(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
06 February 2020 (06.02.2020)

WIPO | PCT

(10) International Publication Number
WO 2020/028307 A1

(51) International Patent Classification:
A01H 5/00 (2018.01) *C12N 15/82* (2006.01)
A01H 9/00 (2006.01) *C12N 15/00* (2006.01)
C12N 15/87 (2006.01) *C07H 21/04* (2006.01)

(21) International Application Number:
PCT/US2019/044048

(22) International Filing Date:
30 July 2019 (30.07.2019)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/711,651 30 July 2018 (30.07.2018) US

(71) Applicant: **UT-BATTELLE, LLC** [US/US]; One Bethel Valley Road, Oak Ridge, Tennessee 37831-6528 (US).

(72) Inventor; and
(71) Applicant: **MUCHERO, Wellington** [US/US]; 122 Hardinberry Street, Oak Ridge, Tennessee 37830 (US).

(72) Inventors: **TUSKAN, Gerald A.**; P.O. Box 5004, Oak Ridge, Tennessee 37831 (US). **CHEN, Jin-Gui**; 118 Hardinberry Street, Oak Ridge, Tennessee 37830 (US). **GUNTER, Lee E.**; 109 Cedar Lane, Oak Ridge, Tennessee 37830 (US).

(74) Agent: **GROLZ, Edward W.**; Scully, Scott, Murphy & Presser, 400 Garden City Plaza, Suite 300, Garden City, New York 11530 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report (Art. 21(3))*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*
— *with sequence listing part of description (Rule 5.2(a))*

(54) **Title:** METHODS FOR IMPROVING CALLUS FORMATION AND REGENERATION IN PLANTS

(57) **Abstract:** This disclosure provides methods of improving callus formation in plants. Specifically, the methods comprising altering in a plant, plant cell or plant tissue the expression of a selected gene, wherein the alteration comprises inactivation of the selected gene in the plant, plant cell or plant tissue. Further disclosed are selected gene sequences. The disclosure further provides genetically engineered plants with improved callus formation.

WO 2020/028307 A1

Methods for Improving Callus Formation and Regeneration in Plants

CROSS REFERENCE TO RELATED APPLICATION

[0001] This application claims the benefit of priority from U.S. Provisional Application No. 62/711,651, filed July 30, 2018, the entire contents of which are incorporated herein by reference.

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT

[0002] This disclosure was made with government support under a research project supported by Prime Contract No. DE-AC05-00OR22725 awarded by the U.S. Department of Energy. The government has certain rights in this invention.

INCORPORATION BY REFERENCE OF SEQUENCE LISTING

[0003] The Sequence Listing in the ASCII text file, named as 37200PCT_4190.2_SequenceListing.txt of 50 KB, created on July 22, 2019, and submitted to the United States Patent and Trademark Office via EFS-Web, is incorporated herein by reference.

BACKGROUND

[0004] Callus arises in plants through the cellular reprogramming of parenchyma cells (Braun AC., *Annual Review of Plant Physiology*. 1954; 5: 133-62), leading to a disorganized amorphic mass of rapidly dividing cells. Callus induction is triggered by variations in endogenous plant hormone levels that occur in response to physical or chemical stimuli (Skoog F. *Am J Bot*. 1944; 30: 19-24; Ikeuchi M. et al., *Plant Cell*.2013; 25: 3159-73). There are several regulatory cascades and pathways that lead to cellular reprogramming, including a cytokinin-based route, an auxin-based route and a wound-induced route (Ikeuchi M. et al., *Plant Cell*.2013; 25: 3159-73). Wound-induced cellular reprogramming can occur due to bacterial, viral, and/or insect attack, as well as physical abrasion. *In vivo* callus formation has been generally observed across all higher plant genera. It was first reported in excised stem tissue of poplar, which was subsequently induced to form roots and shoots (Simon SV., *Jahrb Wiss Bot*. 1908; 45: 351-478). Callus induction is the basis of many in

vitro plant regeneration protocols (Skoog F. et al., *Symp Soc Exp Biol.* 1957; 11: 118-30) that are prerequisites for genetic engineering and genome editing (Liu D. et al., *Curr Opin Plant Biol.* 2016; 30: 70-7). Moreover, plant callus formation shares similar anatomical and physiologic features with human tumor formation (Birnbaum KD. Et al., *Cell.* 2008; 132: 697-710), highlighting the value of understanding the underlying mechanisms callus formation across the tree of life. Fully defining the genetic components of callus induction and formation is therefore of broad general interest to plant and animal biologists.

[0005] Individual species, as well as genotypes within a species, vary in their ability to form callus. Despite significant progress in the field (Butt SJ. et al., *Advancements in Life Sciences.* 2015; 2: 48-57; Gaur A. et al., *Isr J Plant Sci.* 2016; 63: 77-84), some commercially important plant species or genotypes within species often lack effective in vitro culture and callus induction protocols. This is particularly true for non-domesticated *Populus*, and without this capacity, creation of transgenic plants is difficult. Since callus induction and proliferation is under genetic control and regulation, identifying the genes and regulatory elements that control callus formation has the potential to facilitate the development of *in vitro* systems in recalcitrant plant species.

BRIEF SUMMARY OF THE DISCLOSURE

[0006] An aspect of the disclosure is directed to a genetically modified plant, plant cell or plant tissue, wherein the expression of a gene selected from the group consisting of Potri.003G018500, Potri.009G066100, Potri.012G083800, Potri.006G222700, Potri.008G208200, Potri.015G023600, Potri.004G118700, and Potri.018G014800, or a homolog thereof, is altered in the plant, plant cell or plant tissue.

[0007] Another aspect of the disclosure is directed to a method for increasing callus formation in a plant, plant cell or plant tissue comprising altering in a plant, plant cell or plant tissue the expression of a gene selected from the group consisting of Potri.003G018500, Potri.009G066100, Potri.012G083800, Potri.006G222700, Potri.008G208200, Potri.015G023600, Potri.004G118700, and Potri.018G014800, or a homolog thereof.

[0008] In some embodiments, the gene is selected from the group consisting of Potri.003G018500, Potri.009G066100, Potri.012G083800, Potri.006G222700, Potri.008G208200, and Potri.015G023600, and wherein the alteration comprises inactivation of the selected gene in the plant, plant cell or plant tissue, resulting in increased callus formation in the plant, plant cell or plant tissue.

[0009] In some embodiments, the inactivation of the selected gene is achieved by introducing a nucleic acid inhibitor of the selected gene to the plant, plant cell or plant tissue. In some embodiments, the nucleic acid inhibitor is selected from the group consisting of an antisense RNA, a small interfering RNA, an RNAi, a microRNA, an artificial microRNA, and a ribozyme.

[0010] In some embodiments, the inactivation of the selected gene is achieved by genome editing, which is achieved by a method selected from the group consisting of CRISPR/Cas system, Cre/Lox system, TALEN system, ZFNs system and homologous recombination. In some embodiments, the CRISPR-mediated genome editing comprises introducing into the plant a first nucleic acid encoding a Cas9 nuclease, a second nucleic acid comprising a guide RNA (gRNA), wherein said gRNA is specific to the selected gene.

[0011] In some embodiments, the gene is selected from the group consisting of Potri.004G118700, and Potri.018G014800, and wherein the alteration comprises expressing in the plant an exogenous nucleic acid comprising the selected gene, resulting in increased callus formation in the plant, plant cell or plant tissue.

[0012] In some embodiments, the plant is selected from the group consisting of genera *Acer*, *Afzelia*, *Allium*, *Arabidopsis*, *Agrostis*, *Avena*, *Betula*, *Brassica*, *Capsicum*, *Citrullus*, *Cucumis*, *Eucalyptus*, *Fagus*, *Festuca*, *Fraxinus*, *Fragaria*, *Glycine*, *Gossypium*, *Hordeum*, *Ipomoea*, *Jatropha*, *Juglans*, *Lemna*, *Lolium*, *Malus*, *Manihot*, *Medicago*, *Micropus*, *Milium*, *Miscanthus*, *Nicotiana*, *Oryza*, *Pennisetum*, *Phalaris*, *Phleum*, *Picea*, *Pinus*, *Poa*, *Populus*, *Prunus*, *Quercus*, *Rosa*, *Salix*, *Solanum*, *Sorghum*, *Spinacia*, *Tectona*, *Trifolium*, *Triticum*, *Panicum*, *Saccharum*, *Setaria*, *Zea*, and *Zoysia*.

BRIEF DESCRIPTION OF THE FIGURES

[0013] The patent or application file contains at least one drawing executed in color. Copies of this patent or patent application publication with color drawing(s) will be provided by the Office upon request and payment of the necessary fee.

[0014] FIG 1A – 1F. Callus formation on *Populus* leaf disc explants after 30 days on a callus induction medium. (**A**) 12 replicate leaf disk explants with callus along the midrib, (**B**) 12 replicate leaf disk explants with callus across the explant, (**C**) 12 replicate leaf disk explants with callus along the cut margin, (**D**) white friable callus along the midrib, (**E**) light green compact callus, and (**F**) green friable callus.

[0015] FIGs. 2A – 2H. Box plots for (A-D) callus formation and (E-H) callus rating score from a genome-wide association test. (**A**) callus formation of Potri.003G018500, (**B**) callus formation of Potri.004G118700, (**C**) callus formation of Potri.009G066100, (**D**) callus formation of Potri.018G014800, (**E**) callus rating score of Potri.006G222700, (**F**) callus rating score of Potri.008G208200, (**G**) callus rating score of Potri.012G083800, and (H) callus rating score of Potri.015G023600. In each panel, the homozygous rare (i.e., less common) alleles are displayed to the left in the red box, the heterozygous genotypes in the yellow box and the homozygous common alleles in the green box.

[0016] FIG. 3. Co-expression networks for the eight-significant genome-wide association loci related to callus formation and callus rating in *Populus*. Red edges indicate a positive co-expression at $r \geq 0.9$ and blue edges indicate negative coexpression at $r \leq -0.9$.

[0017] FIG. 4. Heat map of differentially expressed *Arabidopsis* orthologs, over 96 hours during callus induction, for *Populus* genes associated with callus formation or callus score in a genome-wide association study. Data taken from the NCBI GEO database.

[0018] FIG. 5. Co-expression network for orthologs of *Arabidopsis* genes tested in transgenic experiments and their association with *Populus* callus formation and callus rating genes identified via genome-wide association approaches. *Arabidopsis* orthologs are presented in parenthesis and *Populus* candidate genes are underlined. The *Populus* genes were discovered using a GWAS approach; the *Arabidopsis* genes were significantly co-

expressed with the candidate genes. Red edges indicate a positive co-expression at r>0.9 and blue edges indicate negative co-expression at r≤-0.9.

[0019] FIGs. 6A -6C. (**A**) Combined genome-wide association results and *Populus* co-expression analyses, with *Populus* homologs of *Arabidopsis*-tested transcription factors, in a proposed regulatory network. Gold boxes are *Populus* homologs of *Arabidopsis*-tested transcription factors; green boxes are GWAS identified *Populus* genes associated with callus formation. Red edges indicate positive co-expression, blue edges indicate negative co-expression. (**B**) The CNDbr, which negatively co-expressed with LEC2, was down-regulated in Populus leaf protoplasts when overexpressing LEC2. The other three genes (SOK1, MAPK3, and RPD3), positively co-expressed with LEC2, were not detected by qRT-PCR. Ctrl refers to the endogenous expression level of CNDbr in protoplasts while OE_LEC2 refers to the expression level of CNDbr when LEC2 was overexpressed in 3 independent replicates. The expression level of CNDbr was normalized to the ubiquitin internal control. (**C**) Expression patterns of five selected genes in co-expression network. LEC2 has extremely low abundance in leaves while CNDbr was highly expressed in leaves. SOK1, MAPK3, and RPD3 showed low abundances in leaf tissues.

DETAILED DESCRIPTION OF THE DISCLOSURE

Definitions

[0020] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs.

[0021] As used herein, the term "about" refers to an approximately +/-10% variation from a given value.

[0022] An "altered level of gene expression" refers to a measurable or observable change in the level of expression of a transcript of a gene, or the amount of its corresponding polypeptide, relative to a control plant or plant cell under the same conditions (e.g., as measured through a suitable assay such as quantitative RT-PCR, a Northern blot, a Western blot or through an observable change in phenotype, chemical profile or metabolic profile). An altered level of gene expression can include up-regulated or down-regulated expression

of a transcript of a gene or polypeptide relative to a control plant or plant cell under the same conditions. Altered expression levels can occur under different environmental or developmental conditions or in different locations than those exhibited by a plant or plant cell in its native state.

[0023] The term "control plant" as used herein refers to a plant cell, an explant, seed, plant component, plant tissue, plant organ, or whole plant used to compare against transgenic or genetically modified plant for the purpose of identifying an enhanced phenotype or a desirable trait in the transgenic or genetically modified plant. A "control plant" may in some cases be a transgenic plant line that comprises an empty vector or marker gene, but does not contain the recombinant polynucleotide of interest that is present in the transgenic or genetically modified plant being evaluated. A control plant may be a plant of the same line or variety as the transgenic or genetically modified plant being tested, or it may be another line or variety, such as a plant known to have a specific phenotype, characteristic, or known genotype. A suitable control plant would include a genetically unaltered or non-transgenic plant of the parental line used to generate a transgenic plant herein.

[0024] As used herein, the term "CRISPR" refers to a RNA-guided endonuclease comprising a nuclease, such as Cas9, and a guide RNA that directs cleavage of the DNA by hybridizing to a recognition site in the genomic DNA.

[0025] The term "DNA," as used herein, refers to a nucleic acid molecule of one or more nucleotides in length, wherein the nucleotide(s) are nucleotides. By "nucleotide" it is meant a naturally-occurring nucleotide, as well as modified versions thereof. The term "DNA" includes double- stranded DNA, single-stranded DNA, isolated DNA such as cDNA, as well as modified DNA that differs from naturally-occurring DNA by the addition, deletion, substitution and/or alteration of one or more nucleotides as described herein.

[0026] The term "exogenous," as used herein, refers to a substance or molecule originating or produced outside of an organism. The term "exogenous gene" or "exogenous nucleic acid molecule," as used herein, refers to a nucleic acid that codes for the expression of an RNA and/or protein that has been introduced ("transformed") into a cell or a progenitor of the cell. An exogenous gene may be from a different species (and so a "heterologous" gene) or from

the same species (and so a "homologous" gene), relative to the cell being transformed. A transformed cell may be referred to as a recombinant or genetically modified cell. An "endogenous" nucleic acid molecule, gene, or protein can represent the organism's own gene or protein as it is naturally produced by the organism.

[0027] The term "expression" refers to the process of converting genetic information of a polynucleotide into RNA through transcription, which is catalyzed by an enzyme, RNA polymerase and into protein, through translation of mRNA on ribosomes. Expression can be, for example, constitutive or regulated, such as, by an inducible promoter (e.g., lac operon, which can be triggered by Isopropyl β-D-1-thiogalactopyranoside (IPTG)). Up-regulation or overexpression refers to regulation that increases the production of expression products (mRNA, polypeptide or both) relative to basal or native states, while inhibition or down-regulation refers to regulation that decreases production of expression products (mRNA, polypeptide or both) relative to basal or native states.

[0028] The phrase "genetically modified," as used herein, refers to an organism whose genetic material has been altered by means of genetic engineering. Genetically modified organisms include genome-edited organisms, transgenic organisms, as well as organisms that were introduced exogenous nucleic acids into their cells.

[0029] The term "homolog" means a gene related to a second gene by descent from a common ancestral DNA sequence, therefore, the corresponding polynucleotide/polypeptide has a certain degree of homology, i.e., sequence identity (at least 40%, at least 60%, at least 65%, particularly preferred at least 66%, 68%, 70%, 75%, 80%, 86%, 88%, 90%, 92%, 95%, 97% or 99% sequence identity). A "homolog" furthermore means that the function is equivalent to the function of the original gene. Homologs of a given gene and homologous positions in the gene can be determined by sequence alignment programs, e.g., including but not limited to, NCBI BLAST, ClustalW, DIAMOND, CS-BLAST, and MAFFT.

[0030] As used herein, the term "nucleic acid" has its general meaning in the art and refers to a coding or non coding nucleic acid sequence. Nucleic acids include DNA (deoxyribonucleic acid) and RNA (ribonucleic acid) nucleic acids. Examples of nucleic acid thus include but are not limited to DNA, mRNA, tRNA, rRNA, tmRNA, miRNA, piRNA,

snoRNA, and snRNA. Nucleic acids thus encompass coding and noncoding region of a genome (i.e., nuclear or mitochondrial).

