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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 April 2017 (13.04.2017)

WIPO | PCT

(10) International Publication Number
WO 2017/062618 A1

(51) International Patent Classification:
C07H 21/04 (2006.01) *C12N 15/82* (2006.01)
C12N 5/10 (2006.01) *C12Q 1/68* (2006.01)

(21) International Application Number:
PCT/US2016/055749

(22) International Filing Date:
6 October 2016 (06.10.2016)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/237,822 6 October 2015 (06.10.2015) US

(71) Applicants: **IOWA STATE UNIVERSITY RESEARCH FOUNDATION, INC.** [US/US]; 2121 State Avenue, Ames, IA 50014 (US). **UNIVERSITY OF BONN** [DE/DE]; Regina-Pacis-Weg 3, Bonn 53113 (DE).

(72) Inventors: **SCHNABLE, Parick, S.**; 2035B Roy J. Carver Co-Lab, Iowa State University, Ames, IA 50011-3650 (US). **LIU, Sanzhen**; 4022 Throckmorton, Department of Plant Pathology, Kansas State University, Manhattan, KS 66506-5502 (US). **LI, Li**; Chinese Academy of Agriculture Science, Crop Science Institute, Key Building 601, Zhongguancun South Street #12, Beijing 100081 (CN). **HOCHHOLDINGER, Frank**; University of Bonn, Friedrich-Ebert-Allee 144, Bonn 53113 (DE). **HEY, Stefan**; University of Bonn, Friedrich-Ebert-Allee 144, Bonn 53113 (DE).

(74) Agents: **ANDERSON, Joseph, P., III** et al.; Ballard Spahr LLP, 999 Peachtree Street, Suite 1000, Atlanta, GA 30309 (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, JP, KE, KG, 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).

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

Published:

— *with international search report (Art. 21(3))*

[Continued on next page]

(54) Title: PLANTS WITH IMPROVED AGRONOMIC CHARACTERISTICS

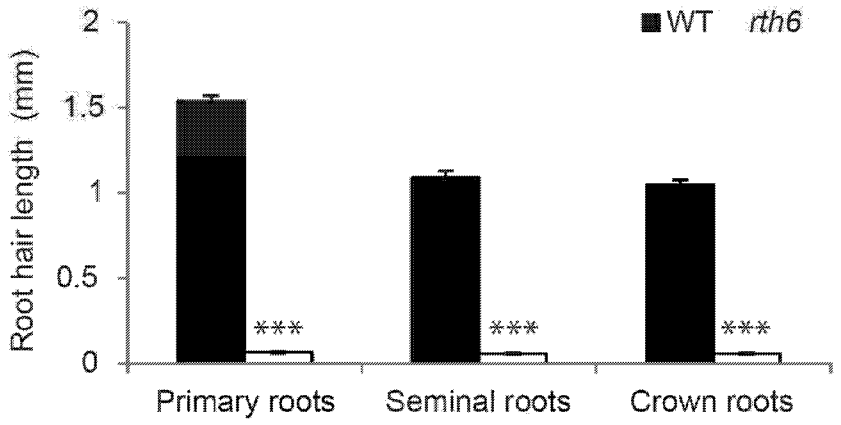


FIG. 1B

(57) Abstract: Disclosed herein is a method for increasing drought tolerance in a plant. The method includes introducing into the plant an expression construct that includes a heterologous promoter operably linked to a polynucleotide that encodes a RTH6 family member. The method can be performed in monocots and dicots. Also, disclosed herein are plants and expression cassettes that include recombinant DNA that contains a heterologous promoter operably linked to polynucleotides that encodes an RTH5 family member.

WO 2017/062618 A1

— *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))*

PLANTS WITH IMPROVED AGRONOMIC CHARACTERISTICS

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application claims benefit of the filing date of U.S. Provisional Application No. 62/237,822, which was filed on October 6, 2015. The content of this earlier filed application is hereby incorporated by reference herein in its entirety.

SEQUENCE LISTING

[0002] The sequence listing submitted October 6, 2016 as a text file named "36446.0203P1_Sequence_Listing," created on August 17, 2016, and having a size of 570,020 bytes is hereby incorporated by reference pursuant to 37 C.F.R. § 1.52(e)(5).

FIELD

[0003] The field of disclosure relates to plant breeding and genetics and, in particular, relates to expression constructs useful in plants for conferring improved agronomic traits.

BACKGROUND

[0004] Improving agronomic traits in crop plants is beneficial to farmers. Several factors influence crop yield. Abiotic stress is the primary cause of crop loss worldwide, causing average yield losses of more than 50% for major crops. Among the various abiotic stresses, drought is a major factor that limits crop productivity worldwide. Exposure of plants to a water-limiting environment during various developmental stages appears to activate various physiological and developmental changes. Molecular mechanisms of abiotic stress responses and the genetic regulatory networks of drought stress tolerance have been studied.

[0005] Natural responses to abiotic stress vary among plant species and among varieties and cultivars within a plant species. Certain species, varieties or cultivars are more tolerant to abiotic stress such as drought than others. Transgenic approaches are needed for improving drought tolerance in crop plants.

SUMMARY

[0006] In accordance with the purpose(s) of the invention, as embodied and broadly described herein, the invention, in one aspect, relates to compositions and methods to increase drought tolerance and to increase nutrient uptake in plants. In various aspects, compositions comprise recombinant RTH6 family polynucleotides, RTH6 polypeptides, DNA constructs, expression cassettes, plants, parts of plants, and seeds.

[0007] Expression of RTH6 polypeptides increases root hair formation and growth. Increased expression of RTH6 polypeptides in the root of a plant thus results in increased drought tolerance, increased nutrient uptake, and decreased soil erosion. In a further aspect,

the methods of the invention comprise increasing the expression of an RTH6 polypeptide in a plant of interest, or in a specific part of a plant. Any method for increasing the expression of the RTH6 polypeptide is encompassed. That is plants can be transformed with a DNA construct comprising an RTH6 polynucleotide operably linked with a heterologous promoter that drives expression in plant roots. Alternatively, expression levels of the endogenous RTH6 polypeptide in the plant can be increased by methods available in the art to enhance the expression of endogenous genes. Expression constructs comprising an RTH6 polynucleotide as well as plants and seed having increased levels of an RTH6 polypeptide are provided.

[0008] Decreased expression of RTH6 polypeptides in the root of a plant can result in increased resistance to infection by a root pathogen. In a further aspect, the methods of the invention comprise decreasing the expression of an RTH6 polypeptide in a plant of interest, or in a specific part of a plant. Any method for decreasing the expression of the RTH6 polypeptide is encompassed. That is plants can be transformed with a DNA construct comprising an RTH6 polynucleotide operably linked with a heterologous promoter that drives expression of a silencing construct, *e.g.* a siRNA construct comprising an RTH6 polynucleotide, in plant roots. Alternatively, expression levels of the endogenous RTH6 polypeptide in the plant can be decreased by methods available in the art to decrease the expression of endogenous genes. Expression constructs comprising an RTH6 polynucleotide as well as plants and seed having decreased levels of an RTH6 polypeptide are provided.

[0009] Disclosed are methods for improving at least one agronomic characteristic in a plant, the method comprising introducing into the plant an expression construct, the construct comprising a promoter operably linked to a polynucleotide, wherein the polynucleotide comprises a nucleotide sequence selected from the group consisting of: (a) a nucleotide sequence comprising SEQ ID NO:2; (b) a nucleotide sequence encoding the amino acid sequence set forth in SEQ ID NO:3; (c) a nucleotide sequence having at least 90% sequence identity to SEQ ID NO:2, wherein the polynucleotide encodes a polypeptide having ß-glycan synthase activity; and (d) a nucleotide sequence encoding an amino acid sequence having at least 90% sequence identity to SEQ ID NO:3, wherein the polynucleotide encodes a polypeptide having ß-glycan synthase activity; wherein the at least one agronomic characteristic is selected from the group consisting of increased root hair formation, increased root growth, increased drought tolerance, enhanced nutrient uptake, increased yield, enhanced nutrient uptake, and combinations thereof.

[0010] Also disclosed are methods for increasing drought tolerance in a plant, the method comprising introducing into the plant an expression construct, the construct comprising a heterologous root-preferred promoter operably linked to a polynucleotide, wherein the polynucleotide comprises a nucleotide sequence selected from the group consisting of: (a) a nucleotide sequence comprising SEQ ID NO:2; (b) a nucleotide sequence encoding the amino acid sequence set forth in SEQ ID NO:3; (c) a nucleotide sequence having at least 90% sequence identity to SEQ ID NO:2, wherein the polynucleotide encodes a polypeptide having ß-glycan synthase activity; and (d) a nucleotide sequence encoding an amino acid sequence having at least 90% sequence identity to SEQ ID NO:3, wherein the polynucleotide encodes a polypeptide having ß-glycan synthase activity.

[0011] Also disclosed are methods for increasing nutrient uptake in a plant, the method comprising introducing into the plant an expression construct, the construct comprising a heterologous root-preferred promoter operably linked to a polynucleotide, wherein the polynucleotide comprises a nucleotide sequence selected from the group consisting of: (a) a nucleotide sequence comprising SEQ ID NO:2; (b) a nucleotide sequence encoding the amino acid sequence set forth in SEQ ID NO:3; (c) a nucleotide sequence having at least 90% sequence identity to SEQ ID NO:2, wherein the polynucleotide encodes a polypeptide having ß-glycan synthase activity; and (d) a nucleotide sequence encoding an amino acid sequence having at least 90% sequence identity to SEQ ID NO:3, wherein the polynucleotide encodes a polypeptide having ß-glycan synthase activity.

[0012] Also disclosed are methods for reducing soil erosion in a crop field, the method comprising the step of planting a crop field with a plant comprising an expression construct, the construct comprising a heterologous root-preferred promoter operably linked to a polynucleotide, wherein the polynucleotide comprises a nucleotide sequence selected from the group consisting of: (a) a nucleotide sequence comprising SEQ ID NO:2; (b) a nucleotide sequence encoding the amino acid sequence set forth in SEQ ID NO:3; (c) a nucleotide sequence having at least 90% sequence identity to SEQ ID NO:2, wherein the polynucleotide encodes a polypeptide having ß-glycan synthase activity; and (d) a nucleotide sequence encoding an amino acid sequence having at least 90% sequence identity to SEQ ID NO:3, wherein the polynucleotide encodes a polypeptide having ß-glycan synthase activity.

