que portaban alelos semi-determinados con número de semillas positivo aún mostraron estimaciones de rendimiento de granos de neutras a positivas. Los alelos *AL411b* y *AL430a*, de los cuales se notó previamente que tenían mediciones de rasgos de vainas significativamente negativas, resultado de una extrema reducción en la altura de las plantas, también mostraron reducciones en semillas por planta y estimación de rendimiento del grano. Esto indica el límite útil en la magnitud del rasgo a la reducción de la terminación de la punta del tallo. Estas observaciones proporcionan respaldo sustancial al potencial uso de alelos que confieren semi-determinación de la magnitud apropiada para aumentar la erectabilidad como así también el potencial de rendimiento en algunos casos.

[0168] Se evaluaron los siguientes rasgos encima del nivel del suelo: calificación de la cobertura del dosel, número de días desde la plantación hasta la etapa R₁ (comienzo de la floración), y número de días desde la plantación hasta la vaina terminal (TP) La calificación de la cobertura del dosel representa el porcentaje de un área de la parcela que está ocupada por vegetación. La evaluación se realiza sobre la base de datos de imágenes remotas RGB (rojo, verde, azul, por sus siglas en inglés) logradas por vehículos aéreos no tripulados que distinguen píxeles verdes de píxeles que no son verdes en una parcela, vistos directamente desde arriba de la parcela. Por ejemplo, una calificación de cobertura de dosel de 40% indica que 60% del espacio de la parcela plantada probablemente está desocupado. La terminación se midió mediante observación de la fecha de floración terminal o de formación de vainas terminales. Se determinaron las duraciones de la etapa R₁ y la etapa de vaina terminal mediante observación secuencial de las parcelas experimentales hasta que todas las parcelas habían alcanzado las etapas de desarrollo mencionadas. El inicio de la floración se midió notando la fecha en la cual existió una flor abierta en cualquier nudo en el tallo principal de la planta. La vaina terminal se midió notando la fecha en la cual existió una vaina con un largo de $\geq \frac{3}{16}$ pulgadas (0,5 cm) en el nudo terminal del tallo principal. Los resultados se muestran en la Tabla 12 a continuación.

Tabla 12. Datos de campo del año 3 sobre rasgos por encima del nivel del suelo para plantas que comprenden alelos homocigotos *Dt1* editados.

Alelo	Calificación de la cobertura del dosel (media WT = 33,93)		Nro. de días - Plantación hasta R ₁ (media WT = 35,6 Días)		Nro. de días - Plantación hasta TP (media WT = 69,1 Días)	
	Δ media (%)	Valor P	Δ media (Días)	Valor P	Δ media (Días)	Valor P
<i>AL391a</i>	0,17	0,79	-0,8	0,01	-12,00	0,00
<i>AL411a</i>	0,62	0,32	-1,0	0,00	-9,75	0,00
<i>AL411b</i>	-1,11	0,08	-1,0	0,00	-12,08	0,00
<i>AL376b</i>	-0,55	0,38	-0,3	0,31	-12,08	0,00
<i>AL410a</i>	0,23	0,70	-1,1	0,00	-9,17	0,00
<i>AL430b</i>	0,93	0,14	-1,0	0,00	-8,67	0,00
<i>AL437</i>	-2,39	0,00	-2,3	0,00	-9,25	0,00
<i>AL455</i>	0,91	0,15	-0,6	0,04	-3,50	0,00

*La información para todos los alelos está ordenada por valores de encamado de las plantas, del más bajo al más alto.

[0169] El rasgo primario impactado por las ediciones en el locus *Dt1* es la fecha de terminación. El espectro de cambios en los rasgos de terminación obtenidos por la edición del locus *Dt1* produjo cambios proporcionales corriente abajo en la altura y erectabilidad de la planta. Reducciones significativas en el número de días desde la plantación hasta la vaina terminal para alelos semi-determinados, dentro del rango de 3,5 a 12 días antes que las plantas indeterminadas de tipo silvestre, están asociadas con cambios mínimos en el inicio de la floración y la madurez. Las plantas portadoras del alelo *AL437* iniciaron la floración en el lapso de 1 día después de las plantas indeterminadas de tipo silvestre. La edición de *Dt1* creó cambios con significación estadística en la terminación de diferentes magnitudes según el alelo semi-determinado. Los rasgos por encima del nivel del suelo para la mayoría de las plantas editadas semi-determinadas medidas no presentaron diferencias significativas con las plantas indeterminadas de tipo silvestre en la etapa de crecimiento R₁.