[0031] The term "operably linked" refers to positioning of a regulatory region and a sequence to be transcribed in a nucleic acid so as to influence transcription or translation of such a sequence. For example, to bring a coding sequence under the control of a regulatory region, the translation initiation site of the translational reading frame of the polypeptide is typically positioned between one and about fifty nucleotides downstream of the promoter. A regulatory region can, however, be positioned as much as about 5,000 nucleotides upstream of the translation initiation site or about 2,000 nucleotides upstream of the transcription start site. A regulatory region typically comprises at least a core (basal) promoter.

[0032] The term "regulatory region" refers to a nucleic acid having nucleotide sequences that influence transcription or translation initiation and rate and stability and/or mobility of a transcription or translation product. Regulatory regions include, without limitation, promoter sequences, enhancer sequences, response elements, protein recognition sites, inducible elements, protein binding sequences, 5' and 3' untranslated regions (UTRs), transcriptional start sites, termination sequences, polyadenylation sequences, introns and combinations thereof.

[0033] A regulatory region also may include at least one control element, such as an enhancer sequence, an upstream element or an upstream activation region (UAR). For example, a suitable enhancer is a cis-regulatory element (-212 to -154) from the upstream region of the octopine synthase (ocs) gene (Fromm et al., *The Plant Cell* 1:977-984 (1989)). The choice of regulatory regions to be included depends upon several factors, including, but not limited to, efficiency, selectability, inducibility, desired expression level and cell- or tissue-preferential expression. It is a routine matter for one of skill in the art to modulate the expression of a coding sequence by appropriately selecting and positioning regulatory regions relative to the coding sequence.

[0034] A "vector" is a replicon, such as a plasmid, phage or cosmid, into which another DNA segment may be inserted so as to bring about the replication of the inserted segment.

Generally, a vector is capable of replication when associated with the proper control elements. Suitable vector backbones include, for example, those routinely used in the art such as plasmids, viruses, artificial chromosomes, BACs, YACs or PACs. The term "vector" includes cloning and expression vectors, as well as viral vectors and integrating vectors. An "expression vector" is a vector that includes a regulatory region. Suitable expression vectors include, without limitation, plasmids and viral vectors derived from, for example, bacteriophage, baculoviruses and retroviruses. Numerous vectors and expression systems are commercially available from such corporations as Novagen (Madison, Wis.), Clontech (Mountain View, Calif.), Stratagene (La Jolla, Calif.) and Invitrogen/Life Technologies (Carlsbad, Calif.).

[0035] The vectors provided herein also can include, for example origins of replication, scaffold attachment regions (SARs) and/or markers. A marker gene can confer a selectable phenotype on a plant cell. For example, a marker can confer biocide resistance, such as resistance to an antibiotic (e.g., kanamycin, G418, bleomycin or hygromycin) or an herbicide (e.g., chlorosulfuron or phosphinothricin). In addition, an expression vector can include a tag sequence designed to facilitate manipulation or detection (e.g., purification or localization) of the expressed polypeptide. Tag sequences, such as green fluorescent protein (GFP), glutathione S-transferase (GST), polyhistidine, c-myc, hemagglutinin or Flag- tag (Kodak, New Haven, Conn.) sequences typically are expressed as a fusion with the encoded polypeptide. Such tags can be inserted anywhere within the polypeptide, including at either the carboxyl or amino terminus. As described herein, plant cells can be transformed with a recombinant nucleic acid construct to express a polypeptide of interest.

General Description

Genetically modified plants

[0036] One aspect of the disclosure is directed to a genetically modified plant, plant cell or plant tissue that has improved or increased callus formation and regeneration capabilities. Also included herein are plant cells and plant tissue, all derived from the genetically modified plant of the disclosure. In addition, seeds which can germinate into a genetically modified plant as described herein are also provided.

[0037] In some embodiments, the expression of a gene selected from the group consisting of Potri.003G018500, Potri.009G066100, Potri.012G083800, Potri.006G222700, Potri.008G208200, Potri.015G023600, Potri.004G118700, and Potri.018G014800, or a homolog thereof, is altered in the genetically modified plant, plant cell or plant tissue as compared to a control plant which was not genetically modified.

[0038] In some embodiments, the gene with the altered expression is selected from the group consisting of Potri.003G018500, Potri.009G066100, Potri.012G083800, Potri.006G222700, Potri.008G208200, and Potri.015G023600, and the alteration comprises inactivation of the selected gene in the plant, plant cell or plant tissue, resulting in increased callus formation in the plant, plant cell or plant tissue. The term "inactivation," as used herein, includes knocking out (e.g., deleting the gene using genome editing), knocking down (reducing the protein expression of a gene at least by 70%, at least 80%, at least 90%, at least 95%, or at least 99%, e.g., by using nucleic acid inhibitor), or abolishing at least one function (e.g., DNA binding ability, enzymatic activity etc.) of a gene.

[0001] In some embodiments, the inactivation of the selected gene is achieved by introducing a nucleic acid inhibitor of the selected gene to the plant. A "nucleic acid inhibitor" is a nucleic acid that can reduce or prevent expression or activity of a target gene. For example, an inhibitor of expression of a gene can reduce or eliminate transcription and/or translation of the gene product, thus reducing the gene protein expression.

[0039] In some embodiments, the nucleic acid inhibitor is selected from the group consisting of an antisense RNA, a small interfering RNA, an RNAi, a microRNA, an artificial microRNA, and a ribozyme.

[0040] In some embodiments, the inactivation of the selected gene is achieved by available gene targeting technologies in the art. Examples of gene targeting technologies include the Cre/Lox system (described in Kühn, R., & M. Torres, R., 2002. *Transgenesis Techniques: Principles and Protocols*, 175-204.), homologous recombination (described in Capecchi, Mario R. 1989. *Science*, 244: 1288-1292), TALENs (described in Sommer et al., 2015. *Chromosome Research*, 23: 43-55, and Cermak et al., 2011. *Nucleic Acids Research*: gkr218.), and CRISPR Cas system as described in Ran FA et al., 2013. *Nature Protocols*.

[0041] In one embodiment, gene modulation is achieved by a CRISPR/Cas system. CRISPR-Cas and similar gene targeting systems are well known in the art with reagents and protocols readily available (Mali, P. et al., 2013. Science, 339(6121), 823-826; Hsu, P.D. et al., 2014. Cell, 157.6: 1262-1278; Jiang et al., 2013. Nature Biotechnology, 31, 233-239). Exemplary genome editing protocols are described in Jennifer Doudna, and Prashant Mali, 2016. "CRISPR-Cas: A Laboratory Manual" (CSHL Press, ISBN: 978-1-621821-30-4) and Ran, F. Ann, et al. 2013. Nature Protocols, 8 (11): 2281-2308.

[0042] A CRISPR endonuclease comprises two components: (1) an RNA-dependent nuclease, typically microbial Cas9; and (2) a short "guide RNA" (gRNA or sgRNA) comprising a 20-nucleotide targeting sequence that directs the nuclease to a location of interest in the genome. When co-expressed with an artificial sgRNA targeting a cellular gene, the Cas9 endonuclease generates double-stranded breaks of DNA at the targeted locus. In addition, when CRISPR endonuclease is supplemented with a stretch of DNA template homologous to the break region, the break is repaired using the supplied homologous DNA template via the process of homologous recombination (HR). CRISPR-mediated HR makes it possible to specifically edit the target DNA sequence and/or alter gene expression.

[0043] In some embodiments, the CRISPR-mediated genome editing comprises introducing into the plant a first nucleic acid encoding a Cas9 nuclease, a second nucleic acid comprising a guide RNA (gRNA), wherein said gRNA is specific to the selected gene.

[0044] In some embodiments, the gene with the altered expression is selected from the group consisting of Potri.004G118700, and Potri.018G014800, and the alteration comprises expressing in the plant an exogenous nucleic acid comprising the selected gene, resulting in increased callus formation in the plant, plant cell or plant tissue.

[0045] Genetically modified plants of the disclosure are capable of self-pollinating or cross-pollinating with other plants of the same species so that the foreign gene, carried in the germ line, can be inserted into or bred into agriculturally useful plant varieties. The term "plant cell" as used herein, includes protoplasts, gamete producing cells, and cells which regenerate into whole plants. Accordingly, a seed comprising multiple plant cells capable of regenerating into a whole plant, is included in the definition of "plant cell."

[0046] In some embodiments, the genetically modified plant of the disclosure belongs to a recalcitrant plant species. The phrase "recalcitrant plant species" refers to plant species that are difficult to propagate *in vitro* under culture conditions. The phrase "recalcitrant plant species" also refers to plant species that do not readily produce calluses. The phrase "recalcitrant plant species" also refers to plant species that cannot be easily manipulated by genetic engineering methods.

[0047] In some embodiments, the genetically modified plant of the disclosure is a monocotyledonous plant. Examples of monocotyledonous plants include, but are not limited to, asparagus, field and sweet corn, barley, wheat, rice, sorghum, onion, pearl millet, rye and oats.

[0048] In some embodiments, the genetically modified plant of the disclosure is a dicotyledonous plant. Examples of dicotyledonous plants include, but are not limited to tomato, tobacco, cotton, rapeseed, field beans, soybeans, peppers, lettuce, peas, alfalfa, clover, cole crops or *Brassica oleracea* (e.g., cabbage, broccoli, cauliflower, brussel sprouts), radish, carrot, beets, eggplant, spinach, cucumber, squash, melons, cantaloupe, sunflowers and various ornamentals. Woody species include poplar, pine, sequoia, cedar, and oak.

[0049] In some embodiments, the plant is selected from the group consisting of genera *Acer, Afzelia, Allium, Arabidopsis, Agrostis, Avena, Betula, Brassica, Capsicum, Citrullus, Cucumis, Eucalyptus, Fagus, Festuca, Fraxinus, Fragaria, Glycine, Gossypium, Hordeum, Ipomoea, Jatropha, Juglans, Lemna, Lolium, Malus, Manihot, Medicago, Micropus, Milium, Miscanthus, Nicotiana, Oryza, Pennisetum, Phalaris, Phleum, Picea, Pinus, Poa, Populus, Prunus, Quercus, Rosa, Salix, Solanum, Sorghum, Spinacia, Tectona, Trifolium, Triticum, Panicum, Saccharum, Setaria, Zea,* and *Zoysia.*

Methods ForIncreasing Callus Formation and Regeneration in a Plant

[0050] This disclosure further provides methods for increasing callus formation and regeneration in a plant, plant cell or plant tissue.

[0051] In some embodiments, a method for increasing callus formation and/or regeneration in a plant, plant cell or plant tissue comprises altering in a plant, plant cell or plant tissue the

expression of a gene selected from the group consisting of Potri.003G018500, Potri.009G066100, Potri.012G083800, Potri.006G222700, Potri.008G208200, Potri.015G023600, Potri.004G118700, and Potri.018G014800, or a homolog thereof.

[0052] In some embodiments, the gene is selected from the group consisting of Potri.003G018500, Potri.009G066100, Potri.012G083800, Potri.006G222700, Potri.008G208200, and Potri.015G023600, and the alteration comprises inactivation of the selected gene in the plant, plant cell or plant tissue, resulting in increased callus formation in the plant, plant cell or plant tissue.

[0053] In some embodiments, the inactivation of the selected gene is achieved by introducing a nucleic acid inhibitor of the selected gene to the plant. In some embodiments, the nucleic acid inhibitor is selected from the group consisting of an antisense RNA, a small interfering RNA, an RNAi, a microRNA, an artificial microRNA, and a ribozyme.

[0054] In some embodiments, the inactivation of the selected gene is achieved by genome editing, which is achieved by a method selected from the group consisting of CRISPR/Cas system, Cre/Lox system, TALEN system, ZFNs system and homologous recombination.

[0055] In some embodiments, the gene is selected from the group consisting of Potri.004G118700, and Potri.018G014800, and the alteration comprises expressing in the plant an exogenous nucleic acid comprising the selected gene, resulting in increased callus formation in the plant, plant cell or plant tissue.

[0056] In some embodiments, the plant belongs a to recalcitrant plant species.

[0057] In some embodiments, the plant is a monocotyledonous plant. Examples of monocotyledonous plants include, but are not limited to, asparagus, field and sweet corn, barley, wheat, rice, sorghum, onion, pearl millet, rye and oats.

[0058] In some embodiments, the plant is a dicotyledonous plant. Examples of dicotyledonous plants include, but are not limited to tomato, tobacco, cotton, rapeseed, field beans, soybeans, peppers, lettuce, peas, alfalfa, clover, cole crops or *Brassica oleracea* (e.g., cabbage, broccoli, cauliflower, brussel sprouts), radish, carrot, beets, eggplant, spinach,

cucumber, squash, melons, cantaloupe, sunflowers and various ornamentals. Woody species include poplar, pine, sequoia, cedar, and oak.

[0059] In some embodiments, the plant is selected from the group consisting of genera *Acer, Afzelia, Allium, Arabidopsis, Agrostis, Avena, Betula, Brassica, Capsicum, Citrullus, Cucumis, Eucalyptus, Fagus, Festuca, Fraxinus, Fragaria, Glycine, Gossypium, Hordeum, Ipomoea, Jatropha, Juglans, Lemna, Lolium, Malus, Manihot, Medicago, Micropus, Milium, Miscanthus, Nicotiana, Oryza, Pennisetum, Phalaris, Phleum, Picea, Pinus, Poa, Populus, Prunus, Quercus, Rosa, Salix, Solanum, Sorghum, Spinacia, Tectona, Trifolium, Triticum, Panicum, Saccharum, Setaria, Zea,* and *Zoysia.*

Nucleic Acid Inhibitors

[0060] An aspect of the disclosure provides a number of nucleic acid based methods, including antisense RNA, ribozyme directed RNA cleavage, post-transcriptional gene silencing (PTGS), e.g., RNA interference (RNAi), microRNA and artificial microRNA and transcriptional gene silencing (TGS) that can be used to inhibit the expression of a gene selected from the group consisting of Potri.003G018500, Potri.009G066100, Potri.012G083800, Potri.006G222700, Potri.008G208200, and Potri.015G023600, or a homolog thereof.

[0061] Suitable nucleic acid inhibitors, i.e., nucleic acids capable of inhibiting the expression of a target gene, include full-length nucleic acids of allelic variants of a gene selected from the group consisting of Potri.003G018500, Potri.009G066100, Potri.012G083800, Potri.006G222700, Potri.008G208200, and Potri.015G023600, or a homolog thereof, or fragments of such full-length nucleic acids. In some embodiments, a complement of the full-length nucleic acid or a fragment thereof can be used. Typically, a fragment is at least 10 nucleotides, e.g., at least 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 30, 35, 40, 50, 80, 100, 200, 500 nucleotides or more. Generally, higher homology can be used to compensate for the use of a shorter sequence.

[0062] Antisense technology is one well-known method. In this method, a nucleic acid segment from a gene to be repressed is cloned and operably linked to a regulatory region and a transcription termination sequence so that the antisense strand of RNA is transcribed.

The recombinant vector is then transformed into plants, as described below and the antisense strand of RNA is produced. The nucleic acid segment need not be the entire sequence of the gene to be repressed, but typically will be substantially complementary to at least a portion of the sense strand of the gene to be repressed.

[0063] In another method, a nucleic acid can be transcribed into a ribozyme or catalytic RNA, which affects expression of an mRNA. See, U.S. Pat. No. 6,423,885. Ribozymes can be designed to specifically pair with a target RNA and cleave the phosphodiester backbone at a specific location, thereby functionally inactivating the target RNA. Heterologous nucleic acids can encode ribozymes designed to cleave particular mRNA transcripts, thus preventing expression of a polypeptide. Hammerhead ribozymes cleave mRNAs at locations dictated by flanking regions that form complementary base pairs with the target mRNA. See, for example, U.S. Pat. No. 5,254,678; Perriman et al., *PNAS* 92(13):6175-6179 (1995); de Feyter and Gaudron, Methods in Molecular Biology, Vol. 74, Chapter 43, Edited by Turner, P. C., Humana Press Inc., Totowa, N.J. RNA endoribonucleases which have been described, such as the one that occurs naturally in *Tetrahymena thermophila*, can be useful. See, for example, U.S. Pat. Nos. 4,987,071 and 6,423,885.