[0013] Also disclosed are methods for increasing infection of a plant with a beneficial microorganism, said method comprising introducing into said plant an expression construct, said construct comprising a promoter that drives expression in a plant root operably linked to a polynucleotide, wherein said polynucleotide comprises a nucleotide sequence selected from

the group consisting of: (a) a nucleotide sequence comprising SEQ ID NO:2; (b) a nucleotide sequence encoding the amino acid sequence set forth in SEQ ID NO:3; (c) a nucleotide sequence having at least 90% sequence identity to SEQ ID NO:2, wherein the polynucleotide encodes a polypeptide having ß-glycan synthase activity; and (d) a nucleotide sequence encoding an amino acid sequence having at least 90% sequence identity to SEQ ID NO:3, wherein the polynucleotide encodes a polypeptide having ß-glycan synthase activity.

[0014] Also disclosed are methods for increasing resistance of a plant to a root pathogen, the method comprising the step of reducing the expression of a *rth6* gene family member or inhibiting the activity a *rth6* polypeptide.

[0015] Also disclosed are methods for identifying a first corn plant or germplasm that displays tolerance, improved tolerance, or susceptibility to drought; the method comprising detecting in the first corn plant or germplasm at least one allele of a quantitative trait locus that is associated with the tolerance, improved tolerance, or susceptibility; wherein the quantitative locus is a chromosomal interval located within about 10 kbp of an interval at 104,616,335 - 104,622,459 of chromosome 1 (SEQ ID NO:1).

[0016] Also disclosed are seeds of a plant produced by any of the disclosed methods.

[0017] Also disclosed are plants comprising an expression construct comprising a heterologous root-preferred promoter operably linked to a polynucleotide, wherein the polynucleotide comprises a nucleotide sequence selected from the group consisting of: (a) a nucleotide sequence comprising SEQ ID NO:2; (b) a nucleotide sequence encoding the amino acid sequence set forth in SEQ ID NO:3; (c) a nucleotide sequence having at least 90% sequence identity to SEQ ID NO:2, wherein the polynucleotide encodes a polypeptide having ß-glycan synthase activity; and (d) a nucleotide sequence encoding an amino acid sequence having at least 90% sequence identity to SEQ ID NO:3, wherein the polynucleotide encodes a polypeptide having ß-glycan synthase activity.

[0018] Also disclosed are expression cassettes comprising a polynucleotide operably linked to a heterologous promoter that drives expression of the polynucleotide in a root of a plant, wherein the polynucleotide comprises a nucleotide sequence selected from the group consisting of: (a) a nucleotide sequence comprising SEQ ID NO:2; (b) a nucleotide sequence encoding the amino acid sequence set forth in SEQ ID NO:3; (c) a nucleotide sequence having at least 90% sequence identity to SEQ ID NO:2, wherein the polynucleotide encodes a polypeptide having ß-glycan synthase activity; and (d) a nucleotide sequence encoding an amino acid sequence having at least 90% sequence identity to SEQ ID NO:3, wherein the polynucleotide encodes a polypeptide having ß-glycan synthase activity.

[0019] Also disclosed are methods for selecting an allelic variant of *rth6* in a maize plant, the method comprising the steps of: (a) obtaining a population of maize plants, wherein the maize plants exhibit an alteration of at least one agronomic characteristic; wherein the at least one agronomic characteristic is selected from the group consisting of increased root hair formation and growth, increased drought tolerance, and enhanced nutrient uptake; (b) evaluating allelic variations with respect to the polynucleotide sequence encoding a protein comprising SEQ ID NO:3, or in the genomic region that regulates the expression of the polynucleotide encoding the protein; (c) associating allelic variations with the alteration of at least one agronomic characteristic; and (d) selecting an allelic variant that is associated with the alteration of at least one agronomic characteristic.

[0020] Also disclosed are methods for selecting a first maize plant or a first maize germplasm that has one or more beneficial alleles of *rth6*, the method comprising: (a) screening a plurality of maize plants or a plurality of maize germplasm for at least one polymorphism within a marker locus, wherein the marker locus is: (i) a first polynucleotide having at least 90% and less than 100% nucleotide sequence identity with SEQ ID NO: 1 or 2; or (ii) a second polynucleotide encoding a polypeptide comprising an amino acid sequence having at least 90% and less than 100% sequence identity to SEQ ID NO: 3, wherein expression of the first or second polynucleotide in a maize plant results in a phenotype comprising an alteration of at least one agronomic characteristic when compared to a control maize plant; wherein the at least one agronomic characteristic is selected from the group consisting of increased root hair formation and growth, increased drought tolerance, and enhanced nutrient uptake; and wherein the control maize plant comprises: a polynucleotide having the nucleotide sequence of SEQ ID NO:1 or 2; or a polynucleotide encoding the amino acid sequence of SEQ ID NO: 3; (b) identifying a first maize plant or a first maize germplasm comprising the at least one polymorphism of the marker locus; and (c) selecting the first maize plant or first maize germplasm of step (b).

[0021] While aspects of the present invention can be described and claimed in a particular statutory class, such as the system statutory class, this is for convenience only and one of skill in the art will understand that each aspect of the present invention can be described and claimed in any statutory class. Unless otherwise expressly stated, it is in no way intended that any method or aspect set forth herein be construed as requiring that its steps be performed in a specific order. Accordingly, where a method claim does not specifically state in the claims or descriptions that the steps are to be limited to a specific order, it is no way intended that an order be inferred, in any respect. This holds for any

possible non-express basis for interpretation, including matters of logic with respect to arrangement of steps or operational flow, plain meaning derived from grammatical organization or punctuation, or the number or type of aspects described in the specification.

BRIEF DESCRIPTION OF THE FIGURES

[0022] The accompanying figures, which are incorporated in and constitute a part of this specification, illustrate several aspects and together with the description serve to explain the principles of the invention.

[0023] **FIGs. 1A-1H** show representative images of *rth6* mutant phenotypes. The images in **FIG. 1A** show microscopic images of 3-4 cm long primary roots, seminal roots and crown roots of wild-type (WT) (upper panels) and *rth6* mutants (lower panels). The data in **FIG. 1B** show root hair length of 3-4 cm long primary roots, seminal roots and crown roots of WT and *rth6* mutants (Scale bar: 1 mm; n=25, error bars indicate SD; *p*-value ***: <0.001 via Students t-test). The images in **FIGs. 1C-1H** are cryo scanning electron microscopic (cSEM) images of root surfaces. WT primary root surface (**FIGs. 1C-1E**) and rth6 primary root surface (**FIGs. 1F-1H**) at 100x magnification (**FIGs. 1C and 1F**), 1000x magnification (**FIGs. 1D and 1G**) and tips of root hairs at 10,000x magnification (**FIGs. 1E and 1H**).

[0024] **FIGs. 2A-2D** show representative data for BSR-Seq mapping of *rth6*. **FIG. 2A** shows BSR-Seq analyses confining rth6 to the centromeric region of chromosome 1. **FIG. 2B** shows BSR-Seq scanning data confining *rth6* to a 15.7 cM interval which corresponds to ~50 Mb. **FIG. 2C** shows correlation of physical and genetic map positions illustrating the low recombination rates in the centromeric region of chromosome 1. The vertical line indicates the position of the *rth6* gene on the physical map. **FIG. 2D** shows the gene structure of *rth6*. Exons: black boxes, introns: thin lines between black boxes, UTR: untranslated region, CDS: coding sequence. The positions of transposon insertions in for additional alleles, *rth6-2*, *rth6-4*, *rth6-5*, and *rth6-6* are indicated. Functional domains of the RTH6 protein are indicated. RING/Ubox: RING/Ubox like Zinc-finger domain, TM: transmembrane domain.

[0025] **FIG. 3** shows a representative phylogenetic reconstruction of the CSLD subfamily of cellulose synthases. Dicot species: blue, monocot species: red, non-seed plants: green. Posterior probabilities are indicated at the branching points.

[0026] **FIGs. 4A-4C** shows representative expression data for *rth6*. **FIG. 4A** shows the expression of *rth6* in primary roots of different stages, seminal roots, crown roots, lateral roots and leaves. **FIG. 4B** shows expression of *rth6* in the meristematic zone (MZ), elongation zone (EZ), cortex (C), stele (S) and root hairs (RH) of 2-4 cm primary roots. **FIG.**

4C shows expression of *rth6* in 2-4 cm primary roots of the mutants *rth1*, *rth2*, *rth3*, *rth5*, and *rth6* and their wild-type (WT) siblings. Bars indicate mean relative expression ± SD (n=4, ten plants per biological replicate). (A) and (B) One way ANOVA; Tukey test ($p \leq 0.01$). (C) Student's t-test (**: $p \leq 0.01$; ***: $p \leq 0.001$).

[0027] **FIGs. 5A-5B** shows the expression of CslD1, CslD2, and CslD5 (*rth6*). The relative expression of CslD1 (white bars), CslD2 (grey bars) and CslD5 (rth6; black bars) is shown for primary roots of different stages, seminal roots, crown roots, lateral roots and leaves **(FIG. 5A)** and the meristematic zone (MZ), elongation zone (EZ), cortex (C), stele (S) and root hairs (RH) of 2-4 cm primary roots **(FIG. 5B)**.

DESCRIPTION

[0028] The present invention can be understood more readily by reference to the following detailed description of the invention and the Examples included therein.

[0029] As used herein the singular forms "a", "and", and "the" include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to "a cell" includes a plurality of such cells and reference to "the protein" includes reference to one or more proteins and equivalents thereof known to those skilled in the art, and so forth. All technical and scientific terms used herein have the same meaning as commonly understood to one of ordinary skill in the art to which this invention belongs unless clearly indicated otherwise.

[0030] Ranges can be expressed herein as from "about" one particular value, and/or to "about" another particular value. When such a range is expressed, a further aspect includes from the one particular value and/or to the other particular value. Similarly, when values are expressed as approximations, by use of the antecedent "about," it will be understood that the particular value forms a further aspect. It will be further understood that the endpoints of each of the ranges are significant both in relation to the other endpoint, and independently of the other endpoint. It is also understood that there are a number of values disclosed herein, and that each value is also herein disclosed as "about" that particular value in addition to the value itself. For example, if the value "10" is disclosed, then "about 10" is also disclosed. It is also understood that each unit between two particular units are also disclosed. For example, if 10 and 15 are disclosed, then 11, 12, 13, and 14 are also disclosed.