[0170] También se midieron en la etapa R₁ los siguientes rasgos por debajo del nivel del suelo: peso de la raíz seca, peso de los brotes secos, y la proporción entre brotes y raíz en relación con la biomasa. Estos rasgos se evaluaron a nivel de las parcelas sobre un segmento interno

predeterminado de 50,8 cm (20 pulgadas) de la parcela que era uniforme y no estaba dañado (que contenía aproximadamente 15 plantas consecutivas). En la etapa R₁, todas las plantas del segmento predeterminado de cada parcela fueron extraídas del suelo, se contaron y se dividieron en fracciones de brotes y raíz y se empacaron para su transferencia al horno. Las muestras de las plantas se secaron por completo y luego se registró el peso por parcela. La biomasa de raíz y brotes se informó sobre la base de la media por planta. La relación entre brotes y raíz se calculó dividiendo el peso medio de brotes secos por el peso medio de raíces secas para cada entrada. Los resultados se muestran en la Tabla 13 a continuación.

Tabla 13. Datos de campo del año 3 sobre rasgos por debajo del nivel del suelo en la etapa R₁ de crecimiento para plantas que comprendían alelos homocigotos *Dt1* editados.

Alelo	Peso raíz seca (media WT = 3,40 g (0,12 onzas))		Peso brote seco (media WT = 7,93 g (0,28 onzas))		Relación brote:raíz (media WT = 2,34)	
	Δ media (g (oz))	Valor P	Δ media (g (oz))	Valor P	Δ media (relación)	Valor P
<i>AL391a</i>	0,142 (0,005)	0,34	0,113 (0,004)	0,84	-0,06	0,64
<i>AL411a</i>	0,113 (0,004)	0,48	0,652 (0,023)	0,20	0,13	0,33
<i>AL411b</i>	-0,142 (-0,005)	0,34	0,595 (0,021)	0,25	0,29	0,04
<i>AL376b</i>	-0,142 (-0,005)	0,29	-0,113 (-0,004)	0,83	0,15	0,31
<i>AL410a</i>	-0,142 (-0,005)	0,30	-0,198 (-0,007)	0,69	0,07	0,62
<i>AL430b</i>	0,085 (0,003)	0,60	0,170 (0,006)	0,72	-0,02	0,88
<i>AL437</i>	0,198 (-0,007)	0,16	-1,559 (-0,055)	0,00	-0,33	0,02
<i>AL455</i>	0,142 (0,005)	0,30	0,482 (0,017)	0,35	0,06	0,70

***La información para todos los alelos está ordenada por valores de encamado de las plantas, del más bajo al más alto.**

[0171] Las mediciones de biomasa por debajo del nivel del suelo y las relaciones para la mayoría de las plantas portadoras de alelos semi-determinados no presentaron diferencias significativas con las plantas indeterminadas de tipo silvestre en la etapa de crecimiento R₁. Uno de los alelos semi-determinados, *AL437*, resultó una excepción y fue significativamente negativo para las tres mediciones de biomasa, indicando un tamaño de planta en general más pequeño con un brote desproporcionadamente pequeño. Esto es notorio a partir de una calificación de la cobertura de dosel significativamente más baja para este alelo, como se muestra en la Tabla 12.

[0172] En general, estos rasgos indican la utilidad de la edición para la semi-determinación para así obtener alelos exclusivos que son la base para variedades de soja más erectas con potencial de rendimiento estable o mejorado. Además, los resultados muestran que la edición del locus *Dt1* no solo puede generar alelos semi-determinados, sino también alelos casi-determinados, determinados y extremadamente indeterminados.

[0173] Habiendo descripto la presente divulgación en detalle, se observará que son posibles modificaciones, variaciones y realizaciones equivalentes sin apartarse del espíritu ni del alcance de la presente divulgación tal como se define ulteriormente en las reivindicaciones adjuntas. Asimismo, se deberá apreciar que todos los ejemplos en la presente divulgación se proporcionan como ejemplos no limitantes.