[0064] PTGS, e.g., RNAi, can also be used to inhibit the expression of a gene. For example, a construct can be prepared that includes a sequence that is transcribed into an RNA that can anneal to itself, e.g., a double stranded RNA having a stem-loop structure. In some embodiments, one strand of the stem portion of a double stranded RNA comprises a sequence that is similar or identical to the sense coding sequence or a fragment thereof, of the polypeptide of interest. The length of the sequence that is similar or identical to the sense coding sequence can be from 10 nucleotides to 500 nucleotides, from 15 nucleotides to 300 nucleotides, from 20 nucleotides to 100 nucleotides or from 25 nucleotides to 100 nucleotides. The other strand of the stem portion of a double stranded RNA comprises a sequence that is similar or identical to the antisense strand or a fragment thereof, of the coding sequence of the polypeptide of interest and can have a length that is shorter, the same as or longer than the corresponding length of the sense sequence. In some cases, one strand of the stem portion of a double stranded RNA comprises a sequence that is similar or identical to the 3' or 5' untranslated region or a fragment thereof, of the mRNA encoding the

polypeptide of interest and the other strand of the stem portion of the double stranded RNA comprises a sequence that is similar or identical to the sequence that is complementary to the 3' or 5' untranslated region, respectively or a fragment thereof, of the mRNA encoding the polypeptide of interest. In other embodiments, one strand of the stem portion of a double stranded RNA comprises a sequence that is similar or identical to the sequence of an intron or a fragment thereof in the pre-mRNA encoding the polypeptide of interest and the other strand of the stem portion comprises a sequence that is similar or identical to the sequence that is complementary to the sequence of the intron or fragment thereof in the pre-mRNA.

[0065] A construct including a sequence that is operably linked to a regulatory region and a transcription termination sequence and that is transcribed into an RNA that can form a double stranded RNA, can be transformed into plants as described below. Methods for using RNAi to inhibit the expression of a gene are known to those of skill in the art. See, e.g., U.S. Pat. Nos. 5,034,323; 6,326,527; 6,452,067; 6,573,099; 6,753,139; and 6,777,588. See also WO 97/01952; WO 98/53083; WO 99/32619; WO 98/36083; and U.S. Patent Publications 20030175965, 20030175783, 20040214330 and 20030180945.

[0066] In some embodiments, a construct containing a nucleic acid having at least one strand that is a template for both sense and antisense sequences that are complementary to each other is used to inhibit the expression of a gene. The sense and antisense sequences can be part of a larger nucleic acid molecule or can be part of separate nucleic acid molecules having sequences that are not complementary. The sense or antisense sequence can be a sequence that is identical or complementary to the sequence of an mRNA, the 3' or 5' untranslated region of an mRNA or an intron in a pre-mRNA encoding a polypeptide of interest or a fragment of such sequences. In some embodiments, the sense or antisense sequence is identical or complementary to a sequence of the regulatory region that drives transcription of the gene encoding a polypeptide of interest. In each case, the sense sequence is the sequence that is complementary to the antisense sequence.

[0067] A nucleic acid having at least one strand that is a template for one or more sense and/or antisense sequences can be operably linked to a regulatory region to drive transcription of an RNA molecule containing the sense and/or antisense sequence(s). In addition, such a nucleic acid can be operably linked to a transcription terminator sequence,

such as the terminator of the nopaline synthase (nos) gene. In some cases, two regulatory regions can direct transcription of two transcripts: one from the top strand and one from the bottom strand. See, for example, Yan et al., *Plant Physiol.*, 141:1508-1518 (2006). The two regulatory regions can be the same or different. The two transcripts can form double-stranded RNA molecules that induce degradation of the target RNA. In some cases, a nucleic acid can be positioned within a P-DNA such that the left and right border-like sequences of the P-DNA are on either side of the nucleic acid.

[0068] In some embodiments, a suitable nucleic acid inhibitor can be a nucleic acid analog. Nucleic acid analogs can be modified at the base moiety, sugar moiety or phosphate backbone to improve, for example, stability, hybridization or solubility of the nucleic acid. Modifications at the base moiety include deoxyuridine for deoxythymidine and 5-methyl-2'-deoxycytidine and 5-bromo-2'-deoxycytidine for deoxycytidine. Modifications of the sugar moiety include modification of the 2' hydroxyl of the ribose sugar to form 2'-O-methyl or 2'-O-allyl sugars. The deoxyribose phosphate backbone can be modified to produce morpholino nucleic acids, in which each base moiety is linked to a six-membered morpholino ring or peptide nucleic acids, in which the deoxyphosphate backbone is replaced by a pseudopeptide backbone and the four bases are retained. See, for example, Summerton and Weller, 1997, *Antisense Nucleic Acid Drug Dev.*, 7:187-195; Hyrup et al., *Bioorgan. Med. Chem.*, 4:5-23 (1996). In addition, the deoxyphosphate backbone can be replaced with, for example, a phosphorothioate or phosphorodithioate backbone, a phosphoroamidite or an alkyl phosphotriester backbone.

Expression Vector Modulators

[0069] This disclosure provides an exogenous nucleic acid vector that comprises a nucleotide sequence that is transcribed into a nucleic acid inhibitor of a gene selected from the group consisting of Potri.003G018500, Potri.009G066100, Potri.012G083800, Potri.006G222700, Potri.008G208200, and Potri.015G023600, operably linked to a regulatory region that is functional in a plant, plant cell or plant tissue as described above, where a plant, plant cell or plant tissue expressing such exogenous nucleic acid vector displays increased callus formation and regeneration properties compared to a control plant that does not comprise the nucleic acid vector.

[0070] In a specific embodiment, the Potri.003G018500 gene comprises a nucleotide sequence shown by SEQ ID NO: 15, encoding the protein SEQ ID NO: 23. In some embodiments, the Potri.003G018500 nucleotide sequence comprises a sequence that shows at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identity to SEQ ID NO: 15. In some embodiments, the Potri.003G018500 the nucleotide sequence encodes a protein that shows at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identity to SEQ ID NO: 23.

[0071] In a specific embodiment, the Potri.009G066100 gene comprises a nucleotide sequence shown by SEQ ID NO: 13, encoding the protein SEQ ID NO: 21. In some embodiments, the Potri.009G066100 nucleotide sequence comprises a sequence that shows at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identity to SEQ ID NO: 13. In some embodiments, the Potri.009G066100 nucleotide sequence encodes a protein that shows at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identity to SEQ ID NO: 21.

[0072] In a specific embodiment, the Potri. 012G083800 gene comprises a nucleotide sequence shown by SEQ ID NO: 20, encoding the protein SEQ ID NO: 28. In some embodiments, the Potri. 012G083800 nucleotide sequence comprises a sequence that shows at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identity to SEQ ID NO: 20. In some embodiments, the Potri. 012G083800 nucleotide sequence encodes a protein that shows at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identity to SEQ ID NO: 28.

[0073] In a specific embodiment, the Potri. 006G222700 gene comprises a nucleotide sequence shown by SEQ ID NO: 16, encoding the protein SEQ ID NO: 24. In some embodiments, the Potri. 006G222700 nucleotide sequence comprises a sequence that shows at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identity to SEQ ID NO: 16. In some embodiments, the Potri. 006G222700 nucleotide sequence encodes a protein that shows at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identity to SEQ ID NO: 24.

[0074] In a specific embodiment, the Potri. 008G208200 gene comprises a nucleotide sequence shown by SEQ ID NO: 19, encoding the protein SEQ ID NO: 27. In some embodiments, the Potri. 008G208200 nucleotide sequence comprises a sequence that shows at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identity to SEQ ID NO: 19. In some embodiments, the Potri. 008G208200 nucleotide sequence encodes a protein that shows at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identity to SEQ ID NO: 27.

[0075] In a specific embodiment, the Potri. 015G023600 gene comprises a nucleotide sequence shown by SEQ ID NO: 18, encoding the protein SEQ ID NO: 26. In some embodiments, the Potri. 015G023600 nucleotide sequence comprises a sequence that shows at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identity to SEQ ID NO: 18. In some embodiments, the Potri. 015G023600 nucleotide sequence encodes a protein that shows at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identity to SEQ ID NO: 26.

[0076] This disclosure further provides an exogenous nucleic acid vector that comprises a nucleotide sequence that is transcribed into expression or overexpression of a gene selected from the group consisting of Potri.004G118700, and Potri.018G014800, operably linked to a regulatory region that is functional in a plant as described above, where a plant, plant cell or plant tissue expressing such exogenous nucleic acid vector displays increased callus formation and regeneration properties compared to a control plant that does not comprise the nucleic acid vector.

[0077] In a specific embodiment, the nucleotide sequence comprises the Potri. 004G118700 gene shown by SEQ ID NO: 14, encoding the protein SEQ ID NO: 22. In some embodiments, the nucleotide sequence comprises a sequence that shows at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identity to SEQ ID NO: 14. In some embodiments, the nucleotide sequence encodes a protein that shows at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identity to SEQ ID NO: 22.

[0078] In a specific embodiment, the nucleotide sequence comprises the Potri. 018G014800 gene shown by SEQ ID NO: 17, encoding the protein SEQ ID NO: 25. In some

embodiments, the nucleotide sequence comprises a sequence that shows at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identity to SEQ ID NO: 17. In some embodiments, the nucleotide sequence encodes a protein that shows at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% identity to SEQ ID NO: 25.

[0079] A variety of promoters are available for use, depending on the degree of expression desired. For example, a broadly expressing promoter promotes transcription in many, but not necessarily all, plant tissues. Non-limiting examples of broadly expressing promoters that can be included in the nucleic acid constructs provided herein include the cauliflower mosaic virus (CaMV) 35S promoter, the mannopine synthase (MAS) promoter, the 1' or 2' promoters derived from T-DNA of *Agrobacterium tumefaciens*, the figwort mosaic virus 34S promoter, actin promoters such as the rice actin promoter and ubiquitin promoters such as the maize ubiquitin-1 promoter.

[0080] Some suitable regulatory regions initiate transcription, only or predominantly, in certain cell types. For example, a promoter that is active predominantly in a reproductive tissue (e.g., fruit, ovule or inflorescence) can be used. Thus, as used herein a cell type- or tissue-preferential promoter is one that drives expression preferentially in the target tissue, but may also lead to some expression in other cell types or tissues as well.

[0081] Root-active and root-preferential promoters confer transcription in root tissue, e.g., root endodermis, root epidermis or root vascular tissues. Root-preferential promoters include the root-specific subdomains of the CaMV 35S promoter (Lam et al., *Proc. Natl. Acad. Sci.* USA, 86:7890-7894 (1989)), root cell specific promoters reported by Conkling et al., *Plant Physiol.*, 93:1203-1211 (1990) and the tobacco RD2 promoter.

[0082] Promoters active in photosynthetic tissue confer transcription in green tissues such as leaves and stems. Examples of such promoters include the ribulose-1,5-bisphosphate carboxylase (RbcS) promoters such as the RbcS promoter from eastern larch (Larix laricina), the pine cab6 promoter (Yamamoto et al., *Plant Cell Physiol.*, 35:773-778 (1994)), the Cab-1 promoter from wheat (Fejes et al., *Plant Mol. Biol.*, 15:921-932 (1990)), the CAB-1 promoter from spinach (Lubberstedt et al., *Plant Physiol.*, 104:997-1006 (1994)), the cab IR promoter from rice (Luan et al., *Plant Cell*, 4:971-981 (1992)), the pyruvate

orthophosphate dikinase (PPDK) promoter from corn (Matsuoka et al., *Proc. Natl. Acad. Sci. USA*, 90:9586-9590 (1993)), the tobacco Lhcb1*2 promoter (Cerdan et al., *Plant Mol. Biol.*, 33:245-255 (1997)), the Arabidopsis SUC2 sucrose-H+ symporter promoter (Truernit et al., *Planta*, 196:564-570 (1995)) and thylakoid membrane protein promoters from spinach (psaD, psaF, psaE, PC, FNR, atpC, atpD, cab, rbcS).

[0083] Lignin biosynthesis promoters are promoters that drive transcription of nucleic acids encoding enzymes involved in lignin biosynthesis. Examples of lignin biosynthesis promoters include promoters of the switchgrass (*Panicum virgatum*), rice (*Oryza sativa*), corn (*Zea mays*) and wheat (*Triticum aestivum*) homologs of the *Populus* cinnamate 4-hydroxylase, caffeoyl-CoA O-methyltransferase and caffeic acid O-methyltransferase genes. Also suitable are promoters of *Arabidopsis* genes encoding phenylalanin ammonia lyase (genomic locus At3g10340), trans-cinnamate 4-hydroxylase (genomic locus At2g30490), 4-coumarate:CoA ligase (genomic locus At1g51680), hydroxycinnamoyl-CoA:shikimate/quinate hydroxycinnamoyltransferase (genomic locus At5g48930), p-coumarate 3-hydroxylase (genomic locus At2g40890), caffeoyl-CoA 3-O-methyltransferase (genomic locus At4g34050), cinnamoyl CoA reductase (genomic locus At1g15950), ferulate 5-hydroxylase (genomic locus At4g36220), caffeic acid O-methyltransferase (genomic locus At5g54160) and cinnamyl alcohol dehydrogenase (genomic locus At4g34230).

[0084] Useful promoters also include cell wall related promoters, such as cellulose biosynthesis promoters. Cellulose biosynthesis promoters are promoters that drive transcription of nucleic acids encoding enzymes involved in cellulose biosynthesis. Examples of cellulose biosynthesis promoters include the promoter of the rice cellulose synthase gene (genomic locus Os08g25710), the promoter of the rice cellulose synthase gene (genomic locus Os08g06380) and the promoter of the rice cellulose synthase-like A2 gene (genomic locus Os10g26630).

[0085] Examples of promoters that have high or preferential activity in vascular bundles include the glycine-rich cell wall protein GRP 1.8 promoter (Keller and Baumgartner, *Plant Cell*, 3(10):1051-1061 (1991)), the Commelina yellow mottle virus (CoYMV) promoter (Medberry et al., *Plant Cell*, 4(2):185-192 (1992)) and the rice tungro bacilliform virus (RTBV) promoter (Dai et al., *Proc. Natl. Acad. Sci. USA*, 101(2):687-692 (2004)).

Promoters having preferential activity in the phloem region (e.g., primary phloem cells, companion cells and sieve cells), the xylem region (e.g., tracheids and vessels), the bundle sheath layer and/or the endodermis are also considered vascular tissue promoters. Promoters that have preferential activity in the pith, cortex, epidermis and/or in the vascular bundles or vascular layers of the stem are considered stem promoters. In some cases, the activity of stem promoters can also be induced by stress like drought.

[0086] Inducible promoters confer transcription in response to external stimuli such as chemical agents or environmental stimuli. For example, inducible promoters can confer transcription in response to hormones such as gibberellic acid or ethylene or in response to light, nitrogen, shade or drought.

[0087] A basal promoter is the minimal sequence necessary for assembly of a transcription complex required for transcription initiation. Basal promoters frequently include a "TATA box" element that may be located between about 15 and about 35 nucleotides upstream from the site of transcription initiation. Basal promoters also may include a "CCAAT box" element (typically the sequence CCAAT) and/or a GGGCG sequence, which can be located between about 40 and about 200 nucleotides, typically about 60 to about 120 nucleotides, upstream from the transcription start site.

[0088] A 5' untranslated region (UTR) can be included in nucleic acid constructs described herein. A 5' UTR is transcribed, but is not translated and lies between the start site of the transcript and the translation initiation codon and may include the +1 nucleotide. A 3' UTR can be positioned between the translation termination codon and the end of the transcript. UTRs can have particular functions such as increasing mRNA stability or attenuating translation. Examples of 3' UTRs include, but are not limited to, polyadenylation signals and transcription termination sequences, e.g., a nopaline synthase termination sequence.

[0089] It will be understood that more than one regulatory region may be present in a recombinant polynucleotide, e.g., introns, enhancers, upstream activation regions, transcription terminators and inducible elements. Thus, for example, more than one regulatory region can be operably linked to the sequence of a polynucleotide encoding a Gene Y homolog or other lignin-modulating polypeptide. Regulatory regions, such as

promoters for endogenous genes, can be obtained by chemical synthesis or by subcloning from a genomic DNA that includes such a regulatory region. A nucleic acid comprising such a regulatory region can also include flanking sequences that contain restriction enzyme sites that facilitate subsequent manipulation.