[0031] As used herein, the terms "optional" or "optionally" means that the subsequently described event or circumstance can or cannot occur, and that the description includes instances where said event or circumstance occurs and instances where it does not.

[0032] Unless otherwise expressly stated, it is in no way intended that any method set

forth herein be construed as requiring that its steps be performed in a specific order. Accordingly, where a method claim does not actually recite an order to be followed by its steps or it is not otherwise specifically stated in the claims or descriptions that the steps are to be limited to a specific order, it is no way intended that an order be inferred, in any respect. This holds for any possible non-express basis for interpretation, including: matters of logic with respect to arrangement of steps or operational flow; plain meaning derived from grammatical organization or punctuation; and the number or type of aspects described in the specification.

I. OVERVIEW

[0033] In one aspect, the invention relates to methods for increasing the expression of an RTH6 polypeptide in a plant. The methods comprise transforming the plant with a DNA construct comprising an RTH6 polynucleotide operably linked to a heterologous promoter that drives expression in a plant root. Alternatively the plant can be altered to increase the levels of expression of the endogenous RTH6 polypeptide. Such methods for altering the expression of the endogenous gene are known in the art and are encompassed by the present invention.

[0034] Using the methods and compositions of the present invention, drought tolerance and/or the uptake of nutrients in a plant can be increased. In particular, polynucleotides that encode polypeptides from the roothairless6 (RTH6) family are provided. Polypeptides from the RTH6 family promote root hair formation and growth.

[0035] Root hairs are important for the uptake of water and nutrients in plants. As shown herein, the roothairless6 (RTH6) family of polynucleotides and polypeptides are important for the formation of root hairs. Expression of RTH6 polypeptides in plant roots promotes the formation and growth of root hairs, thus improving water and nutrient uptake by the plant.

[0036] As referred to herein, the "RTH6 family" refers to a family of polynucleotides and polypeptides that have high sequence identity with and include the *Zea mays* L. (Maize) *roothairless6* (*rth6*) gene, cDNA, or polypeptide described herein as SEQ ID NOs: 1, 2, and 3, respectively, and promote the formation and growth of root hairs in a plant. An "RTH6 family member" refers to a specific polynucleotide or polypeptide included in the RTH6 family (an *rth6* polynucleotide or *RTH6* polypeptide, respectively). As an example, maize *roothairless6* is an RTH6 family member. The maize wild-type "*roothairless6*" and "*RTH6*" genomic sequence, cDNA, and polypeptide are provided in SEQ ID NOs: 1, 2, and 3, respectively.

[0037] Polynucleotides of the invention are polynucleotides encoding RTH6 family members and include without limitation, the polynucleotides corresponding to SEQ ID NOs: 1 and 2, as well as active fragments and variants thereof. Polypeptides of the invention include members of the RTH6 family. Such RTH6 polypeptides include those set forth in SEQ ID NO: 3, and active fragments and variants thereof. RTH6 family members include polynucleotides having sequences with about 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.3%, 99.5% or more sequence identity to SEQ ID NOs: 1 and 2, and which encode RTH6 polypeptides that have the ability to promote root hair formation and growth in a plant. RTH6 family members also include polypeptides having sequences with about 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.3%, 99.5% or more sequence identity to SEQ ID NO: 3 and which promote root hair formation and growth.

[0038] The RTH6 family are structurally related to Cellulose Synthase-like D (CSLD) protein which belongs to the D-type subfamily of the Cellulose Synthase superfamily. Without wishing to be limited to a particular mechanism of action, cellulose synthase and CLSD proteins are thought to play an important role in root hair tip growth. The methods of the invention utilize RTH6 family polynucleotides to increase the expression of RTH6 family polypeptides, which advantageously increase root hair formation and growth.

[0039] In general, concentration and/or activity of the RTH6 family member is increased by at least 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, or 90% relative to a native control plant, plant part, or cell which did not have the sequence of the invention introduced. The increased expression in the present invention can occur during and/or subsequent to growth of the plant to the desired stage of development. In specific aspects, the levels and/or activity of RTH6 polypeptides of the present invention are increased in monocots, particularly maize.

[0040] The expression level of the RTH6 family polypeptide may be measured directly, for example, by assaying for the level of the RTH6 family polypeptide in the plant.

[0041] Root hair cells include long tubular projections referred to herein as "root hairs." Root hairs are thought to aid plants in nutrient uptake, anchorage, and microbial interactions. Root hair growth is divided into three phases: first, defined swelling to form a bulge; second, transition to tip growth; and finally, tip growth by oriented exocytosis. Root hairs increase the surface area on a plant root, thereby increasing the ability of the root to take up water and nutrients.

[0042] Root hair formation and growth can be measured by phenotypic analysis. For example, root hair formation and growth can be measured using microscopy techniques described in the experimental section and as described by Foreman *et al.* (2003) *Nature* 422:442-446, herein incorporated by reference.

[0043] The formation and growth of root hairs leads to increased drought tolerance. As referred to herein, "drought" refers to a decrease in water availability to a plant that, especially when prolonged, can cause damage to the plant or prevent its successful growth (e.g., limiting plant growth or seed yield). Accordingly, "drought tolerance" is a trait of a plant to survive under drought conditions over prolonged periods of time without exhibiting substantial physiological or physical deterioration.

[0044] Embodiments of the present methods and compositions promote "increased drought tolerance" of a plant. Increased drought tolerance is measured relative to a reference or control plant, and is a trait of the plant to survive under drought conditions over prolonged periods of time, without exhibiting the same degree of physiological or physical deterioration relative to the reference or control plant grown under similar drought conditions. Typically, when a transgenic plant comprising an expression construct in its genome exhibits increased drought tolerance relative to a reference or control plant due to the presence of the construct, the reference or control plant does not comprise in its genome the expression construct.

[0045] One of ordinary skill in the art is familiar with protocols for simulating drought conditions and for evaluating drought tolerance of plants that have been subjected to simulated or naturally occurring drought conditions. For example, one can simulate drought conditions by giving plants less water than normally required or no water over a period of time, and one can evaluate drought tolerance by looking for differences in physiological and/or physical condition, including (but not limited to) vigor, growth, size, or root length, or in particular, leaf color or leaf area size. Other techniques for evaluating drought tolerance include measuring chlorophyll fluorescence, photosynthetic rates and gas exchange rates. See, for example, WO 2013/006345 herein incorporated by reference in its entirety.

[0046] One can also evaluate drought tolerance by the ability of a plant to maintain sufficient yield (at least about 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% yield) in field testing under simulated or naturally occurring drought conditions (e.g., by measuring for substantially equivalent yield under drought conditions compared to non-drought conditions, or by measuring for less yield loss under drought conditions compared to a control or reference plant).

[0047] In addition, increased formation and growth of root hairs leads to increased nutrient uptake. As referred to herein, "nutrient uptake" refers to a plant's ability it to remove nutrients from a soil or growth medium. "Increased nutrient uptake" refers to a plant's ability to remove nutrients from a soil or growth medium relative to a reference or control plant, and is a trait of the plant wherein the plant demonstrates an improved agronomic characteristic compared to the reference or control plant. Typically, when a transgenic plant comprising an expression construct in its genome exhibits increased nutrient uptake relative to a reference or control plant due to the presence of the construct, the reference or control plant does not comprise in its genome the expression construct. Nonlimiting examples of nutrients taken up by plant roots include nitrogen, phosphorus, potassium, and carbon.

[0048] "Agronomic characteristic" or "agronomic parameter" is a measurable trait including but not limited to, abiotic stress tolerance, greenness, stay-green, yield, growth rate, biomass, fresh weight at maturation, dry weight at maturation, fruit yield, seed yield, total plant nitrogen content, fruit nitrogen content, seed nitrogen content, nitrogen content in a vegetative tissue, total plant free amino acid content, fruit free amino acid content, seed free amino acid content, free amino acid content in a vegetative tissue, total plant protein content, fruit protein content, seed protein content, protein content in a vegetative tissue, drought tolerance, nitrogen stress tolerance, nitrogen uptake, root lodging, root mass, harvest index, stalk lodging, plant height, ear height, ear length, salt tolerance, early seedling vigor and seedling emergence under low temperature stress.

[0049] Yield can be measured in many ways, including, for example, test weight, seed weight, seed number per plant, seed number per unit area (i.e. seeds, or weight of seeds, per acre), bushels per acre, tonnes per hectare, tonnes per acre, tons per acre and kilograms per hectare.

[0050] It is understood and herein contemplated that the plants and germplasms disclosed herein can be identified as having an alternation in at least one agronomic characteristic through the identification of a polymorphism in a polynucleotide having at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.3%, or 99.5% identity to SEQ ID NO: 1 or 2 or the identification of a polymorphism in a polynucloeotide that encodes a polypeptide having at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.3%, or 99.5% identity to SEQ ID NO: 3. Thus, in one aspect, disclosed herein are methods of selecting or identifying a first maize plant or a first maize germplasm that has one or more beneficial alleles of RTH6, the method comprising:

(a) screening a plurality of maize plants or a plurality of maize germplasm for at least one polymorphism within a marker locus, wherein the marker locus is:

(i) a first polynucleotide having at least 90% and less than 100% nucleotide sequence identity with SEQ ID NO:1 or 2; or

(ii) a second polynucleotide encoding a polypeptide comprising an amino acid sequence having at least 90% and less than 100% sequence identity to SEQ ID NO: 3, wherein expression of the first or second polynucleotide in a maize plant results in a phenotype comprising an alteration of at least one agronomic characteristic when compared to a control maize plant; wherein the at least one agronomic characteristic is selected from the group consisting of increased root hair formation and growth, increased drought tolerance, and enhanced nutrient uptake; and wherein the control maize plant comprises:

a polynucleotide having the nucleotide sequence of SEQ ID NO:1 or 2; or

a polynucleotide encoding the amino acid sequence of SEQ ID NO:3;

(b) identifying a first maize plant or a first maize germplasm comprising the at least one polymorphism of the marker locus; and

(c) selecting the first maize plant or first maize germplasm of step (b).

[0051] Abiotic stress can be at least one condition selected from the group consisting of: drought, water deprivation, flood, high light intensity, high temperature, low temperature, salinity, etiolation, defoliation, heavy metal toxicity, anaerobiosis, nutrient deficiency, nutrient excess, UV irradiation, atmospheric pollution (e.g., ozone) and exposure to chemicals (e.g., paraquat) that induce production of reactive oxygen species (ROS).