[0090] In one aspect, a plant cell comprising a Potri.003G018500, Potri.009G066100, Potri.012G083800, Potri.006G222700, Potri.008G208200, or Potri.015G023600 nucleic acid inhibitor is provided. The plant cell comprises an exogenous nucleic acid, the exogenous nucleic acid comprising a regulatory region operably linked to a polynucleotide that is transcribed into an interfering RNA effective for inhibiting expression of Potri.003G018500, Potri.009G066100, Potri.012G083800, Potri.006G222700, Potri.008G208200, or an allelic variant of any one of these genes. The exogenous nucleic acid can further comprise a 3' UTR operably linked to the polynucleotide. The polynucleotide can be transcribed into an interfering RNA comprising a stem-loop structure. The stem-loop structure can comprise an inverted repeat of the 3' UTR.

Methods of Use of Genetically Modified (Transgenic) Plants

[0091] This disclosure provides methods of using the disclosed plants with increased callus formation and regeneration properties in biofuel production processes. Methods of pretreatment and saccharification of biomass to fermentable sugars, followed by fermentation of the sugars to ethanol, are known in the art.

[0092] This disclosure further provides methods of using the disclosed genetically modified plants with increased callus formation and regeneration properties as starter plants for further plant propagation and/or genetic engineering.

[0093] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one skilled in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can also be used in the practice or testing of the present invention, the preferred methods and materials are now described. All publications mentioned herein are incorporated herein by reference to disclose and describe the methods and/or materials in connection with which the publications are cited.

[0094] The present disclosure is further illustrated by the following non-limiting examples.

EXAMPLES

Example 1: Materials and Methods

Plant materials

[0095] From within 1084 genotypes of *Populus trichocarpa* contained in the GWAS population (Geraldes A. et al, *PLoS One*, PMID: 23311503; Evans LM. et al, *Nat Genet.* 2014; 46: 1089-96) callus induction was tested in 280 genotypes. To avoid potential bias in allele frequencies, these genotypes were selected to represent the latitudinal gradient across the natural range of this species in the Pacific Northwest of North America. Global Position Systems (GPS) co-ordinates recorded when each genotype was originally collected were used to uniformly sample across river systems in this range (Slavov GT. et al., *New Phytol.* 2012; 196: 713-25). Clonal replicates of each genotype were grown in the greenhouse for three months prior to sampling leaf tissue for explant establishment. Each genotype had been re-sequenced to a minimum of an 18x depth and a SNP library with 8.2 million SNPs was available for the GWAS analyses. Whole-genome resequencing, alignment of Illumina short reads to the reference *P. trichocarpa* genome, SNP calling and data curation parameters are fully described by Evans et al. (Evans LM. et al, *Nat Genet.* 2014; 46: 1089-96).

Callus induction

[0096] Multiple fully expanded leaves were collected from each genotype and surface disinfested using 1% (v/v) Tween-20 solution for 3 to 5 min, 70% (v/v) ethanol for 1 min, and 10% bleach solution (5.25% sodium hypochlorite) for 10 min, followed by 3 rinses with sterile, distilled water (Kang B-g. et al., *Plant Cell*, Tissue and Organ Culture (PCTOC). 2009; 99: 251). Explants were aseptically cut from leaves using a 1 cm^2 diameter cork borer and placed adaxial-side up on medium previously proven successful for callus induction of *Populus trichocarpa* (Meilan R, and Ma C., *Agrobacterium Protocols Volume* 2. Humana Press; 2007. pp. 143-51). Specifically, a Murashige and Skoog (MS) medium (Murashige T. et al. *Physiol Plant*. 1962; 15: 473-97) was supplemented with 0.5 μM 6-benzylaminopurine (BA), 0.5 μM zeatin, 5 μM naphthaleneacetic acid (NAA), 5 μM 2,4-D, and 1.28 mM 1-

morpholinoethanesulfonic acid (MES), adjusted to a pH of 5.8, and solidified using 0.3% Phytoagar and 0.1% Gelrite. Midveins within the leaf explants were targeted as explants due to their organogenic potential. Three replications with 12 leaf disks per plate per replication were initiated. Cultures were then incubated for 4 weeks under constant dark at 25°C.

[0097] A second callus induction experiment was conducted using the seven genotypes each with the most and least prolific callus formation. Following the same protocol described above, leaf explants were cultured on media with varying phytohormone levels. Combinations of high and low cytokinin to auxin were tested: high cytokinin/low auxin with 5 μM BA + 0.5 μM TDZ + 0.5 μM NAA; high cytokinin/medium auxin with 5 μM BA + 0.5 μM TDZ + 1 μM NAA; high cytokinin/high auxin with 5 μM BA + 0.5 μM TDZ + 5 μM NAA; low cytokinin/low auxin with 1 μM BA + 0.5 μM TDZ + 0.5 μM NAA; low cytokinin/medium auxin with 1 μM BA + 0.5 μM TDZ + 1 μM NAA; low cytokinin/high auxin with 1 μM BA + 0.5 μM TDZ + 5 μM NAA. Three replications with 12 leaf disks per plate per replication were examined. Cultures were incubated for 4 weeks under constant dark at 25°C.

Callus rating

[0098] The number of explants forming callus was counted and scored based on the amount of callus formed. Callus formation was recorded as a percent of the 12 explants per replicate forming callus. Callus ratings score were assigned as follows: 0 for no callus formation, 1 for compact callus, 2 for green friable callus, and 3 for white friable callus. Callus organogenic potential is known to vary by callus appearance (Meilan R, and Ma C., *Agrobacterium Protocols Volume* 2. Humana Press; 2007. pp. 143-51), with white friable callus leading to greater shoot induction potential. Location of callus formation on the explant was also noted as initiating from the midvein and/or from the cut edge.

Analysis of variance

[0099] A two-way ANOVA, with genotype (G), replication (R) and GxR interaction as random effect sources of variation, was used to test for significant differences among genotypes in callus formation and callus rating ($p \leq 0.05$). Broad-sense heritability was calculated as the variance due to genotype divided by the summation of the error variance

plus the genotype variance. Heritability was only calculated when there were significant genotype effects. A one-tailed t-test ($p \leq 0.05$) was used to test difference among hormone treatments in the second callus induction experiment.

Genome-wide association test

[0100] To determine genetic loci associated with callus formation or callus rating, the EMMAX algorithm was used, with kinship as the correction factor for genetic background effects (Kang HM. et al, *Nat Genet.* 2010; 42: 348-54), to compute genotype-to-phenotype associations using 8.2 million SNPs with minor allele frequencies ≥ 0.05 as described by Zhang et al. (2018) (Zhang J. et al., *New Phytol.* 2018 July 11). Callus formation and callus rating candidate genes were identified based SNP association which exceeded the chromosome-wide–log10 (p) = 4.46 [p = 3.47E−05] Bonferroni-adjusted significance threshold. GWAS tests were run independently by replicates and only those associations that were significant across all three replicates are reported here.

Gene Atlas analysis

[0101] Gene Atlas data (Goodstein DM. et al., *Nucleic Acids Res.* 2011; 40(D1): D1178-D86) for four callus formation and four callus rating genes were collected from Phytomine database integrated in Phytozome (v.11.0) with FPKM value (Grigoriev IV. et al., *Nucleic Acids Res.* 2011; 40: D26-D32; Nordberg H. et al., *Nucleic Acids Res.* 2013; 42: D26-D31). The $\log_2$ scaled FPKM from a total of 24 different tissue types or conditions were subjected to 'gplots' of R package and summarized in heat maps (Warnes GR. et al., *R package version.* 2009; 2: 1).

Gene Co-expression network construction and gene ontology enrichment

[0102] Gene Atlas data across seven tissues was also used to calculate Pearson correlation coefficients between the expression profiles of all pairs of genes using the mcxarray and mcxdump programs from the MCL-Edge software package (Van Dongen S. *SIAM Journal on Matrix Analysis and Applications.* 2008; 30: 121-41). Correlation were calculated in a parallel fashion making use of the Parallel::MPI::Simple Perl module available on the Comprehensive Perl Archive Network (CPAN). A respective 0.8, and -0.8 Pearson threshold was applied and subnetworks of genes that co-express (positive or negative) with

the eight candidate genes identified by GWAS were created and visualized in Cytoscape (Shannon P. et al., Genome Res. 2003; 13: 2498-504).

Arabidopsis callus orthologs

[0103] *Arabidopsis*-based microarray expression data was obtained from GSE29543, probes were mapped to the Affymetrix ATH1-121501 *Arabidopsis* annotation V35, expression was normalized using robust multi-array averaging (RMA) and then Linear Models for Microarray and RNA-Seq Data (Limma) was used to calculate differential expression. Time points of 12 h, 24 h, 48 h and 96 h of callus induction were compared to 0 h, representing establishment of shoot explants introduced to callus inducing media. A heat map was then constructed based on fold-change values from *Arabidopsis* genes that were significantly differentially expressed in shoot (adjusted p-value≤0.05) in at least one comparison and that were orthologs to *Populus* genes associated with callus formation.

Transient overexpression in Populus protoplast and quantitative RT-PCR (qRT-PCR)

[0104] Protoplasts from hybrid poplar 717 (*Populus tremula X alba*) leaves were isolated and subsequently transfected as previously described (Guo J. et al., *PLoS One*. 2012; 7: e44908). The full-length CDS of LEC2 (Potri.004G045800) was determined according to the sequence information available at Phytozome. The CDS of LEC2 was introduced into the pENTR/D-TOPO vector (Life Technologies), and subsequently transferred into a Gateway destination vector via LR reaction. The Gateway destination vector was constructed by amplifying the 35S promoter, the Gateway cassette and the Tnos terminator from pGWB502 (Nakagawa T. et al., *Biosci, Biotechnol, Biochem*. 2007; 71: 2095-100), using primers 5'-ATGGTACCTGAGACTTTTCAACAAAGGGTA-3' (SEQ ID NO: 1) and 5'-ATAAGCTTGATCTAGTAACATAGATGACAC-3' (SEQ ID NO: 2), was subcloned into the pUC19 vector using restriction enzymes *KpnI* and *HindIII*.

[0105] Total RNA from transfected and control *Populus* protoplasts was extracted using the Spectrum Plant Total RNA isolation kit (Sigma). One μg of total RNA were reversely transcribed to cDNA using RevertAid Reverse Transcriptase (Thermo Fisher Scientific). qRT-PCR was performed using Maxima SYBR Green/ROX qPCR Master Mix (Thermo Fisher Scientific). *Populus* Ubiquitin (UBQ10b) was used as an internal control for

normalizing the relative transcript level. All PCR reactions were completed with at least three replicates. The primers used for qRT-PCR are listed in Table 1.

Table 1. qRT-PCR primers used in the transient protoplast assay.

Gene Name	Gene ID	Primer	Primer Sequences
LEC2	Potri.004G045800	Forward	GGTGCTAGTACTTGTGGCCAAAGA (SEQ ID NO: 3)
		Reverse	TTCCTAAGCACCGCTCTGAGTC (SEQ ID NO: 4)
CNDbr	Potri.018G014800	Forward	ATATTTGACACAGGCAGTGGTCT (SEQ ID NO: 5)
		Reverse	GTTAAGTAGGTGCACTTCGGAGA (SEQ ID NO: 6)
MAPK3	Potri.009G066100	Forward	AGATCTCAAACCCAGCAATTTACTGC (SEQ ID NO: 7)
		Reverse	ACACATCAATTGCAGCAGTATAGTCG (SEQ ID NO: 8)
SOK1	Potri.003G018500	Forward	CAGCTTGCTTGTCTGATTGAATCAACA (SEQ ID NO: 9)
		Reverse	GGTGATCAATGTTTTCCAAGCTGGAG (SEQ ID NO: 10)
UBQ10b (Ctrl)	Potri.001G263000	Forward	GCCTTCGTGGTGGTTATTAAGC (SEQ ID NO: 11)
		Reverse	TCCAACAATGGCCAGTAAACAC (SEQ ID NO: 12)

Example 2: Callus formation and rating is genotype dependent

[0106] Among the 280 *P. trichocarpa* genotypes tested for callus induction, 21 genotypes produced no callus and 30 genotypes formed callus from 100% of their explants (FIG. 1). The mean callus formation frequency across all genotypes was 53%+1.9% (mean+s.e.). Among those genotypes that did form callus; the mean callus rating was 1.3+0.05, with only 49 genotypes averaging a rating of 2.5 or higher. In total, 101 genotypes had a mean callus rating score of 1.0 or greater. Callus formation and callus rating were positively correlated

with r2 = 0.77. Of the explants that formed callus, 73% initiated from the midrib and 25% formed callus along the cut edge of the leaf explant. Genotype had a significant effect on callus formation (F279, 558 = 7.16, p-value = 4.28E-86) and callus rating (F279, 558 = 6.56, p-value = 5.42E-79). Broad-sense heritability for callus formation was h2 = 0.67 and heritability for callus rating h2 = 0.65.

Example 3: Candidate genes associated with callus formation and rating

[0107] Among the 11 significant GWAS associations (Table 2), seven were significant for callus formation and four were selected for further study based on their repeated occurrence across biological replicates (FIG. 2A- 2D)—Potri.003G018500, Potri.004G118700, Potri.009G066100, and Potri.018G014800 (p-value = 9.90E-08, 4.27E-07, 9.72E-08 and 3.83E-07, respectively).

Table 2. Chromosome location for single nucleotide polymorphisms associations with *Populus* callus phenotypes that exceeded a Bonferroni-adjusted significance. Chr. = Chromosome.

Chr.	Position (bp)	-log p-value	Callus Phenotype
3	2236428	9.90E-08	Formation
4	11078307	4.27E-07	Formation
6	22645602	9.89E-07	Formation
8	15806984	8.15E-07	Formation
9	6655602	9.72E-08	Formation
9	6655627	6.98E-07	Formation
18	1206685	3.83E-07	Formation
1	23447383	7.15E-07	Rating
6	23519186	1.61E-07	Rating
12	11117487	5.44E-07	Rating

15	1835111	6.53E-07	Rating

[0108] Potri.003G018500 encodes a SOK1 kinase containing a Testis-complex protein 11 motif that is highly expressed in early male flower development and co-expressed with Potri.015G078200—a gene of unknown function and Potri.016G082400—a kinesin motor protein-related protein (r^2 = 0.89 and r^2 = 0.86, respectively). Potri.004G118700, is a targeting protein for XKLP2 and is highly expressed in fully opened buds, immature leaves and root tips and is co-expressed with numerous genes including: Potri.002G080000—a mitotic-specific cyclin-B protein, Potri.016G033000—a cyclin G protein, Potri.017G081000—a tubulin, Potri.005G257500—a cyclin-dependent kinase and Potri.005G258300—a spindle checkpoint protein (r2 = 0.99, r2 = 0.98, r2 = 0.98 and r2 = 0.96, respectively). Potri.009G066100 encodes a mitogen-activated protein kinase (MAPK3) which is highly expressed in roots under high nitrogen and urea and is co-expressed with many genes including Potri.008G082100—a cell cycle control protein and Potri.016G009700—a scarecrow-like protein (r2 = 0.92 and r2 = 0.92, respectively). Finally, Potri.018G014800 is Chloroplast Nucleoid DNA-binding-related gene (CNDbr) which includes an aspartyl protease family protein domain, and is highly expressed in young leaves, stem nodes and internodes, root tips and in roots under high ammonia and nitrogen.

[0109] Among the 11 significant GWAS associations for callus rating, four were significant across all biological replicates (Table 2)—Potri.006G222700, Potri.008G208200, Potri.012G083800 and Potri.015G02360 (p-value = 1.61E-07, 8.15E-07, 5.44E-07 and 6.53E-07, respectively) (FIG. 2E – 2H). Potri.006G222700 is a gene of unknown function and is expressed in late development female flowers and dormant buds and is found in *Salix purpurea*, *Theobroma cacao*, and *Manihot exculenta* with >80% amino acid similarity. Potri.008G208200, a RALF-LIKE protein 22, is highly expressed in early developing male flowers. Potri.012G083800, a RPD3 histone deacetylase protein, is moderately expressed in dormant buds and is co-expressed with multiple genes including Potri.010G213700—a LEUKOCYTE RECEPTOR CLUSTER MEMBER 8 protein and Potri.009G137200 –a transcriptional coactivator CAPER RRM superfamily protein (r2 = 0.93 and r2 = 0.92, respectively). Potri.015G023600, a second gene of unknown function, is moderately

expressed in multiple tissues and is found in *S. purpurea* with >95% amino acid similarity. Interestingly, Potri.015G023600 contains a non-annotated RNA transcribed from the sequence between the 4th and 5th exons. This RNA is found in various tissues and contains no known domains or motifs. Potri.015G023600 is co-expressed with several zinc-finger proteins (r^2 = 0.85–0.91) and Potri.003G195400 encodes an ARMADILLO repeat-containing protein (r^2 = 0.92).

[0110] By examining the boxplots for each of the eight candidate genes, the inventors found that the rare allele (defined as the less frequent allele in the test population and depicted in the left column of each boxplot) for Potri.003G018500 and Potri.018G014800 lead to reduced callus formation, whereas the rare allele for Potri.004G118700 and Potri.009G066100 lead to increased callus formation (FIG. 2B and FIG. 2C). Interestingly, genotypes with homozygous rare alleles for Potri.003G018500 and Potri.018G014800 were not found in the tested population, suggesting that this condition may be lethal. Callus rating scores were all higher for the rare alleles for Potri.006G222700, Potri.008G208200, Potri.012G083800, and Potri.015G023600 (FIG 2E – 2H). Three of the candidate genes identified via the GWAS analysis for callus rating were associated with small frameshift INDELs.

Table 3. Sequences for the Eight Candidate Genes

Name	Nucleotide SEQ ID NO	Protein SEQ ID NO
Potri.009G066100	13	21
Potri.004G118700	14	22
Potri.003G018500	15	23
Potri.006G222700	16	24
Potri.018G014800	17	25
Potri.015G023600	18	26

Potri.008G208200	19	27
Potri.012G083800	20	28

Example 4: Callus formation validation

[0111] Callus formation *in vitro*, which is dependent on the plant source tissue and genetic background, varies with the concentration and ratios of added exogenous phytohormones to the plant media (Thorpe TA. *Journal of Plant Biotechnology*. 2000; 27: 245-58). The inventors therefore hypothesized that the *Populus* genotypes with the alleles associated with increased callus formation will consistently perform better in the different phytohormone treatments while those genotypes with the alleles associated with reduced callus formation will maintain reduced callus formation capacity due to their genetic background. To validate the initial callus formation experiment, and to leverage the information contained in the GWAS analyses, the inventors initiated an independent phytohormone treatment experiment based on six phytohormone combinations and seven genotypes that initially produced abundant callus with higher rating scores and contained the alleles associated with increased callus formation (i.e., BESC-18, BESC-233, BESC-823, GW-9795, GW-9877, GW-9920, and HOMB-21-2) and seven genotypes that had low occurrence of callus formation (i.e., BESC-100, BESC-106, BESC-352, BESC-856, BESC-89, GW-9904, and YALD-27-2). These genotypes were selected specifically because they contained high impact mutations (i.e., frameshifts or premature stop codons) predicted by genotype resequencing data using SnpEff in one or more of the significant loci identified in the GWAS results. Based on a one-tailed t-test, there were significant differences between the high callus producing genotypes and the low callus producing genotypes across all phytohormone combinations tested ($t = 3.70$, $p = 2.03E-3$). The abundant callus forming genotypes also had consistently higher callus rating scores across all phytohormone combinations, with the exception of genotypes BESC-18 and GW-9877.

Example 5: Callus formation genes co-expressed with genes related to cell differentiation and growth

[0112] Candidate genes from the GWAS were used as query in a co-expression of expressed genes in the Gene Atlas dataset (FIG. 3). The genome-wide co-expression network revealed that among the eight candidate genes, Potri.006G222700 and Potri.015G023600, were generally negatively co-expressed with their respective neighboring gene nodes in the co-expression network; while Potri.003G018500, Potri.012G083800, Potri.008G208200, Potri.009G066100, and Potri.004G118700, were overwhelmingly positively co-expressed with their respective neighboring gene nodes in the co-expression network. Potri.015G023600 and Potri.004G118700 were the only two candidate genes that were co-expressed with each other. These two genes were also consistently and commonly negatively or positively co-expressed with 332 other genes, respectively, including 35 putative transcriptional regulators, 44 protein kinases, and 10 cell-cycle-related genes. Potri.006G222700 and Potri.012G083800 were also in a reciprocal co-expression network involving 77 genes including KNUCKLES (KNU) that mediates the repression of WUSCHEL (WUS), a floral meristem determinacy gene (homologous to AT5g14010), a phosphoribosyl transferase family protein involved in cellular biosynthesis (homologous to AT2g35390) and two genes related to microtubule organization. A group of genes which co-expressed simultaneously with three candidate genes (Potri.004G118700, Potri.015G023600 and Potri.018G014800) were identified. Generally, Potri.015G023600 was negatively co-expressed with this set of genes, while Potri.004G118700 and Potri.018G014800 were positively co-expressed with this set of genes. This subnetwork involving co-expression with Potri.004G118700, Potri.015G023600 and Potri.018G014800 includes genes related to arrested embryo development (Potri.010G020600, homologous to AT3g06350 (MEE32)) and a microtubule-binding protein (Potri.005G033200, homologous to AT3g05330 (TANGLED1)). In addition, the co-expressed gene neighborhoods for Potri.004G118700 and Potri.015G023600 were enriched for cell cycle and microtubule formation genes, whereas the neighborhood between Potri.006G222700 and Potri.008G208200 contained quite a few transcription factors and genes of unknown

function. The distinctive positive and negative co-expression subnetworks (FIG. 3) strongly indicate tight orchestration of gene expression related to callus induction and repression.

[0113] Analysis of differential expression in *Arabidopsis thaliana* callus formation data from GEO (GSE29543) revealed that five orthologs to the candidate *Populus* genes were significantly differentially expressed in shoot callus formation in *Arabidopsis* (FIG. 4). Two of these orthologs (orthologous to Potri.004G118700 and Potri.012G083800) were upregulated during callus formation, while two alternate orthologs (Potri.009G066100 and Potri.003G018500) were downregulated during callus formation, again suggesting a network of genes that induce or repress callus formation.

[0114] Interestingly, orthologs of genes reported in *Arabidopsis* transgenesis experiments do occur in the co-expression network. Two LBD16 genes, Potri.005G221900 (orthologous to AT2g42430) and Potri.002G041200 (orthologous to AT2g23380), are negatively co-expressed with Potri.009G066100, along with Potri.002G044100 (orthologous to AT1g231970, LEC1, (Lotan T. et al., *Cell.* 1998; 93: 1195-205)), Potri.002G071200 (orthologous to AT5g49720, TSD1, (Frank M. et al., *The Plant J.* 2002; 29: 73-85)), Potri.005G188500 (orthologous to AT2g30580, BM1A, (Bratzel F. et al., *Curr Biol.* 2010; 20: 1853-9)), and Potri.011G054000 (orthologous to AT1g28300, LEC2, (Stone SL. et al., *Proc Natl Acad Sci U S A.* 2001; 98: 11806-11)) (FIG. 5). Potri.011G054000 is also negatively co-expressed with the candidate gene Potri.018G014800. A paralog of Potri.011G054000, Potri.004G045800 is positively co-expressed with both Potri.003G018500 and Potri.012G083800. Potri.007G012100 (orthologous to AT2g17950, WUS, (Zuo J. et al., *The Plant J.* 2002; 30: 349-59)) is positively co-expressed with Potri.012G083800 and negatively co-expressed with Potri.012G083800. Potri.005G140200 (orthologous to AT2g23380, CLF, (Chanvivattana Y. et al., *Development.* 2004; 131: 5263-76. PMID: 15456723)) was negatively co-expressed with Potri.015G023600.

[0115] Co-expression of the candidate genes from this disclosure, with orthologs of genes functionally validated in callus formation in the model plant *Arabidopsis*, provides support for the GWAS approach used to identify genes targets involved in this process in *Populus*. Based on both GWAS results and the co-expression analyses of the *Populus* candidate genes with the tested and published *Arabidopsis* transgene results, the inventors propose a

regulatory gene network for callus formation (FIG. 6A). Within this regulatory network, the gene encoding the transcription factor LEC2 containing the B3 domain showed either a positive or negative correlation to 4 of the 8 candidate GWAS genes identified in this study and may function as a hub gene control downstream expression of other transcription factors and kinases. Using a transient expression system in protoplast and quantitative RT-PCR (qRT-PCR), the inventors examined the ability of LEC2 to negatively regulate the expression of the Chloroplast Nucleoid DNA-binding-related gene (Potri.018G014800, CNDbr) and positively regulate the expression of SOK1, MAPK3 and RPD3 (Potri.003G018500, Potri.009G066100 and Potri.012G083800, respectively). The inventors found that when LEC2 was constitutively overexpressed, CNDbr was significantly repressed (FIG. 6B); however, the three positively regulated candidate GWAS genes which also showed low abundance in leaf tissue, were not detected in the transient expression assay (FIG. 6C).

Example 6

[0116] Completely defining the genetic components of cell de-differentiation and callus formation is of broad interest and application. Induction of pluripotency has implications in understanding orchestrated cell proliferation as well as normal tissue and organ development. Here, the inventors identified eight genes associated with callus formation or callus rating in *Populus*. These eight loci were distributed across the *Populus* genome on chromosomes III, IV, XI, VIII, IX, XII, XV and XVII. All eight loci have paralogs within the *Populus* genome that were the result of the Salicoid duplication event that occurred approximately 64 mya (Tuskan GA. et al., *Science*. 2006; 313: 1596-604). None of the paralogs showed significant association with callus formation or callus rating, suggesting that subfunctionalization may have occurred in these gene lineages. Among the eight significant associations, Potri.004G118700, Potri.008G082100 and Potri.009G066100 are co-expressed with genes annotated with functions related to cell division and cell differentiation; Potri.012G083800 is known to affect chromatin remodeling and an ortholog of Potri.008G208200 has been reported to be potentially involved in callus formation in sugarcane (Mingossi FB. et al., *Plant Mol Biol*. 2010; 73: 271-81). In total, the evidence

suggests that there are networks of genes that tightly regulate the cell division and cell differentiation cascade controlling callus formation.

[0117] Potri.004G118700, LEC2, may function as upstream regulator of several genes related to callus formation, including Potri.003G018500, Potri.009G066100, Potri.012G083800 and Potri.018G014800 (FIG. 6A). Specifically, Potri.003G018500, a SUPPRESSOR OF KINASE (SOK1) kinase, belongs to the STE20/SPS1/GC kinase family (Pfam PF05794), and there are multiple frameshift mutants at this locus within the GWAS population that cause a gained stop codon at position Chr03:2242626 bp. STE20 kinases in general are thought to regulate MAPK cascades, including several eukaryotic T-complex protein 11 (Tcp11)-related sequences. In yeast, a SOK1 protein, sharing sequence homology to a testis-specific mouse gene, suppresses cyclic AMP-dependent protein kinase mutants. Deletions in SOK1 in Saccharomyces can lead to an increase in lifespan of 15% or higher (Managbanag J. et al., *PloS One*. 2008; 3: e3802). The human homolog to mouse Tcp11 is only expressed in fertile adult testes and is thought to be important in sperm function and fertility (Ma Y. et al., *Mol Human Reprod*. 2002; 8: 24-31). The SOK1 or MST4 family of kinases are known signaling molecules for cell proliferation in multicellular organisms and have been implicated in cancer (Thompson BJ. et al., *J Cell Biol*. 2015; 210: 871-82).

[0118] Potri.009G066100 (MPK3), a member of a 21-gene family comprised of four groups and is orthologous to the defense-related gene AtMPK3 (Nicole M-C. et al., BMC Genomics.2006; 7: 223). MPKs are generally involved in directing cellular responses to a variety of stimuli, such as osmotic stress and heat shock, and they regulate cell functions, including proliferation, gene expression, differentiation, mitosis, cell survival, and apoptosis (Pearson G. et al., Endocr Rev. 2001; 22: 153-83). Interestingly, homologs of PtMPK3 in humans have been linked to various forms of cancer Chano T. et al., *Nat Genet*. 2002; 31: 285-9.

[0119] Potri.012G083800, a RPD3 HISTONE DEACETYLASE (RDP3), is present as a single copy gene in *Populus* and is found as co-orthologs in all sequenced plant genomes. Potri.012G083800 shares sequence similarity with two Arabidopsis RNA-MEDIATED TRANSCRIPTIONAL SILENCING 1 genes (At5g63110 and At5g35600). Histone

acetylation/deacetylation, in combination with various MAPKs, has been reported to play a role in plant defense (Hollender C. et al., *J Integr Plant Biol.* 2008; 50: 875-85). Histone deacetylases are primarily involved in regulating DNA transcription via modification of histone and chromatin structure and are often implicated in cellular processes such as cell growth, cell cycle and apoptosis. Posttranslational modification of histones has an intriguing but not fully understood role in human cancer (Cohen I. et al., *Genes & Cancer.* 2011; 2: 631-47). Moreover, histone acetylase PRZ1 in *Arabidopsis* acts as a transcriptional coactivator to modulate auxin effects on gene expression. Whereas auxin promotes formation of lateral roots in wild type, and both auxin and cytokinin are necessary for callus formation, prz1-1 mutants will produce callus in the presence of either auxin or cytokinin (Sieberer T. et al., *Curr Biol.* 2003; 13: 837-42; Anzola JM. et al., *Proc Natl Acad Sci U S A.* 2010; 107: 10308-13). In humans, histone acetylation/deacetylation has been linked to chronic myeloid leukemia. Histone deacetylase has also been reported to impact open chromatin and increase gene expression in pluripotent human cancer cells (Gaspar-Maia A. et al., *Nat. Rev. Mol. Cell Biol.* 2011; 12: 36). Potri.012G083800 appears to be a candidate gene for midstream control of signal transduction of cell proliferation in *Populus*.

[0120] Potri.018G014800 is a CHLOROPLAST NUCLEIOD DNA-BINDING-RELATED (CNDbr)/Aspartyl protease (Pfam00026) and variants within the GWAS contain a premature stop codon at position Chr18:1196058 bp that is associated with higher callus formation. In tobacco, CNDbr proteins have proteolytic activity and have been shown to bind to DNA (Diaz-Mendoza M. et al., *Genet Mol Biol.* 2016; 39: 329-38). CNDbr proteins have also been linked to leaf senescence (Kato Y. et al., *Planta.* 2005; 222: 643-51). In humans, proteins containing aspartyl protease domains includes the gene encoding Cathepsin D (CTSD), which has been implicated in breast cancer, and the gene encoding Cathepsin E (CTSE), which has been implicated in stomach cancer (Olson OC. et al., *Nat. Rev. Cancer.* 2015; 15: 712-29). Although annotated as a CHLOROPLAST NUCLEIOD DNA-BINDING-RELATED protein, Potri.018G014800 may primarily be related to general cell differentiation.

[0121] Several transcription factors, including LEC2, have been implicated in ectopic callus formation in Arabidopsis through transgenic studies. Ikeda and Ohme-Takagi (2014) have

implicated WIND1, WUS and TCP as genes that regulate callus formation (Ikeda M. et al., *Front Plant Sci.* 2014; 5). LATERAL ORGAN BOUNDARIES DOMAIN (LBD16) transcription factors have also been reported to induce callus formation in Arabidopsis (Fan M. et al., *Cell Res.* 2012; 22: 1169). And, ectopic overexpression of OPB4, another transcription factor, resulted in enhanced callus formation in *Arabidopsis* (Ramirez-Parra E. et al., *New Phytol.* 2017; 213: 1787-801). And, Iwase et al. (2013) successfully overexpressed AtWND1 to promote callus formation in phytohormone-free medium in tobacco (Iwase A. et al., *Plant Signal Behav.* 2013; 8: e27432). Surprisingly, none of the orthologs to the transcription factors described above showed significant associations with callus formation in *Populus* in the GWAS analysis. This difference could be related to species-specific differences in regulating and inducing callus, however it is more likely that these differences are due to experimental approach. The GWAS approach was conducted with no *a priori* assumptions concerning which genes were controlling callus formation, and thus identified only those loci that satisfied the statistical thresholds. The GWAS-identified genes, particularly, SOK1 and MAPK3, may be acting as checkpoints that monitor environmental queues, as discussed above. Such checkpoint genes could be overwhelmed by ectopic regulator expression in *Arabidopsis*. Human cell checkpoint genes are known to sense environmental signals such as ribonucleotide pools or oxygen tension and can lead to tumor formation if mutated (McDonald ER. et al., Ann Med. 2001; 33(2):113-22. PMID: 11327114). It may also be that the orthologs of those genes tested in Arabidopsis did not vary in the population and therefore were not detectable using GWAS approaches. However, there is substantial SNP variation across *Populus* orthologs of these Arabidopsis genes. It is also possible that the *Arabidopsis* orthologs are indeed influencing callus formation in *Populus*, but to a lesser degree than the genes identified in the GWAS test. Ectopic overexpression approaches may overwhelm innate gene and gene network influences on callus formation and impair de novo gene discovery. Ectopic overexpression of transcription factors likely leads to perturbations in multiple downstream phenotypes.