[0052] "Increased stress tolerance" of a plant is measured relative to a reference or control plant, and is a trait of the plant to survive under stress conditions over prolonged periods of time, without exhibiting the same degree of physiological or physical deterioration relative to the reference or control plant grown under similar stress conditions.

[0053] A plant with "increased stress tolerance" can exhibit increased tolerance to one or more different stress conditions.

[0054] "Stress tolerance activity" of a polypeptide indicates that over-expression of the polypeptide in a transgenic plant confers increased stress tolerance to the transgenic plant relative to a reference or control plant.

[0055] Methods of assaying increased nutrient uptake are known in the art. For example, one can grow a transgenic plant comprising an expression construct and evaluate nutrient uptake by looking for differences in physiological and/or physical condition, including (but

not limited to) vigor, growth, size, or root length, leaf color, leaf area size, or crop yield. Other methods include measurement of the kinetics of plant root nutrient uptake. See, for example, the methods reviewed by Bassirirad (2000) *New Phytol.* 147:155-169, incorporated by reference herein.

[0056] In another aspect, increased root hair formation and growth can reduce soil erosion in crop fields. It is recognized that root hairs help cling to soil particles and are therefore important to preventing soil erosion.

[0057] In addition, increased root hair formation and growth is expected to promote the infection of the plant with beneficial microorganisms, such as rhizobium and *Pseudomonas putida* KT2440. Rhizobia enter the plant via root hairs, which results in root nodule formation and increased nitrogen fixation.

[0058] In addition, plants are known to interact with a wide range of rhizosphere-colonizing bacteria. These are attracted to root surfaces by chemical components in root exudates, which are rapidly assimilated into microbial biomass (Rangel-Castro JI, *et al.* (2005) *Environ Microbiol* 7: 828–838). This so-called rhizosphere effect supports bacterial cell densities in the root vicinity up to 100-fold greater than in surrounding soil (Whipps JM *et al.* (2001) *J. Exp. Bot.* 52: 487–511).

[0059] In another aspect, it is recognized that the reduction of RTH6 family expression or mutation in RTH6 family members can result in reduced susceptibility of the plant to plant pathogens which infect the plant roots. For example, *Barssica* are susceptible to *Plasmodiophora brassicae*, which is the pathogen responsible for clubroot. Clubroot is caused when *P. brassicae* enter the plant via the root hairs. Therefore, it is thought that reducing the number of root hairs can limit the mode of entry of plant pathogens.

[0060] Thus, in one aspect, disclosed herein are methods of increasing the resistance to root pathogens in a plant, said method comprising reducing the expression of an RTH6 family member, or inhibiting the function of the RTH6 family member. Maize root pathogens can be any organisms that cause a dileterious effecto on a plant and can infect the plant through the root or root hairs. On non-limiting list of maize root pathogens include, but are not limited to, *Fusarium* species, like *F. verticillioides* and *F. graminearu.*

II. DNA CONSTRUCTS

[0061] The RTH6 family polynucleotides disclosed herein can be provided in expression cassettes for expression in a plant of interest. An "expression cassette" comprises a polynucleotide that is operably linked to a promoter, wherein the promoter drives expression of the polynucleotide. It is understood that the polynucleotide operably linked to the

promoter is a "recombinant" nucleic acid is made by a combination of two otherwise separated segments of nucleic acid sequence, for example, by chemical synthesis or by the manipulation of isolated segments of polynucleic acids by genetic engineering techniques. The term "expression construct" refers to any agent such as a plasmid, cosmid, virus, autonomously replicating sequence, phage, or linear or circular single-stranded or double-stranded DNA or RNA nucleotide sequence, derived from any source, capable of genomic integration or autonomous replication, comprising a DNA molecule where one or more DNA sequences have been linked in a functionally operative manner. Such expression constructs are capable of introducing a 5′ regulatory sequence or promoter region and a DNA sequence for a selected gene product into a cell in such a manner that the DNA sequence is transcribed into a functional mRNA that is translated and therefore expressed. An expression construct (and an expression cassette) may comprise a polynucleotide of interest linked to a heterologous polynucleotide such as, for example, a heterologous promoter.

[0062] The use of the term "polynucleotide" is not intended to be limiting to polynucleotides comprising DNA. Those of ordinary skill in the art will recognize that polynucleotides can comprise ribonucleotides and combinations of ribonucleotides and deoxyribonucleotides. Such deoxyribonucleotides and ribonucleotides include both naturally occurring molecules and synthetic analogues. The disclosed polynucleotides also encompass all forms of sequences including, but not limited to, single-stranded forms, double-stranded forms, hairpins, stem-and-loop structures, and the like.

[0063] The cassette can include 5' and 3' regulatory sequences operably linked to an RTH6 polynucleotide. "Operably linked" is intended to mean a functional linkage between two or more elements. For example, an operable linkage between a polynucleotide of interest and a regulatory sequence (i.e., a promoter) is functional link that allows for expression of the polynucleotide of interest. Operably linked elements may be contiguous or non-contiguous. When used to refer to the joining of two protein coding regions, by operably linked is intended that the coding regions are in the same reading frame. The cassette may additionally contain at least one additional gene to be cotransformed into the organism. Alternatively, the additional gene(s) can be provided on multiple expression cassettes. Such an expression cassette is provided with a plurality of restriction sites and/or recombination sites for insertion of the RTH6 polynucleotide to be under the transcriptional regulation of the regulatory regions. The expression cassette may additionally contain selectable marker genes.

[0064] In a further aspect, the expression cassette includes a polynucleotide operably linked to a root-preferred heterologous promoter. The polynucleotide can include a nucleotide sequence encompassed by SEQ ID NO:1 or 2. In various aspects, the nucleotide sequence has at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.3%, or 99.5% sequence identity to SEQ ID NO: 1 or 2, and the polynucleotide encodes a polypeptide having NADPH oxidase activity. In various aspects, the nucleotide sequence is the sequence of SEQ ID NO: 1. In a further aspect, the nucleotide sequence is the sequence of SEQ ID NO: 2.

[0065] In various aspects, the expression cassette includes a nucleotide sequence that encodes the amino acid sequence set forth in SEQ ID NO: 1 or 2. In some aspects, the nucleotide sequence encodes an amino acid sequence having at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.3%, or 99.5% sequence identity to SEQ ID NO: 1 or 2, wherein said polynucleotide encodes a polypeptide that has cellulose synthase. In some aspects the polynucleotide encodes a polypeptide having the amino acid sequence of SEQ ID NO: 3.

[0066] The expression cassette can include in the 5'-3' direction of transcription, a transcriptional and translational initiation region (i.e., a promoter), an RTH6 family polynucleotide as disclosed herein, and a transcriptional and translational termination region (i.e., termination region) functional in plants. The regulatory regions (i.e., promoters, transcriptional regulatory regions, and translational termination regions) and/or the RTH6 family polynucleotide of the invention may be native/analogous to the host cell or to each other. Alternatively, the regulatory regions and/or the RTH6 family polynucleotide of the invention can be heterologous to the host cell or to each other. As used herein, "heterologous" in reference to a sequence is a sequence that originates from a foreign species, or, if from the same species, is substantially modified from its native form in composition and/or genomic locus by deliberate human intervention. For example, a promoter operably linked to a heterologous polynucleotide is from a species different from the species from which the polynucleotide was derived, or, if from the same/analogous species, one or both are substantially modified from their original form and/or genomic locus, or the promoter is not the native promoter for the operably linked polynucleotide. As used herein, a "chimeric gene" comprises a coding sequence operably linked to a transcription initiation region that is heterologous to the coding sequence.

[0067] In other aspects, a double stranded RNA can be expressed from a suppression cassette. Such a cassette can comprise two convergent promoters that drive transcription of

an operably linked silencing element. "Convergent promoters" refers to promoters that are oriented on either terminus of the operably linked silencing element such that each promoter drives transcription of the silencing element in opposite directions, yielding two transcripts. In such aspects, the convergent promoters allow for the transcription of the sense and anti-sense strand and thus allow for the formation of a dsRNA. Such a cassette may also comprise two divergent promoters that drive transcription of one or more operably linked silencing elements. "Divergent promoters" refers to promoters that are oriented in opposite directions of each other, driving transcription of the one or more silencing elements in opposite directions. In such embodiments, the divergent promoters allow for the transcription of the sense and antisense strands and allow for the formation of a dsRNA. In such embodiments, the divergent promoters also allow for the transcription of at least two separate hairpin RNAs. In another embodiment, one cassette comprising two or more silencing elements under the control of two separate promoters in the same orientation is present in a construct. In another embodiment, two or more individual cassettes, each comprising at least one silencing element under the control of a promoter, are present in a construct in the same orientation.

[0068] While it may be optimal to express the sequences using heterologous promoters, the native promoter sequences can be used. Such constructs can change expression levels of RTH6 family proteins in the plant or plant cell. Thus, the phenotype of the plant or plant cell can be altered.

[0069] The termination region can be native with the transcriptional initiation region, can be native with the operably linked RTH6 family polynucleotide of interest, can be native with the plant host, or can be derived from another source (i.e., foreign or heterologous) to the promoter, the RTH6 polynucleotide of interest, the plant host, or any combination thereof. Convenient termination regions are available from the Ti-plasmid of *A. tumefaciens*, such as the octopine synthase and nopaline synthase termination regions. See also Guerineau *et al.* (1991) *Mol. Gen. Genet.* 262:141-144; Proudfoot (1991) *Cell* 64:671-674; Sanfacon *et al.* (1991) *Genes Dev.* 5:141-149; Mogen *et al.* (1990) *Plant Cell* 2:1261-1272; Munroe *et al.* (1990) *Gene* 91:151-158; Ballas *et al.* (1989) *Nucleic Acids Res.* 17:7891-7903; and Joshi *et al.* (1987) *Nucleic Acids Res.* 15:9627-9639.

[0070] Where appropriate, the polynucleotides may be optimized for increased expression in the transformed plant. That is, the polynucleotides can be synthesized using plant-preferred codons for improved expression. See, for example, Campbell and Gowri (1990) *Plant Physiol.* 92:1-11 for a discussion of host-preferred codon usage. Methods are available in the art for synthesizing plant-preferred genes. See, for example, U.S. Patent Nos.

5,380,831, and 5,436,391, and Murray *et al.* (1989) *Nucleic Acids Res.* 17:477-498, herein incorporated by reference.

[0071] Additional sequence modifications are known to enhance gene expression in a cellular host. These include elimination of sequences encoding spurious polyadenylation signals, exon-intron splice site signals, transposon-like repeats, and other such well-characterized sequences that may be deleterious to gene expression. The G-C content of the sequence can be adjusted to levels average for a given cellular host, as calculated by reference to known genes expressed in the host cell. When possible, the sequence is modified to avoid predicted hairpin secondary mRNA structures.

[0072] The expression cassettes may additionally contain 5' leader sequences. Such leader sequences can act to enhance translation. Translation leaders are known in the art and include: picornavirus leaders, for example, EMCV leader (Encephalomyocarditis 5' noncoding region) (Elroy-Stein *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:6126-6130); potyvirus leaders, for example, TEV leader (Tobacco Etch Virus) (Gallie *et al.* (1995) *Gene* 165(2):233-238), MDMV leader (Maize Dwarf Mosaic Virus) (*Virology* 154:9-20), and human immunoglobulin heavy-chain binding protein (BiP) (Macejak *et al.* (1991) *Nature* 353:90-94); untranslated leader from the coat protein mRNA of alfalfa mosaic virus (AMV RNA 4) (Jobling *et al.* (1987) *Nature* 325:622-625); tobacco mosaic virus leader (TMV) (Gallie *et al.* (1989) in *Molecular Biology of RNA*, ed. Cech (Liss, New York), pp. 237-256); and maize chlorotic mottle virus leader (MCMV) (Lommel *et al.* (1991) *Virology* 81:382-385). See also, Della-Cioppa *et al.* (1987) *Plant Physiol.* 84:965-968.

[0073] In preparing the expression cassette, the various DNA fragments can be manipulated, so as to provide for the DNA sequences in the proper orientation and, as appropriate, in the proper reading frame. Toward this end, adapters or linkers can be employed to join the DNA fragments or other manipulations can be involved to provide for convenient restriction sites, removal of superfluous DNA, removal of restriction sites, or the like. For this purpose, *in vitro* mutagenesis, primer repair, restriction, annealing, resubstitutions, e.g., transitions and transversions, can be involved.

[0074] A number of promoters can be used in the practice of the invention. The promoters can be selected based on the desired outcome. The nucleic acids can be combined with constitutive, tissue-preferred, inducible, or other promoters for expression in the host organism. Promoters useful in the methods of the invention include those promoters that drive expression of a polypeptide in roots of a plant.

[0075] Constitutive promoters can be used such as the core promoter of the Rsyn7 promoter and other constitutive promoters disclosed in WO 99/43838 and U.S. Patent No. 6,072,050; the core CaMV 35S promoter (Odell *et al.* (1985) *Nature* 313:810-812); rice actin (McElroy *et al.* (1990) *Plant Cell* 2:163-171); ubiquitin (Christensen *et al.* (1989) *Plant Mol. Biol.* 12:619-632 and Christensen *et al.* (1992) *Plant Mol. Biol.* 18:675-689); pEMU (Last *et al.* (1991) *Theor. Appl. Genet.* 81:581-588); MAS (Velten *et al.* (1984) *EMBO J.* 3:2723-2730); ALS promoter (U.S. Patent No. 5,659,026), and the like. Other constitutive promoters include, for example, U.S. Patent Nos. 5,608,149; 5,608,144; 5,604,121; 5,569,597; 5,466,785; 5,399,680; 5,268,463; 5,608,142; and 6,177,611.

[0076] Depending on the desired outcome, it may be beneficial to express the gene from an inducible promoter. An inducible promoter, for instance, a pathogen-inducible promoter could also be employed. Such promoters include those from pathogenesis-related proteins (PR proteins), which are induced following infection by a pathogen; e.g., PR proteins, SAR proteins, beta-1,3-glucanase, chitinase, etc. See, for example, Redolfi *et al.* (1983) *Neth. J. Plant Pathol.* 89:245-254; Uknes *et al.* (1992) *Plant Cell* 4:645-656; and Van Loon (1985) *Plant Mol. Virol.* 4:111-116. See also WO 99/43819, herein incorporated by reference.

[0077] Chemical-regulated promoters can be used to modulate the expression of a gene in a plant through the application of an exogenous chemical regulator. Depending upon the objective, the promoter may be a chemical-inducible promoter, where application of the chemical induces gene expression, or a chemical-repressible promoter, where application of the chemical represses gene expression. Chemical-inducible promoters are known in the art and include, but are not limited to, the maize In2-2 promoter, which is activated by benzenesulfonamide herbicide safeners, the maize GST promoter, which is activated by hydrophobic electrophilic compounds that are used as pre-emergent herbicides, and the tobacco PR-1a promoter, which is activated by salicylic acid. Other chemical-regulated promoters of interest include steroid-responsive promoters (see, for example, the glucocorticoid-inducible promoter in Schena *et al.* (1991) *Proc. Natl. Acad. Sci. USA* 88:10421-10425 and McNellis *et al.* (1998) *Plant J.* 14(2):247-257) and tetracycline-inducible and tetracycline-repressible promoters (see, for example, Gatz *et al.* (1991) *Mol. Gen. Genet.* 227:229-237, and U.S. Patent Nos. 5,814,618 and 5,789,156), herein incorporated by reference.

[0078] Tissue-preferred promoters can be utilized to target enhanced expression within a particular plant tissue. A "root-preferred" promoter, as referred to herein, is a promoter which favors spatial expression of a polynucleotide of interest in the root of a plant compared

to expression in other plant tissue within the same plant. Root-preferred promoters can be heterologous or native to the plant receiving the polynucleotide of interest. Root-preferred promoters are known and can be selected from the many available from the literature or isolated de novo from various compatible species. See, for example, Hire *et al.* (1992) *Plant Mol. Biol.* 20(2):207-218 (soybean root-specific glutamine synthetase gene); Keller and Baumgartner (1991) *Plant Cell* 3(10):1051-1061 (root-specific control element in the GRP 1.8 gene of French bean); Sanger *et al.* (1990) *Plant Mol. Biol.* 14(3):433-443 (root-specific promoter of the mannopine synthase (MAS) gene of *Agrobacterium tumefaciens*); and Miao *et al.* (1991) *Plant Cell* 3(1):11-22 (full-length cDNA clone encoding cytosolic glutamine synthetase (GS), which is expressed in roots and root nodules of soybean). See also Bogusz *et al.* (1990) *Plant Cell* 2(7):633-641, where two root-specific promoters isolated from hemoglobin genes from the nitrogen-fixing nonlegume *Parasponia andersonii* and the related non-nitrogen-fixing nonlegume *Trema tomentosa* are described. Leach and Aoyagi (1991) describe their analysis of the promoters of the highly expressed rolC and rolD root-inducing genes of *Agrobacterium rhizogenes* (see *Plant Science* (Limerick) 79(1):69-76). They concluded that enhancer and tissue-preferred DNA determinants are dissociated in those promoters. Teeri *et al.* (1989) used gene fusion to lacZ to show that the *Agrobacterium* T-DNA gene encoding octopine synthase is especially active in the epidermis of the root tip and that the TR2' gene is root specific in the intact plant and stimulated by wounding in leaf tissue, an especially desirable combination of characteristics for use with an insecticidal or larvicidal gene (see *EMBO J.* 8(2):343-350). The TR1' gene, fused to *nptII* (neomycin phosphotransferase II) showed similar characteristics. Additional root-preferred promoters include the VfENOD-GRP3 gene promoter (Kuster *et al.* (1995) *Plant Mol. Biol.* 29(4):759-772); and rolB promoter (Capana *et al.* (1994) *Plant Mol. Biol.* 25(4):681-691. See also U.S. Patent Nos. 5,837,876; 5,837,848; 5,750,386; 5,633,363; 5,459,252; 5,401,836; 5,110,732; 5,023,179; and 7,554,005.

[0079] Other root preferred promoters include the following: the maize NAS2 promoter, the maize Cyclo promoter (US 2006/0156439, published July 13, 2006), the maize ROOTMET2 promoter (WO05063998, published July 14, 2005), the CR1BIO promoter (WO06055487, published May 26, 2006), the CRWAQ81 promoter (WO05035770, published April 21, 2005) and the maize ZRP2.47 promoter (NCBI accession number: U38790; GI No. 1063664).

[0080] The above list of promoters is not meant to be limiting. Any root-preferred promoter or promoter that drives expression in the root of a plant can be used with the

polynucleotides disclosed herein.

[0081] The expression cassette can also comprise a selectable marker gene for the selection of transformed cells. Selectable marker genes are utilized for the selection of transformed cells or tissues. Many selectable markers are known in the art and any can be used in the practice of the invention. Marker genes include genes encoding antibiotic resistance, such as those encoding neomycin phosphotransferase II (NEO) and hygromycin phosphotransferase (HPT), as well as genes conferring resistance to herbicidal compounds, such as glufosinate ammonium, bromoxynil, imidazolinones, and 2,4-dichlorophenoxyacetate (2,4-D). Additional selectable markers include phenotypic markers such as β-galactosidase and fluorescent proteins such as green fluorescent protein (GFP) (Su et *al.* (2004) *Biotechnol Bioeng 85:*610-9 and Fetter *et al.* (2004) *Plant Cell 16:*215-28), cyan florescent protein (CYP) (Bolte *et al.* (2004) *J. Cell Science 117*:943-54 and Kato *et al.* (2002) *Plant Physiol 129*:913-42), and yellow florescent protein (PhiYFP™ from Evrogen, see, Bolte *et al.* (2004) *J. Cell Science 117*:943-54). For additional selectable markers, see generally, Yarranton (1992) *Curr. Opin. Biotech.* 3:506-511; Christopherson *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:6314-6318; Yao *et al.* (1992) *Cell* 71:63-72; Reznikoff (1992) *Mol. Microbiol.* 6:2419-2422; Barkley *et al.* (1980) in *The Operon*, pp. 177-220; Hu *et al.* (1987) *Cell* 48:555-566; Brown *et al.* (1987) *Cell* 49:603-612; Figge *et al.* (1988) *Cell* 52:713-722; Deuschle *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:5400-5404; Fuerst *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:2549-2553; Deuschle *et al.* (1990) *Science* 248:480-483; Gossen (1993) Ph.D. Thesis, University of Heidelberg; Reines *et al.* (1993) *Proc. Natl. Acad. Sci. USA* 90:1917-1921; Labow *et al.* (1990) *Mol. Cell. Biol.* 10:3343-3356; Zambretti *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:3952-3956; Baim *et al.* (1991) *Proc. Natl. Acad. Sci. USA* 88:5072-5076; Wyborski *et al.* (1991) *Nucleic Acids Res.* 19:4647-4653; Hillenand-Wissman (1989) *Topics Mol. Struc. Biol.* 10:143-162; Degenkolb *et al.* (1991) *Antimicrob. Agents Chemother.* 35:1591-1595; Kleinschnidt *et al.* (1988) *Biochemistry* 27:1094-1104; Bonin (1993) Ph.D. Thesis, University of Heidelberg; Gossen *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:5547-5551; Oliva *et al.* (1992) *Antimicrob. Agents Chemother.* 36:913-919; Hlavka *et al.* (1985) *Handbook of Experimental Pharmacology*, Vol. 78 (Springer-Verlag, Berlin); Gill *et al.* (1988) *Nature* 334:721-724. Such disclosures are herein incorporated by reference. The above list of selectable marker genes is not meant to be limiting. Any selectable marker gene can be used with the compositions and methods described herein.