[0122] In support of *de novo* gene discovery via GWAS approaches, the Affymetrix resource developed for callus induction in *Arabidopsis* was examined and significant fold change was found in four orthologs of the eight candidate genes. Interestingly, the two

kinases discovered, Potri.003G018500 and Potri.009G066100, display significant negative fold change after 96 hours, while a gene with strong homology to human malignancy, Potri.008G208200, displayed a significant 4-fold change in expression after 96 hours. In further support of *de novo* discovery approaches, the eight genes reported here are significantly co-expressed with genes related to cytokinesis, tubulin, spindle function, and cell differentiation.

What is claimed is:

1. A genetically modified plant, plant cell or plant tissue, wherein the expression of a gene selected from the group consisting of Potri.003G018500, Potri.009G066100, Potri.012G083800, Potri.006G222700, Potri.008G208200, Potri.015G023600, Potri.004G118700, and Potri.018G014800, or a homolog thereof, is altered in the plant, plant cell or plant tissue.

2. The genetically modified plant, plant cell or plant tissue of claim 1, wherein the gene is selected from the group consisting of Potri.003G018500, Potri.009G066100, Potri.012G083800, Potri.006G222700, Potri.008G208200, and Potri.015G023600, and wherein the alteration comprises inactivation of the selected gene in the plant, plant cell or plant tissue, resulting in increased callus formation in the plant, plant cell or plant tissue.

3. The genetically modified plant, plant cell or plant tissue of claim 2, wherein the inactivation of the selected gene is achieved by introducing a nucleic acid inhibitor of the selected gene to the plant, plant cell or plant tissue.

4 The genetically modified plant, plant cell or plant tissue of claim 3, wherein the nucleic acid inhibitor is selected from the group consisting of an antisense RNA, a small interfering RNA, an RNAi, a microRNA, an artificial microRNA, and a ribozyme.

5. The genetically modified plant, plant cell or plant tissue of claim 2, wherein the inactivation of the selected gene is achieved by genome editing, which is achieved by a method selected from the group consisting of CRISPR/Cas system, Cre/Lox system, TALEN system, ZFNs system and homologous recombination.

6. The genetically modified plant, plant cell or plant tissue of claim 5, wherein the CRISPR-mediated genome editing comprises introducing into the plant a first nucleic acid encoding a Cas9 nuclease, a second nucleic acid comprising a guide RNA (gRNA), wherein said gRNA is specific to the selected gene.

7. The genetically modified plant, plant cell or plant tissue of claim 1, wherein the gene is selected from the group consisting of Potri.004G118700, and Potri.018G014800, and wherein the alteration comprises expressing in the plant an exogenous nucleic acid

comprising the selected gene, resulting in increased callus formation in the plant, plant cell or plant tissue.

8. The genetically modified plant, plant cell or plant tissue of claim 1, wherein the plant is selected from the group consisting of genera *Acer, Afzelia, Allium, Arabidopsis, Agrostis, Avena, Betula, Brassica, Capsicum, Citrullus, Cucumis, Eucalyptus, Fagus, Festuca, Fraxinus, Fragaria, Glycine, Gossypium, Hordeum, Ipomoea, Jatropha, Juglans, Lemna, Lolium, Malus, Manihot, Medicago, Micropus, Milium, Miscanthus, Nicotiana, Oryza, Pennisetum, Phalaris, Phleum, Picea, Pinus, Poa, Populus, Prunus, Quercus, Rosa, Salix, Solanum, Sorghum, Spinacia, Tectona, Trifolium, Triticum, Panicum, Saccharum, Setaria, Zea,* and *Zoysia.*

9. A method for increasing callus formation in a plant, plant cell or plant tissue comprising altering in a plant, plant cell or plant tissue the expression of a gene selected from the group consisting of Potri.003G018500, Potri.009G066100, Potri.012G083800, Potri.006G222700, Potri.008G208200, Potri.015G023600, Potri.004G118700, and Potri.018G014800, or a homolog thereof.

10. The method of claim 9, wherein the gene is selected from the group consisting of Potri.003G018500, Potri.009G066100, Potri.012G083800, Potri.006G222700, Potri.008G208200, and Potri.015G023600, and wherein the alteration comprises inactivation of the selected gene in the plant, plant cell or plant tissue, resulting in increased callus formation in the plant.

11. The method of claim 10, wherein the inactivation of the selected gene is achieved by introducing a nucleic acid inhibitor of the selected gene to the plant, plant cell or plant tissue.

12 The method of claim 11, wherein the nucleic acid inhibitor is selected from the group consisting of an antisense RNA, a small interfering RNA, an RNAi, a microRNA, an artificial microRNA, and a ribozyme.

13. The method of claim 10, wherein the inactivation of the selected gene is achieved by genome editing, which is achieved by a method selected from the group consisting of

CRISPR/Cas system, Cre/Lox system, TALEN system, ZFNs system and homologous recombination.

14. The method of claim 13, wherein the CRISPR-mediated genome editing comprises introducing into the plant a first nucleic acid encoding a Cas9 nuclease, a second nucleic acid comprising a guide RNA (gRNA), wherein said gRNA is specific to the selected gene.

15. The method of claim 9, wherein the gene is selected from the group consisting of Potri.004G118700, and Potri.018G014800, and wherein the alteration comprises expressing in the plant an exogenous nucleic acid comprising the selected gene, resulting in increased callus formation in the plant, plant cell or plant tissue.

16. The method of claim 9, wherein the plant, plant cell or plant tissue is selected from the group consisting of genera *Acer, Afzelia, Allium, Arabidopsis, Agrostis, Avena, Betula, Brassica, Capsicum, Citrullus, Cucumis, Eucalyptus, Fagus, Festuca, Fraxinus, Fragaria, Glycine, Gossypium, Hordeum, Ipomoea, Jatropha, Juglans, Lemna, Lolium, Malus, Manihot, Medicago, Micropus, Milium, Miscanthus, Nicotiana, Oryza, Pennisetum, Phalaris, Phleum, Picea, Pinus, Poa, Populus, Prunus, Quercus, Rosa, Salix, Solanum, Sorghum, Spinacia, Tectona, Trifolium, Triticum, Panicum, Saccharum, Setaria, Zea,* and *Zoysia.*

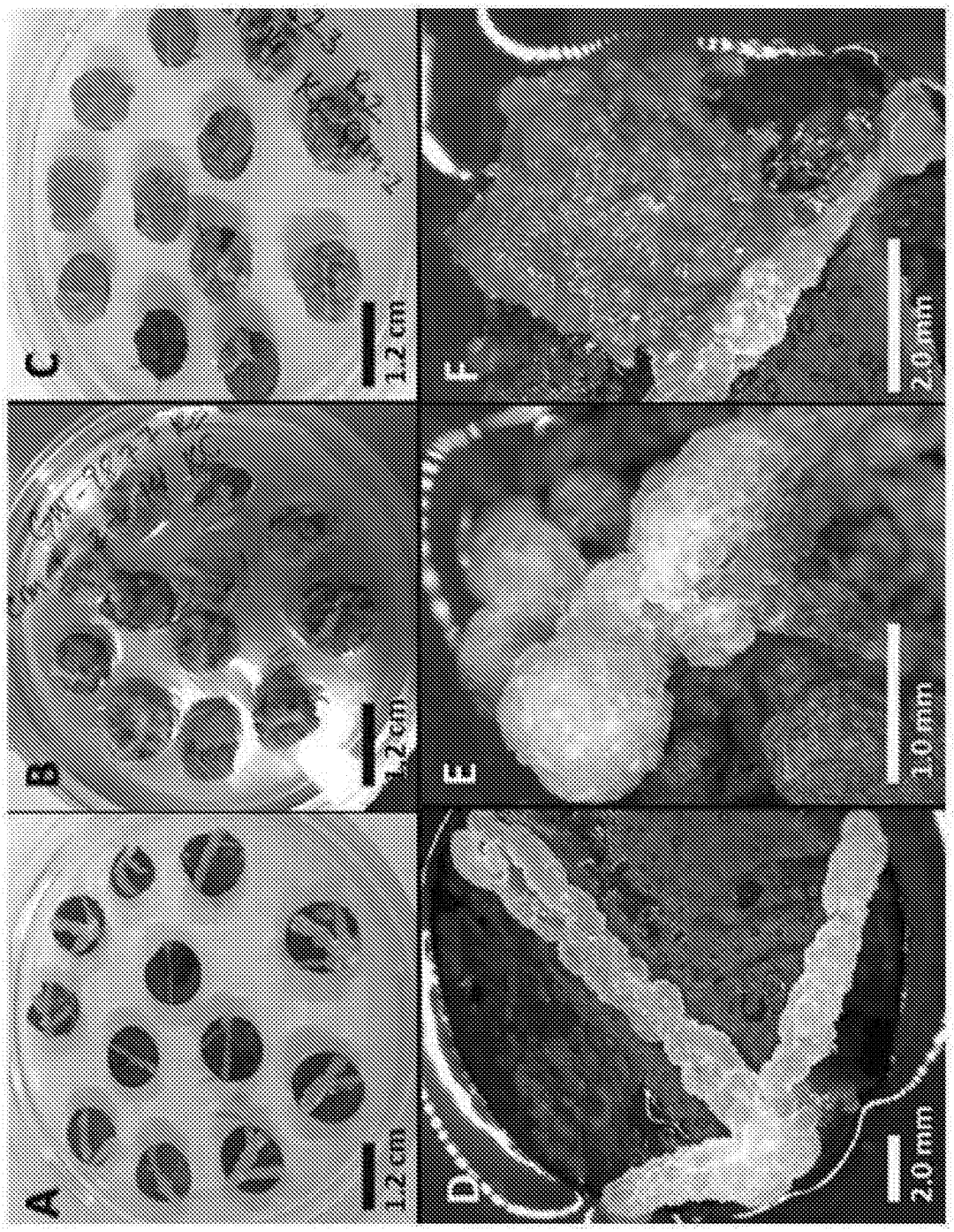

FIG. 1A – 1F

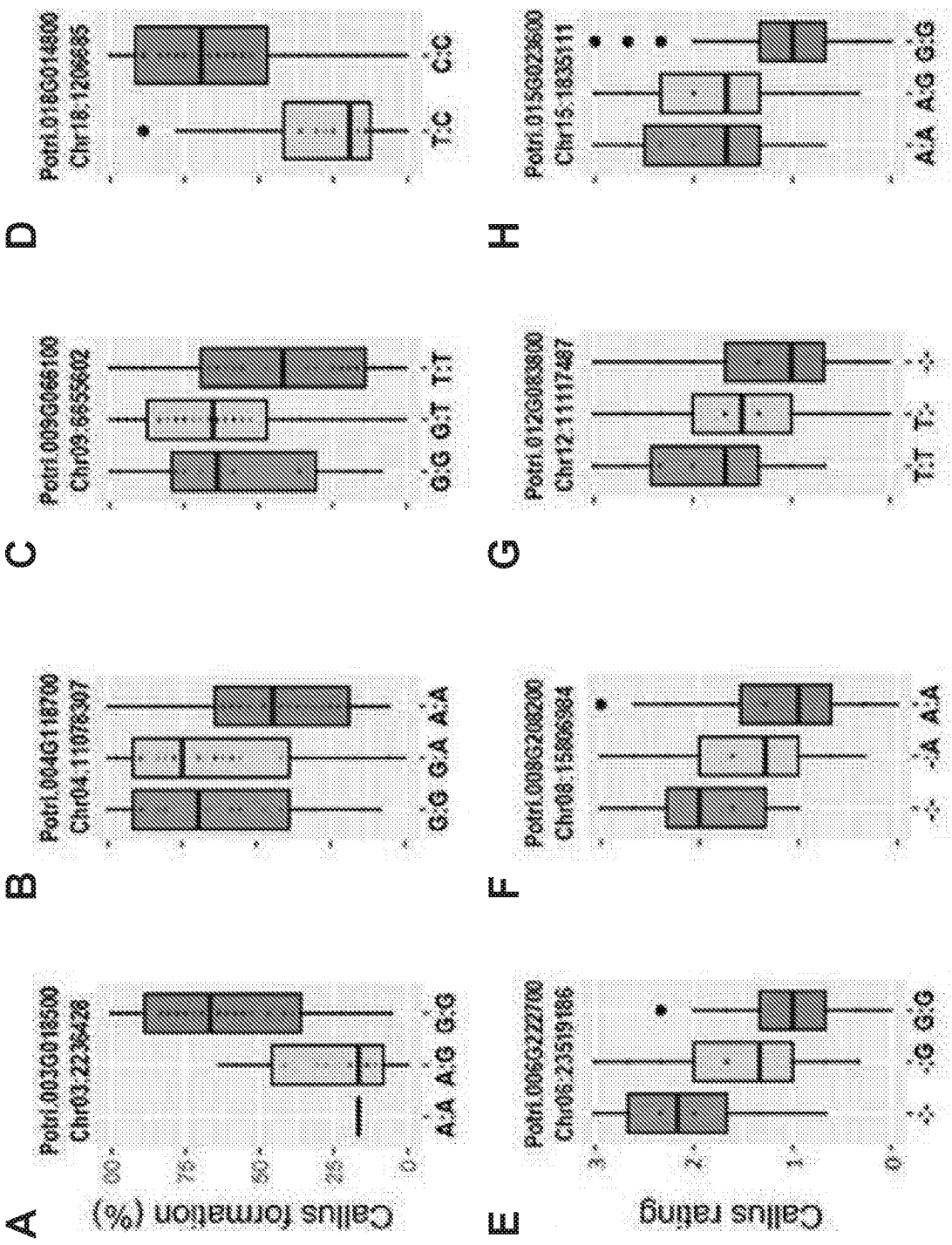
A
Potri.003G016500
Chr03:2228428
Callus formation (%)
100
75
50
25
0
A:A A:G G:G
B
Potri.004G118200
Chr04:11078307
G:G G:A A:A
C
Potri.009G066100
Chr09:6655602
G:G G:T T:T
D
Potri.018G014800
Chr18:1206585
T:C C:C
E
Potri.006G222700
Chr06:23519186
Callus rating
3
2
1
0
:G G:G
F
Potri.008G208200
Chr08:15806304
:A A:A
G
Potri.012G083800
Chr12:11117487
T:T T:
H
Potri.015G023500
Chr15:1835111
A:A A:G G:G

FIG. 2A – 2H

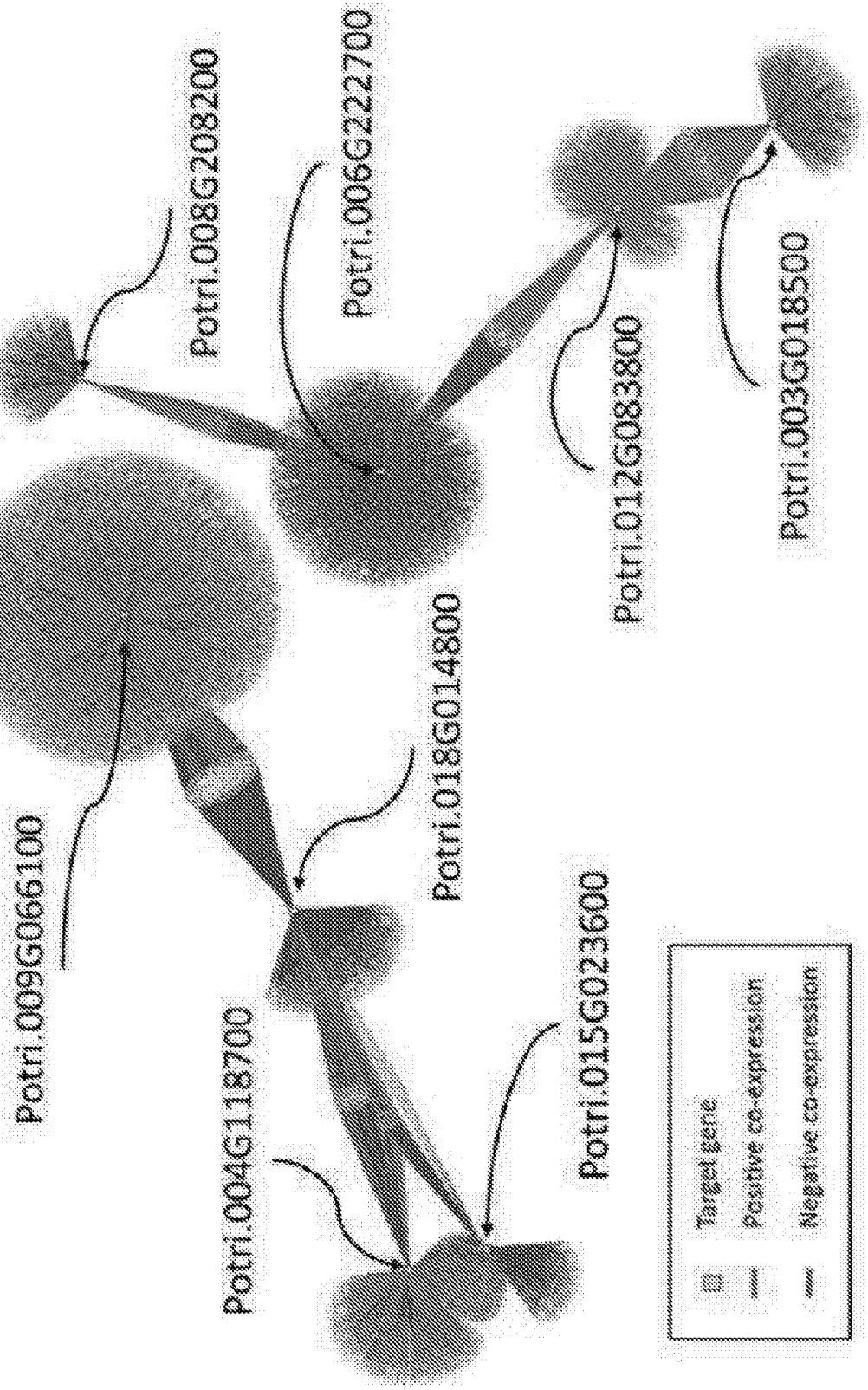
Potri.009G066100
Potri.008G208200
Potri.006G222700
Potri.018G014800
Potri.004G118700
Potri.015G023600
Potri.012G083800
Potri.003G018500
Target gene
Positive co-expression
Negative co-expression

FIG. 3

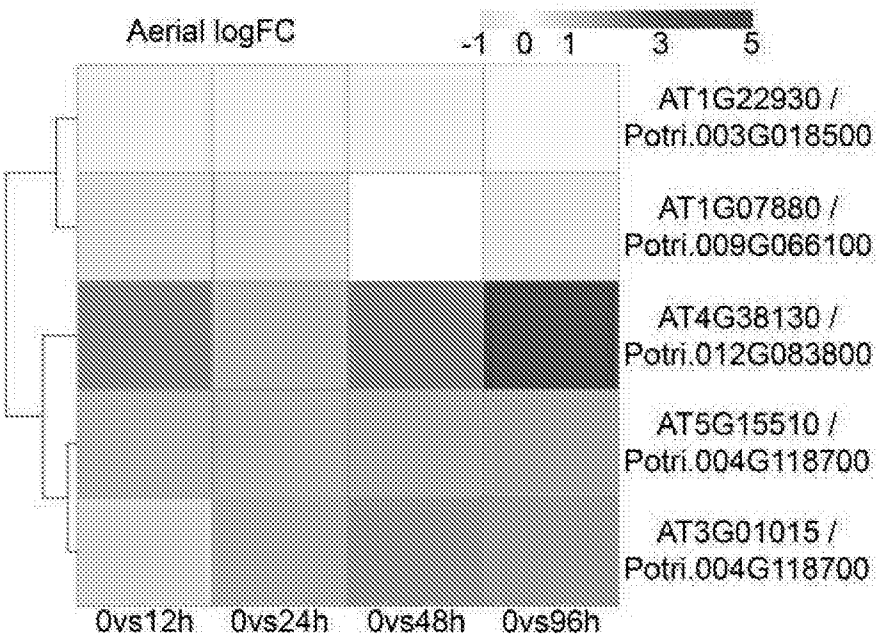
Aerial logFC
-1 0 1 3 5
AT1G22930 / Potri.003G018500
AT1G07880 / Potri.009G066100
AT4G38130 / Potri.012G083800
AT5G15510 / Potri.004G118700
AT3G01015 / Potri.004G118700
0vs12h
0vs24h
0vs48h
0vs96h

FIG. 4

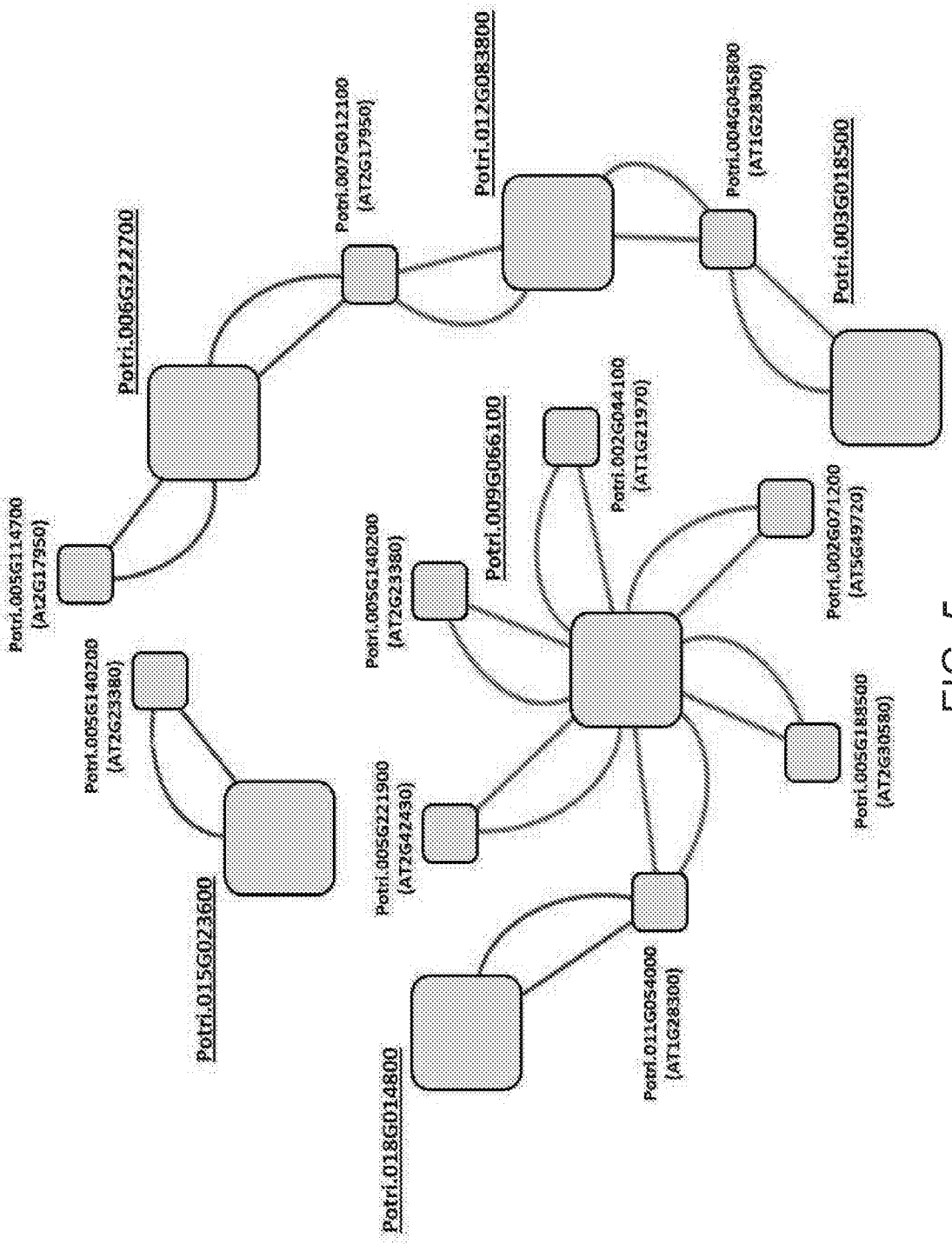
Potri.006G222700
Potri.007G012100
(AT2G17950)
Potri.012G083800
Potri.004G045800
(AT1G28300)
Potri.003G018500
Potri.005G114700
(AT2G17950)
Potri.009G066100
Potri.002G044100
(AT1G21970)
Potri.005G140200
(AT2G23380)
Potri.002G071200
(AT5G49720)
Potri.005G140200
(AT2G23380)
Potri.005G221900
(AT2G42430)
Potri.005G188500
(AT2G30580)
Potri.015G023600
Potri.011G054000
(AT1G28300)
Potri.018G014800

FIG. 5

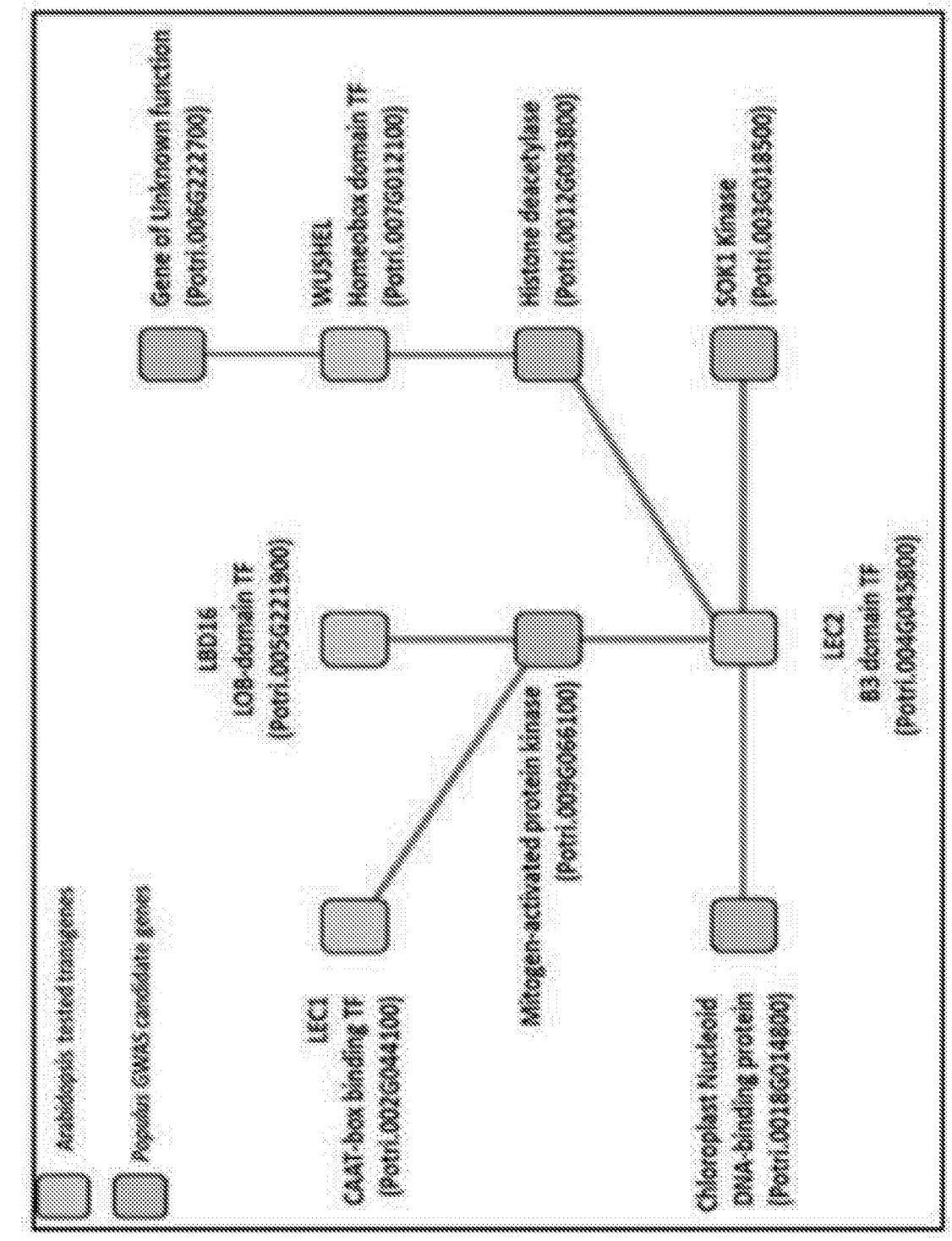
A
Arabidopsis tested transgenes
Populus GWAS candidate genes
LEC1
CAAT-box binding TF
(Potri.002G044100)
LBD16
LOB-domain TF
(Potri.005G221900)
Mitogen-activated protein kinase
(Potri.009G066100)
Chloroplast nucleoid
DNA-binding protein
(Potri.001G014200)
Gene of Unknown function
(Potri.006G222700)
WUSHEL
Homeobox domain TF
(Potri.007G012100)
Histone deacetylase
(Potri.001G023800)
SOK1 Kinase
(Potri.003G018500)
LEC2
B3 domain TF
(Potri.004G045800)

FIG. 6A

B

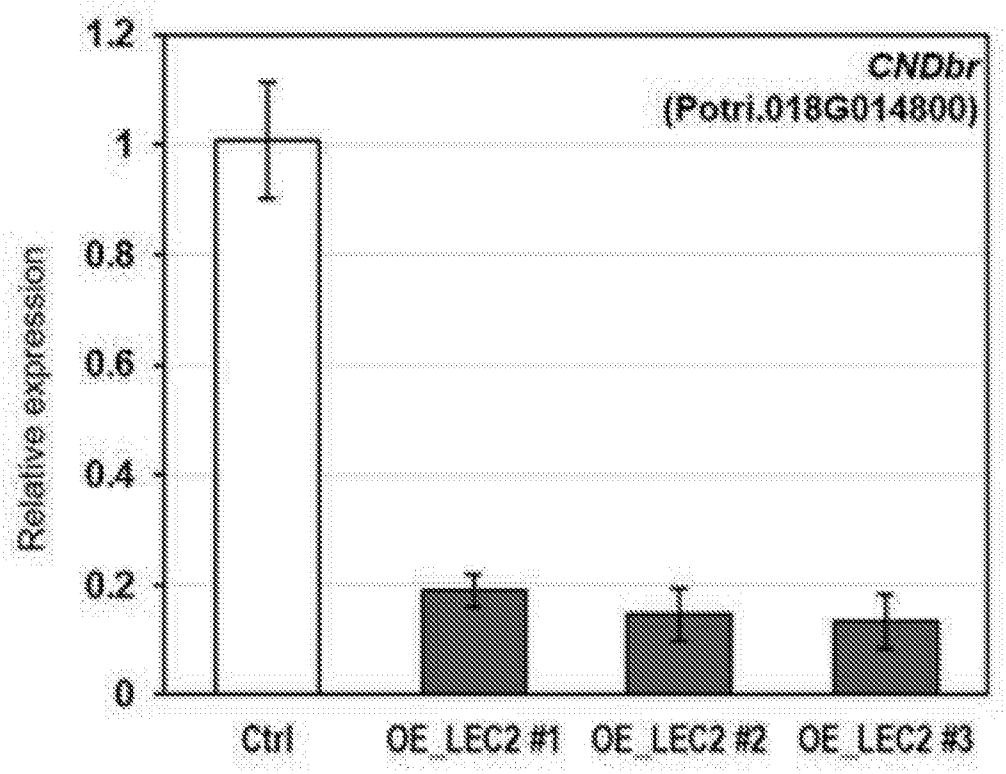
CNDbr
(Potri.018G014800)
Relative expression
1.2
1
0.8
0.6
0.4
0.2
0
Ctrl
OE_LEC2 #1
OE_LEC2 #2
OE_LEC2 #3