[0082] Embodiments found herein encompass isolated, substantially purified, or recombinant polynucleotide or protein compositions. An "isolated" or "purified" polynucleotide or protein, or biologically active portion thereof, is substantially or essentially free from components that normally accompany or interact with the polynucleotide or protein as found in its naturally occurring environment. Thus, an isolated or purified polynucleotide or protein is substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized. Optimally, an "isolated" polynucleotide is free of sequences (optimally protein encoding sequences) that naturally flank the polynucleotide (i.e., sequences located at the 5' and 3' ends of the polynucleotide) in the genomic DNA of the organism from which the polynucleotide is derived. A protein that is substantially free of cellular material includes preparations of protein having less than about 30%, 20%, 10%, 5%, or 1% (by dry weight) of contaminating protein. When the protein of the invention or biologically active portion thereof is recombinantly produced, optimally culture medium represents less than about 30%, 20%, 10%, 5%, or 1% (by dry weight) of chemical precursors or non-protein-of-interest chemicals.

III. FRAGMENTS AND VARIANTS

[0083] Fragments and variants of the disclosed polynucleotides and proteins encoded thereby are also encompassed by the present invention. By "fragment" is intended a portion of the polynucleotide or a portion of the amino acid sequence and hence protein encoded thereby. Fragments of a polynucleotide can encode protein fragments that retain the biological activity of the native protein and hence promote root hair formation and growth. Alternatively, fragments of a polynucleotide that are useful as a silencing elements or as hybridization probes do not need to encode fragment proteins that retain biological activity. Thus, fragments of a nucleotide sequence may range from at least about 20 nucleotides, about 50 nucleotides, about 100 nucleotides, and up to the full-length polynucleotide encoding the proteins disclosed herein. A fragment of a polynucleotide or an amino acid sequence may comprise contiguous nucleic acid or amino acid residues, respectively, of the polynucleotide or the amino acid sequence.

[0084] A fragment of an RTH6 family polynucleotide that encodes a biologically active portion of an RTH6 family protein disclosed herein can encode at least 15, 25, 30, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, or 1159 contiguous amino acids, or up to the total number of amino acids present in a full-length RTH6 family protein, for example, 1,159 amino acids for SEQ ID NO: 3.

[0085] Fragments of an RTH6 family polynucleotide that are useful as hybridization probes or PCR primers generally need not encode a biologically active portion of an RTH6 family protein. Thus, a fragment of an RTH6 family polynucleotide can encode a biologically active portion of an RTH6 family protein, or it can be a fragment that can be used as a hybridization probe or PCR primer using methods disclosed below. A biologically active portion of an RTH6 family protein can be prepared by isolating a portion of one of the RTH6 family polynucleotide disclosed herein, expressing the encoded portion of the RTH6 family protein (e.g., by recombinant expression *in vitro*), and assessing the activity of the encoded portion of the RTH6 family protein. Polynucleotides that are fragments of an RTH6 family nucleotide sequence comprise at least 16, 20, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 800, 900, 1,000, 1,100, 1,200, 1,300, 1,400, 1,500, 1,600, 1,700, 1,800, 1,900, 2,000, 2,100, 2,200, 2,300, 2,400, 2,500, 2,600, 2,700, 2,800, 2,900, 3,000, 3,100, 3,200, 3,300, 3,400, 3,500, 3,600, 3,700, 3,800, 3,900, 4,000, 4,100, 4,200, 4,300, 4,400, 4,500, 4,600, 4,700, 4,800, 4,900, 5,000, 5,100, 5,200, 5,300, 5,500, 5,600, 5,700, 5,800, 5,900, 6,000, or 6,100 contiguous nucleotides, or up to the number of nucleotides present in a full-length RTH6 family polynucleotide disclosed herein, for example, 6125 and 3866 nucleotides for SEQ ID NO: 1 and 2, respectively.

[0086] "Variants" is intended to mean substantially similar sequences. For polynucleotides, a variant comprises a deletion and/or addition of one or more nucleotides at one or more internal sites within the native polynucleotide and/or a substitution of one or more nucleotides at one or more sites in the native polynucleotide. A variant of a polynucleotide that is useful as a silencing element will retain the ability to reduce expression of the target polynucleotide and, in some embodiments, thereby control a plant insect pest of interest. As used herein, a "native" polynucleotide or polypeptide comprises a naturally occurring nucleotide sequence or amino acid sequence, respectively. For polynucleotides, conservative variants include those sequences that, because of the degeneracy of the genetic code, encode the amino acid sequence of one of the RTH6 family polypeptides as disclosed herein. Naturally occurring allelic variants such as these can be identified with the use of well-known molecular biology techniques, as, for example, with polymerase chain reaction (PCR) and hybridization techniques as outlined below. Variant polynucleotides also include synthetically derived polynucleotide, such as those generated, for example, by using site-directed mutagenesis but which still encode a RTH6 family protein. Generally, variants of a particular polynucleotide of the invention will have at least about 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%,

99.3%, 99.5% or more sequence identity to that particular polynucleotide as determined by sequence alignment programs and parameters described elsewhere herein.

[0087] Variants of a particular polynucleotide of the invention (i.e., the reference polynucleotide) can also be evaluated by comparison of the percent sequence identity between the polypeptide encoded by a variant polynucleotide and the polypeptide encoded by the reference polynucleotide. Thus, for example, isolated polynucleotides that encodes a polypeptide with a given percent sequence identity to the polypeptide of SEQ ID NO: 3 are disclosed. Percent sequence identity between any two polypeptides can be calculated using sequence alignment programs and parameters described elsewhere herein. Where any given pair of polynucleotides of the invention is evaluated by comparison of the percent sequence identity shared by the two polypeptides they encode, the percent sequence identity between the two encoded polypeptides is at least about 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.3%, 99.5% or more sequence identity.

[0088] "Variant" protein is intended to mean a protein derived from the native protein by deletion or addition of one or more amino acids at one or more internal sites in the native protein and/or substitution of one or more amino acids at one or more sites in the native protein. Variant proteins encompassed by the present invention are biologically active, that is they continue to possess the desired biological activity of the native protein, that is, root hair formation and growth and/or cellulose synthase activity, e.g., ß-glycan synthase activity, as described herein. Such variants may result from, for example, genetic polymorphism or from human manipulation. Biologically active variants of a native RTH6 family protein will have at least about 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.3%, 99.5% or more sequence identity to the amino acid sequence for the native protein as determined by sequence alignment programs and parameters described elsewhere herein. A biologically active variant of a protein of the invention can differ from that protein by as few as 1-15 amino acid residues, as few as 1-10, as few as 6-10, as few as 5, as few as 4, 3, 2, or even 1 amino acid residue.

[0089] In one aspect, disclosed herein are methods of selecting or identifying an allelic variant of RTH6 in a maize plant, the method comprising the steps of a) obtaining a population of maize plants, wherein said maize plants exhibit an alteration of at least one agronomic characteristic; wherein the at least one agronomic characteristic is selected from the group consisting of increased root hair formation and growth, increased drought tolerance, and enhanced nutrient uptake; b) evaluating allelic variations with respect to the

polynucleotide sequence (for example by performing the sequence identity comparisons described above) encoding a protein comprising SEQ ID NO:3, or in the genomic region that regulates the expression of the polynucleotide encoding the protein; c) associating allelic variations with said alteration of at least one agronomic characteristic; and d) selecting or identifying an allelic variant that is associated with said alteration of at least one agronomic characteristic.

[0090] The proteins disclosed herein may be altered in various ways including amino acid substitutions, deletions, truncations, and insertions. Methods for such manipulations are generally known in the art. For example, amino acid sequence variants and fragments of the RTH6 family proteins can be prepared by mutations in the DNA. Methods for mutagenesis and polynucleotide alterations are well known in the art. See, for example, Kunkel (1985) *Proc. Natl. Acad. Sci. USA* 82:488-492; Kunkel *et al.* (1987) *Methods in Enzymol.* 154:367-382; U.S. Patent No. 4,873,192; Walker and Gaastra, eds. (1983) *Techniques in Molecular Biology* (MacMillan Publishing Company, New York) and the references cited therein. Guidance as to appropriate amino acid substitutions that do not affect biological activity of the protein of interest may be found in the model of Dayhoff *et al.* (1978) *Atlas of Protein Sequence and Structure* (Natl. Biomed. Res. Found., Washington, D.C.), herein incorporated by reference. Conservative substitutions, such as exchanging one amino acid with another having similar properties, may be optimal.

[0091] Thus, the genes and polynucleotides of the invention include both the naturally occurring sequences as well as mutant forms. Likewise, the proteins of the invention encompass both naturally occurring proteins as well as variations and modified forms thereof. Such variants continue to possess the desired root hair formation and growth activity and/or cellulose synthase activity, e.g., ß-glycan synthase activity, . The mutations that are made in the DNA encoding the variant must not place the sequence out of reading frame and optimally do not create complementary regions that produce secondary mRNA structure. See, EP Patent Application Publication No. 75,444.