FIG. 6B

C

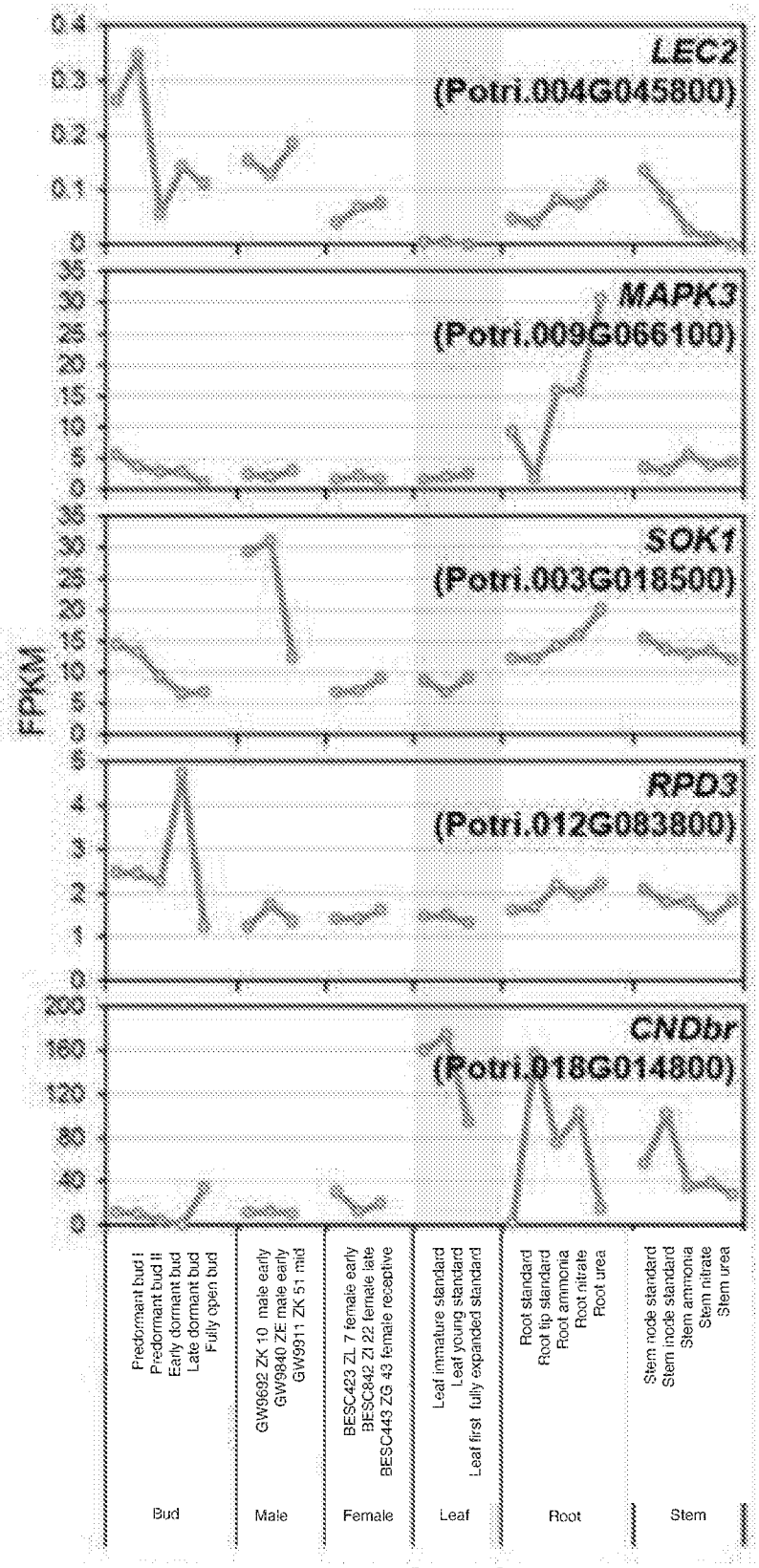
LEC2
(Potri.004G045800)
MAPK3
(Potri.009G066100)
SOK1
(Potri.003G018500)
RPD3
(Potri.012G083800)
CNDbr
(Potri.018G014800)
FPKM
Predormant bud I
Predormant bud II
Early dormant bud
Late dormant bud
Fully open bud
GW9692 ZK 10 male early
GW9840 ZE male early
GW9911 ZK 51 mid
BESC423 ZL 7 female early
BESC842 ZI 22 female late
BESC443 ZG 43 female receptive
Leaf immature standard
Leaf young standard
Leaf first fully expanded standard
Root standard
Root tip standard
Root ammonia
Root nitrate
Root urea
Stem node standard
Stem inode standard
Stem ammonia
Stem nitrate
Stem urea
Bud
Male
Female
Leaf
Root
Stem

FIG. 6C

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 19/44048

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:

   a. ☒ forming part of the international application as filed:

      ☒ in the form of an Annex C/ST.25 text file.

      ☐ on paper or in the form of an image file.

   b. ☐ furnished together with the international application under PCT Rule 13*ter*.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.

   c. ☐ furnished subsequent to the international filing date for the purposes of international search only:

      ☐ in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).

      ☐ on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).

2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 19/44048

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - A01H 5/00; A01H 9/00; C12N 15/87; C12N 15/82; C12N 15/00; C07H 21/04 (2019.01)
CPC - C07K 14/415; C12N 15/8216; C12N 15/00; C12N 15/82; C12N 15/52; C12N 15/113

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
See Search History Document

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
See Search History Document

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
See Search History Document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MEI et al., Tomato leaf curl Yunnan virus-encoded C4 induces cell division through enhancing stability of Cyclin D 1.1 via impairing NbSK.Eta-mediated phosphorylation in Nicotiana benthamiana. PLoS Pathog. 2018 Jan, Vol. 14(1):e1006789. PDF File: pg 1-27. Abstract; pg 2, para 3; pg 7, last para; pg 8, Fig 4A and Fig 4F; pg 9, last para; pg 10, up para, and Fig 5; and pg 15, para 2	1-6, 8-14, 16
A	WO 2017/173318 A1 (KANSAS STATE UNIVERSITY RESEARCH FOUNDATION et al.) 05 October 2017 (05.10.2017), Abstract; pg 1, ln 23-24; pg 3, ln 4-7 and 17-24; pg 7, ln 14-15; pg 8, ln 9-13; pg 10, ln 7-8 and 30-33; and pg 27, ln 2-3 and 18-19	1-6, 8-14, 16
A	WO 2017/117633 A1 (COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANISATION) 13 July 2017 (13.07.2017), Abstract; pg 147, ln 3-12; pg 151, ln 33-34; pg 152, ln 5-12, 14-25, and 29-30; and pg 153, ln 1-2	1-6, 8-14, 16
A	GenBank_JT915654, TSA: Hevea brasiliensis contig09485, mRNA sequence. GenBank Accession Number: JT915654. 04 February 2013. [online]. [Retrieved on 2019.09.27]. Retrieved from the Internet: < URL: https://www.ncbi.nlm.nih.gov/nuccore/JT915654 > Source, and Origin, 3896 nucleotides, the region between the region between nucleotides 546-3755	1-6, 8-14, 16
A	ANZOLA et al., Putative Arabidopsis Transcriptional Adaptor Protein (PROPORZ1) is required to modulate histone acetylation in response to auxin. Proc Natl Acad Sci U S A. 2010, Vol. 107(22), p.10308-13. Abstract; and pg 10311, col 1, para 2, and Fig 4	1-6, 8-14, 16
A	IKEUCHI et al., Plant Callus: Mechanisms of Induction and Repression. Plant Cell. 2013, Vol. 25(9), p. 3159-73. pg 3164, Fig 4	1-6, 8-14, 16

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
30 September 2019	27 NOV 2019

Name and mailing address of the ISA/US	Authorized officer:
Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-8300	Lee W. Young PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 19/44048

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	ENDO et al., Multigene Knockout Utilizing Off-Target Mutations of the CRISPR/Cas9 System in Rice. Plant Cell Physiol. 2015, Vol. 56(1), p. 41-7. Entire documentation, especially Abstract; pg 42, col 1, last para, and col 2, top para, and middle para; pg 43, Fig 1; pg 44, Fig 2; and pg 46, col 1, para 3	1-6, 8-14, 16
A, P	TUSKAN et al., Defining the genetic components of callus formation: A GWAS approach. PLoS One. 2018 Aug, Vol. 13(8):e0202519. PDF File: pg 1-18. Entire documentation, especially Abstract	1-6, 8-14, 16

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 19/44048

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:
This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Groups I+, claims 1-16, directed to a genetically modified plant, plant cell or plant tissue, wherein the expression of a gene is altered in the plant, plant cell or plant tissue; and a method for increasing callus formation in a plant, plant cell or plant tissue comprising altering in a plant, plant cell or plant tissue the expression of a gene. The gene will be searched to the extent that the gene encompasses Potri.003G018500, represented by SEQ ID NO: 15 [Specification: pg 31, Table 3 - 'Potri.003G018500...Nucleotide SEQ ID NO: 15']. It is believed that claims 1-2 (in part), 3-6, 8, 9-10 (in part), 11-14, 16 encompass this first named invention, and thus these claims will be searched without fee to the extent that they encompass Potri.003G018500, represent by SEQ ID NO: 15.
**********Continued in the extra sheet**********

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-2 (in part), 3-6, 8, 9-10 (in part), 11-14, 16, limited to gene Potri.003G018500, represent by SEQ ID NO: 15

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 19/44048

Continuation of:
Box No III (unity of invention is lacking)

(Continuation of Groups I+) Additional gene(s) will be searched upon the payment of additional fees. Applicants must specify the claims that encompass any additionally elected gene(s). Applicants must further indicate, if applicable, the claims which encompass the first named invention, if different than what was indicated above for this group. Failure to clearly identify how any paid additional invention fees are to be applied to the "+" group(s) will result in only the first claimed invention to be searched. An exemplary election would be the gene encompasses Potri.009G066100, represented by SEQ ID NO: 13 [Specification: pg 31, Table 3 - 'Potri.009G066100 ...Nucleotide SEQ ID NO: 13'] [claims: 1-2 (in part), (3-6, 8), 9-10 (in part), (11-14, 16)]. [Please Note, each gene will be searched based on 100% identical to the sequence of the associated SEQ ID NO disclosed (Specification: pg 31-32, Table 3), because the genetically modified plant is generated using nucleic acid sequence based methods, and it will be unpredictable by using nucleic acid sequences with an undefined percentage of homology between 40%-99%, because the definition of "homolog" in the Specification is too broad, which will include unrelated genes, and function determination can not be done in this ISR (Specification: para [0029] - 'term "homolog" means a gene related to a second gene by descent from a common ancestral DNA sequence, ...has a certain degree of homology, i.e., sequence identity (at least 40%, at least 60%, at least 65% particularly preferred at least 66%, 68%, 70%, 75%, 80%, 86%, 88%, 90%, 92%, 95%, 97% or 99% sequence identity). A "homolog" furthermore means that the function is equivalent to the function of the original gene')].

The inventions listed as Groups I+ do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

Special Technical Feature
Among Groups I+, each gene represented by an associated SEQ ID NO represents a structurally different nucleotide sequence from all others.

Common Technical Features
The inventions of Groups I+ share the technical features of a genetically modified plant, plant cell or plant tissue, wherein the expression of a specified gene is altered in the plant, plant cell or plant tissue (part of claim 1); and/or
-- a method for increasing callus formation in a plant, plant cell or plant tissue comprising altering in a plant, plant cell or plant tissue the expression of a specified gene (part of claim 9).

However, these shared technical features do not represent a contribution over prior art as being anticipated by an article entitled 'Tomato leaf curl Yunnan virus-encoded C4 induces cell division through enhancing stability of Cyclin D 1.1 via impairing NbSK.Eta-mediated phosphorylation in Nicotiana benthamiana' by MEI et al. (hereinafter 'Mei'; PLoS Pathog. 2018 Jan, Vol. 14(1):e1006789. PDF File: pg 1-27); or as being anticipated by an article entitled 'Putative Arabidopsis Transcriptional Adaptor Protein (PROPORZ1) is required to modulate histone acetylation in response to auxin' by ANZOLA et al. (hereinafter 'Anzola'; Proc Natl Acad Sci U S A. 2010, Vol. 107(22), p.10308-13) as follows:

Mei discloses a genetically modified plant, plant cell or plant tissue (Abstract - 'NbSK.Eta-RNAi, ... transgenic N. benthamiana plants ...on callus-like tissue formation'; pg 15, para 2 - 'genetic evidence:...(2) NbSK.Eta-silenced N. benthamiana plants'; pg 2, para 3 - 'AtSK.Eta is a homologue of mammalian glycogen synthase kinase 3 (GSK3)'; pg 7, lower para - 'stable NbSK.Eta knock-down transgenic N. benthamiana (NbSK.Eta-RNAi)'; pg 10, up para - 'N. benthamian infiltrated with NbSK.Eta-interaction compromised mutants, ...NbSK.Eta is essential for the induction of cell division in N. benthamiana'; pg 10, Fig 5),
--wherein the expression of a specified gene is altered in the plant, plant cell or plant tissue (Abstract; pg 7, last para - 'silenced NbSK.Eta gene in tobacco plants using a TRV-based vector. Real-time quantitative PCR (qRT-PCR) analysis revealed that the transcripts of NbSK.Eta were significantly reduced (over 75%) in NbSK.Eta-silenced plants compared with the amount in mock plants inoculated with vector control (TRV-GFP) (Fig 4A)'; pg 8, Fig 4A; pg 8, Fig 4, Legend - '(F)...qRT-PCR analysis of NbSK.Eta gene expression in WT and NbSK.Eta-RNAi transgenic plants (right)'; Fig 4F).

Mei further discloses a method for increasing callus formation in a plant or plant tissue comprising altering in a plant or plant tissue the expression of a specified gene (pg 7, last para - 'we silenced NbSK.Eta gene in tobacco plants using a TRV-based vector. Real-time quantitative PCR (qRT-PCR) analysis revealed that the transcripts of NbSK.Eta were significantly reduced (over 75%) in NbSK.Eta-silenced plants compared with the amount in mock plants inoculated with vector control (TRV-GFP) (Fig 4A)'; pg 8, Fig 4A; pg 9, last para - 'Silencing NbSK.Eta in N. benthamiana through TRV vector could promote formation of callus-like tissues (Fig 5E), ... NbSK.Eta-RNAi plants showed the callus-like tissues in their leaves and roots (Fig 5F and 5G). ... NbSK.Eta-RNAi plants were obviously larger than that of wild-type plants'; pg 10, Fig 5).

Furthermore, Anzola also discloses a method for increasing callus formation in a plant comprising altering in a plant the expression of a specified gene (pg 10311, col 1, para 2 - 'we generated KRP silencing lines, to affect expression of multiple KRP genes ... altered KRP expression interfere with plant morphogenesis ... a significant fraction of KRP silencer plants spontaneously started to generate callus-like tissue', wherein 'we generated KRP silencing lines,... KRP genes ... altered KRP expression interfere with plant' is 'altering in a plant the expression of a specified gene', and wherein 'we generated KRP silencing lines, ... KRP genes ... altered KRP expression ? KRP silencer plants spontaneously started to generate callus-like tissue' is 'a method for increasing callus formation in a plant comprising altering in a plant the expression of a specified gene', see the quotations that follow; Abstract - 'gene expression of Arabidopsis KIP RELATED PROTEIN (KRP) CDK inhibitor genes in response to auxin'; pg 10311, Fig 4, Legend - 'Analysis of KRP overexpression and silencing lines. (A Left) ...leaves of a KRP silencer control plant (c; 22 DAG). ... (Right) Microcallus-formation on a leaf of silencer line 2.7'; Fig 4A; Please also see an article entitled 'Plant Callus: Mechanisms of Induction and Repression' by Ikeuchi et al.; Plant Cell. 2013, Vol. 25(9), p. 3159-73, for the association of the effects of auxin and KRP on plant callus formation: pg 3164, Fig 4, Legend - 'Auxin-induced callus formation. ... auxin-dependent repression of CDK inhibitors, KRP2, KRP3, and KRP7'; pg 3164, Fig 4A).

**********Continued in the next extra sheet**********

Continuation of:
The previous extra sheet - Box No III (unity of invention is lacking)

Without a shared special technical feature, the inventions lack unity with one another.

Groups I+ therefore lack unity under PCT Rule 13 because they do not share a same or corresponding special technical feature.

Note:

I) For claims 1 and 9, each gene will be searched based on 100% identical to the sequence of the associated SEQ ID NO disclosed (Specification: pg 31-32, Table 3), because the genetically modified plant is generated using nucleic acid sequence based methods, and it will be unpredictable by using nucleic acid sequences with an undefined percentage of homology between 40%-99%, because the definition of "homolog" in the Specification is too broad, which will include unrelated genes, and function determination can not be done in this ISR (Specification: para [0029] - 'term "homolog" means a gene related to a second gene by descent from a common ancestral DNA sequence, ...has a certain degree of homology, i.e., sequence identity (at least 40%, at least 60%, at least 65% particularly preferred at least 66%, 68%, 70%, 75%, 80%, 86%, 88%, 90%, 92%, 95%, 97% or 99% sequence identity). A "homolog" furthermore means that the function is equivalent to the function of the original gene').