[0092] The deletions, insertions, and substitutions of the protein sequences encompassed herein are not expected to produce radical changes in the characteristics of the protein. However, when it is difficult to predict the exact effect of the substitution, deletion, or insertion in advance of doing so, one skilled in the art will appreciate that the effect can be evaluated by routine screening assays. That is, the activity can be evaluated by, for example, phenotypic analysis of root hair formation and growth and/or by assaying for cellulose synthase activity, e.g., ß-glycan synthase activity, .

[0093] Variant polynucleotides and proteins also encompass sequences and proteins derived from a mutagenic and recombinogenic procedure such as DNA shuffling. With such a procedure, one or more different RTH6 family members coding sequences can be manipulated to create a new RTH6 family member possessing the desired properties. In this manner, libraries of recombinant polynucleotides are generated from a population of related sequence polynucleotides comprising sequence regions that have substantial sequence identity and can be homologously recombined *in vitro* or *in vivo*. For example, using this approach, sequence motifs encoding a domain of interest can be shuffled between the *RTH6* gene disclosed herein and other *rboh* and *rth* genes to obtain a new gene coding for a protein with an improved property of interest, such as an increased K_m in the case of an enzyme. Strategies for such DNA shuffling are known in the art. See, for example, Stemmer (1994) *Proc. Natl. Acad. Sci. USA* 91:10747-10751; Stemmer (1994) *Nature* 370:389-391; Crameri *et al.* (1997) *Nature Biotech.* 15:436-438; Moore *et al.* (1997) *J. Mol. Biol.* 272:336-347; Zhang *et al. (1997) Proc. Natl. Acad. Sci. USA* 94:4504-4509; Crameri *et al.* (1998) *Nature* 391:288-291; and U.S. Patent Nos. 5,605,793 and 5,837,458.

[0094] The following terms are used to describe the sequence relationships between two or more polynucleotides or polypeptides: (a) "reference sequence", (b) "comparison window", (c) "sequence identity", and, (d) "percentage of sequence identity."

[0095] As used herein, "reference sequence" is a defined sequence used as a basis for sequence comparison. A reference sequence may be a subset or the entirety of a specified sequence; for example, as a segment of a full-length cDNA or gene sequence, or the complete cDNA or gene sequence.

[0096] As used herein, "comparison window" makes reference to a contiguous and specified segment of a polynucleotide sequence, wherein the polynucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two polynucleotides. Generally, the comparison window is at least 20 contiguous nucleotides in length, and optionally can be 30, 40, 50, 100, or longer. Those of skill in the art understand that to avoid a high similarity to a reference sequence due to inclusion of gaps in the polynucleotide sequence a gap penalty is typically introduced and is subtracted from the number of matches.

[0097] Methods of alignment of sequences for comparison are well known in the art. Thus, the determination of percent sequence identity between any two sequences can be accomplished using a mathematical algorithm. Non-limiting examples of such mathematical

algorithms are the algorithm of Myers and Miller (1988) *CABIOS* 4:11-17; the local alignment algorithm of Smith *et al.* (1981) *Adv. Appl. Math.* 2:482; the global alignment algorithm of Needleman and Wunsch (1970) *J. Mol. Biol.* 48:443-453; the search-for-local alignment method of Pearson and Lipman (1988) *Proc. Natl. Acad. Sci.* 85:2444-2448; the algorithm of Karlin and Altschul (1990) *Proc. Natl. Acad. Sci. USA* 872264, modified as in Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-5877.

[0098] Computer implementations of these mathematical algorithms can be utilized for comparison of sequences to determine sequence identity. Such implementations include, but are not limited to: CLUSTAL in the PC/Gene program (available from Intelligenetics, Mountain View, California); the ALIGN program (Version 2.0) and GAP, BESTFIT, BLAST, FASTA, and TFASTA in the GCG Wisconsin Genetics Software Package, Version 10 (available from Accelrys Inc., 9685 Scranton Road, San Diego, California, USA). Alignments using these programs can be performed using the default parameters. The CLUSTAL program is well described by Higgins *et al.* (1988) *Gene* 73:237-244 (1988); Higgins *et al.* (1989) *CABIOS* 5:151-153; Corpet *et al.* (1988) *Nucleic Acids Res.* 16:10881-90; Huang *et al.* (1992) *CABIOS* 8:155-65; and Pearson *et al.* (1994) *Meth. Mol. Biol.* 24:307-331. The ALIGN program is based on the algorithm of Myers and Miller (1988) *supra*. A PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used with the ALIGN program when comparing amino acid sequences. The BLAST programs of Altschul *et al* (1990) *J. Mol. Biol.* 215:403 are based on the algorithm of Karlin and Altschul (1990) *supra*. BLAST nucleotide searches can be performed with the BLASTN program, score = 100, wordlength = 12, to obtain nucleotide sequences homologous to a nucleotide sequence encoding a protein of the invention. BLAST protein searches can be performed with the BLASTX program, score = 50, wordlength = 3, to obtain amino acid sequences homologous to a protein or polypeptide of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST (in BLAST 2.0) can be utilized as described in Altschul *et al.* (1997) *Nucleic Acids Res.* 25:3389. Alternatively, PSI-BLAST (in BLAST 2.0) can be used to perform an iterated search that detects distant relationships between molecules. See Altschul *et al.* (1997) *supra*. When utilizing BLAST, Gapped BLAST, PSI-BLAST, the default parameters of the respective programs (e.g., BLASTN for nucleotide sequences, BLASTX for proteins) can be used. Alignment may also be performed manually by inspection.

[0099] Unless otherwise stated, sequence identity/similarity values provided herein refer to the value obtained using GAP Version 10 using the following parameters: % identity and

% similarity for a nucleotide sequence using GAP Weight of 50 and Length Weight of 3, and the nwsgapdna.cmp scoring matrix; % identity and % similarity for an amino acid sequence using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix; or any equivalent program thereof. By "equivalent program" is intended any sequence comparison program that, for any two sequences in question, generates an alignment having identical nucleotide or amino acid residue matches and an identical percent sequence identity when compared to the corresponding alignment generated by GAP Version 10.

[00100] GAP uses the algorithm of Needleman and Wunsch (1970) *J. Mol. Biol.* 48:443-453, to find the alignment of two complete sequences that maximizes the number of matches and minimizes the number of gaps. GAP considers all possible alignments and gap positions and creates the alignment with the largest number of matched bases and the fewest gaps. It allows for the provision of a gap creation penalty and a gap extension penalty in units of matched bases. GAP must make a profit of gap creation penalty number of matches for each gap it inserts. If a gap extension penalty greater than zero is chosen, GAP must, in addition, make a profit for each gap inserted of the length of the gap times the gap extension penalty. Default gap creation penalty values and gap extension penalty values in Version 10 of the GCG Wisconsin Genetics Software Package for protein sequences are 8 and 2, respectively. For nucleotide sequences the default gap creation penalty is 50 while the default gap extension penalty is 3. The gap creation and gap extension penalties can be expressed as an integer selected from the group of integers consisting of from 0 to 200. Thus, for example, the gap creation and gap extension penalties can be 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65 or greater.

[00101] GAP presents one member of the family of best alignments. There may be many members of this family, but no other member has a better quality. GAP displays four figures of merit for alignments: Quality, Ratio, Identity, and Similarity. The Quality is the metric maximized in order to align the sequences. Ratio is the quality divided by the number of bases in the shorter segment. Percent Identity is the percent of the symbols that actually match. Percent Similarity is the percent of the symbols that are similar. Symbols that are across from gaps are ignored. A similarity is scored when the scoring matrix value for a pair of symbols is greater than or equal to 0.50, the similarity threshold. The scoring matrix used in Version 10 of the GCG Wisconsin Genetics Software Package is BLOSUM62 (see Henikoff and Henikoff (1989) *Proc. Natl. Acad. Sci. USA* 89:10915).

[00102] As used herein, "sequence identity" or "identity" in the context of two polynucleotides or polypeptide sequences makes reference to the residues in the two

sequences that are the same when aligned for maximum correspondence over a specified comparison window. When percentage of sequence identity is used in reference to proteins it is recognized that residue positions which are not identical often differ by conservative amino acid substitutions, where amino acid residues are substituted for other amino acid residues with similar chemical properties (e.g., charge or hydrophobicity) and therefore do not change the functional properties of the molecule. When sequences differ in conservative substitutions, the percent sequence identity may be adjusted upwards to correct for the conservative nature of the substitution. Sequences that differ by such conservative substitutions are said to have "sequence similarity" or "similarity". Means for making this adjustment are well known to those of skill in the art. Typically this involves scoring a conservative substitution as a partial rather than a full mismatch, thereby increasing the percentage sequence identity. Thus, for example, where an identical amino acid is given a score of 1 and a non-conservative substitution is given a score of zero, a conservative substitution is given a score between zero and 1. The scoring of conservative substitutions is calculated, e.g., as implemented in the program PC/GENE (Intelligenetics, Mountain View, California).

[00103] As used herein, "percentage of sequence identity" means the value determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison, and multiplying the result by 100 to yield the percentage of sequence identity.

[00104] Sequence alignments and percent identity calculations may be determined using the MEGALIGN® program of the LASERGENE® bioinformatics computing suite (DNASTAR® Inc., Madison, WI). Multiple alignment of the sequences provided herein may be performed using the Clustal V method of alignment (Higgins and Sharp (1989) CABIOS. 5:151 153) with the default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10). Default parameters for pairwise alignments and calculation of percent identity of protein sequences using the Clustal V method are KTUPLE=1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5. For nucleic acids these parameters are KTUPLE=2, GAP PENALTY=5, WINDOW=4 and DIAGONALS

SAVED=4. After alignment of the sequences, using the Clustal V program, it is possible to obtain "percent identity" and "divergence" values by viewing the "sequence distances" table on the same program; unless stated otherwise, percent identities and divergences provided and claimed herein were calculated in this manner.

[00105] Alternatively, the Clustal W method of alignment may be used. The Clustal W method of alignment (described by Higgins and Sharp, CABIOS. 5:151-153 (1989); Higgins, D. G. et al., Comput. Appl. Biosci. 8:189-191 (1992)) can be found in the MEGALIGN® v6.1 program of the LASERGENE® bioinformatics computing suite (DNASTAR® Inc., Madison, Wis.). Default parameters for multiple alignment correspond to GAP PENALTY=10, GAP LENGTH PENALTY=0.2, Delay Divergent Sequences=30%, DNA Transition Weight=0.5, Protein Weight Matrix=Gonnet Series, DNA Weight Matrix=IUB. For pairwise alignments the default parameters are Alignment=Slow-Accurate, Gap Penalty=10.0, Gap Length=0.10, Protein Weight Matrix=Gonnet 250 and DNA Weight Matrix=IUB. After alignment of the sequences using the Clustal W program, it is possible to obtain "percent identity" and "divergence" values by viewing the "sequence distances" table in the same program.

[00106] The use of the term "polynucleotide" is not intended to limit the present invention to polynucleotides comprising DNA. Those of ordinary skill in the art will recognize that polynucleotides can comprise ribonucleotides and combinations of ribonucleotides and deoxyribonucleotides. Such deoxyribonucleotides and ribonucleotides include both naturally occurring molecules and synthetic analogues. The polynucleotides of the invention also encompass all forms of sequences including, but not limited to, single-stranded forms, double-stranded forms, hairpins, stem-and-loop structures, and the like.

IV. PLANTS, PLANT PARTS, AND METHODS OF INTRODUCING SEQUENCES INTO PLANTS

[00107] The methods of the invention involve introducing a polypeptide or polynucleotide into a plant. "Introducing" is intended to mean presenting to the plant the polynucleotide or polypeptide in such a manner that the sequence gains access to the interior of a cell of the plant. The methods of the invention do not depend on a particular method for introducing a sequence into a plant, only that the polynucleotide or polypeptides gains access to the interior of at least one cell of the plant. Methods for introducing polynucleotide or polypeptides into plants are known in the art including, but not limited to, stable transformation methods, transient transformation methods, and virus-mediated methods.

[00108] "Stable transformation" is intended to mean that the nucleotide construct introduced into a plant integrates into the genome of the plant and is capable of being

inherited by the progeny thereof. "Transient transformation" is intended to mean that a polynucleotide is introduced into the plant and does not integrate into the genome of the plant or a polypeptide is introduced into a plant.

[00109] Transformation protocols as well as protocols for introducing polypeptides or polynucleotide sequences into plants may vary depending on the type of plant or plant cell, i.e., monocot or dicot, targeted for transformation. Suitable methods of introducing polypeptides and polynucleotides into plant cells include microinjection (Crossway *et al.* (1986) *Biotechniques* 4:320-334), electroporation (Riggs *et al.* (1986) *Proc. Natl. Acad. Sci. USA* 83:5602-5606, *Agrobacterium*-mediated transformation (U.S. Patent No. 5,563,055 and U.S. Patent No. 5,981,840), direct gene transfer (Paszkowski *et al.* (1984) *EMBO J.* 3:2717-2722), and ballistic particle acceleration (see, for example, U.S. Patent Nos. 4,945,050; U.S. Patent No. 5,879,918; U.S. Patent No. 5,886,244; and, 5,932,782; Tomes *et al.* (1995) in *Plant Cell, Tissue, and Organ Culture: Fundamental Methods*, ed. Gamborg and Phillips (Springer-Verlag, Berlin); McCabe *et al.* (1988) *Biotechnology* 6:923-926); and Lec1 transformation (WO 00/28058). Also see Weissinger *et al.* (1988) *Ann. Rev. Genet.* 22:421-477; Sanford *et al.* (1987) *Particulate Science and Technology* 5:27-37 (onion); Christou *et al.* (1988) *Plant Physiol.* 87:671-674 (soybean); McCabe *et al.* (1988) *Bio/Technology* 6:923-926 (soybean); Finer and McMullen (1991) *In Vitro Cell Dev. Biol.* 27P:175-182 (soybean); Singh *et al.* (1998) *Theor. Appl. Genet.* 96:319-324 (soybean); Datta *et al.* (1990) *Biotechnology* 8:736-740 (rice); Klein *et al.* (1988) *Proc. Natl. Acad. Sci. USA* 85:4305-4309 (maize); Klein *et al.* (1988) *Biotechnology* 6:559-563 (maize); U.S. Patent Nos. 5,240,855; 5,322,783; and, 5,324,646; Klein *et al.* (1988) *Plant Physiol.* 91:440-444 (maize); Fromm *et al.* (1990) *Biotechnology* 8:833-839 (maize); Hooykaas-Van Slogteren *et al.* (1984) *Nature (London)* 311:763-764; U.S. Patent No. 5,736,369 (cereals); Bytebier *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:5345-5349 (Liliaceae); De Wet *et al.* (1985) in *The Experimental Manipulation of Ovule Tissues,* ed. Chapman *et al.* (Longman, New York), pp. 197-209 (pollen); Kaeppler *et al.* (1990) *Plant Cell Reports* 9:415-418 and Kaeppler *et al.* (1992) *Theor. Appl. Genet.* 84:560-566 (whisker-mediated transformation); D'Halluin *et al.* (1992) *Plant Cell* 4:1495-1505 (electroporation); Li *et al.* (1993) *Plant Cell Reports* 12:250-255 and Christou and Ford (1995) *Annals of Botany* 75:407-413 (rice); Osjoda *et al.* (1996) *Nature Biotechnology* 14:745-750 (maize via *Agrobacterium tumefaciens*); all of which are herein incorporated by reference.

[00110] In certain aspects, the *rth6* polynucleotides can be provided to a plant using a variety of transient transformation methods. Such transient transformation methods include,

but are not limited to, the introduction of the protein or variants or fragments thereof directly into the plant or the introduction of the transcript into the plant. Such methods include, for example, microinjection or particle bombardment. See, for example, Crossway et al. (1986) *Mol. Gen. Genet.* 202:179-185; Nomura et al. (1986) *Plant Sci.* 44:53-58; Hepler et al. (1994) *Proc. Natl. Acad. Sci.* 91: 2176-2180 and Hush et al. (1994) *The Journal of Cell Science* 107:775-784, all of which are herein incorporated by reference. Alternatively, polynucleotides can be transiently transformed into the plant using techniques known in the art. Such techniques include viral vector systems and the precipitation of the polynucleotide in a manner that precludes subsequent release of the DNA. Such methods include the use of particles coated with polyethylimine (PEI; Sigma #P3143).

[00111] In other aspects, the *rth6* polynucleotides can be introduced into plants by contacting plants with a virus or viral nucleic acids. Generally, such methods involve incorporating a nucleotide construct of the invention within a viral DNA or RNA molecule. Further, it is recognized that promoters of the invention also encompass promoters utilized for transcription by viral RNA polymerases. Methods for introducing polynucleotides into plants and expressing a protein encoded therein, involving viral DNA or RNA molecules, are known in the art. See, for example, U.S. Patent Nos. 5,889,191, 5,889,190, 5,866,785, 5,589,367, 5,316,931, and Porta *et al.* (1996) *Molecular Biotechnology 5*:209-221; herein incorporated by reference.

[00112] Methods are known in the art for the targeted insertion of a polynucleotide at a specific location in the plant genome. In one aspect, the insertion of the polynucleotide at a desired genomic location is achieved using a site-specific recombination system. See, for example, WO99/25821, WO99/25854, WO99/25840, WO99/25855, and WO99/25853, all of which are herein incorporated by reference. Briefly, the expression cassette disclosed herein can be contained in transfer cassette flanked by two non-recombinogenic recombination sites. The transfer cassette is introduced into a plant having stably incorporated into its genome a target site which is flanked by two non-recombinogenic recombination sites that correspond to the sites of the transfer cassette. An appropriate recombinase is provided and the transfer cassette is integrated at the target site. The polynucleotide of interest is thereby integrated at a specific chromosomal position in the plant genome.

[00113] It is therefore recognized that methods of the present invention do not depend on the incorporation of the entire polynucleotide into the genome, only that the plant or cell thereof is altered as a result of the introduction of the polynucleotide into a cell. In one aspect of the invention, the genome may be altered following the introduction of the

polynucleotide into a cell. For example, the polynucleotide, or any part thereof, may incorporate into the genome of the plant. Alterations to the genome of the present invention include, but are not limited to, additions, deletions, and substitutions of nucleotides into the genome. While the methods of the present invention do not depend on additions, deletions, and substitutions of any particular number of nucleotides, it is recognized that such additions, deletions, or substitutions comprises at least one nucleotide.

[00114] As indicated, the invention also encompasses plants where the endogenous RTH6 expression is enhanced. Thus, the level and/or activity of an RTH6 polypeptide can be increased by altering the gene encoding the RTH6 family polypeptide or its promoter. See, e.g., Kmiec, U.S. Patent 5,565,350; and Zarling *et al.*, PCT/US93/03868. Therefore mutagenized plants that carry mutations in RTH6 family genes, where the mutations increase expression of the RTH6 family gene or increase the root hair formation and growth activity of the encoded RTH6 family polypeptide are provided.

[00115] Levels of endogenous RTH6 can be increased by methods known in the art. For example, endogenous RTH6 can be increased by the introduction of sequences of interest that upregulate endogenous expression. Such sequences of interest can include heterologous promoters, siRNA targeted to genomic regulatory elements, enhancer elements, and/or booster sequences. See, e.g., U. S. Patent 5,939,541; U.S. Patent 6,576,442; US 2012/0036594; and WO 1991/009955. Other methods of inserting sequences of interest to upregulate endogenous expression can include the use of various meganucleases to target polynucleotides. Such meganucleases are set forth in WO 2009/114321 (herein incorporated by reference), which describes "custom" meganucleases. See, also, Gao *et al.* (2010) *Plant Journal* 1:176-187. Additional methods of inserting sequences of interest to upregulate endogenous expression of a sequence of interest that can be employed, include but are not limited to the use of ZnFingers, meganucleases, and, TAL nucleases. See, for example, WO2010079430, WO2011072246, and US20110201118, each of which is herein incorporated by reference in their entirety.

[00116] Various aspects include methods of detection and use of polymorphisms in the maize RTH6 promoter that are associated with increased expression of the RTH6 polypeptide.

[00117] The cells that have been transformed can be grown into plants in accordance with conventional ways. See, for example, McCormick *et al.* (1986) *Plant Cell Reports* 5:81-84. These plants can then be grown, and either pollinated with the same transformed strain or different strains, and the resulting progeny having constitutive expression of the desired

phenotypic characteristic identified. Two or more generations can be grown to ensure that expression of the desired phenotypic characteristic is stably maintained and inherited and then seeds harvested to ensure expression of the desired phenotypic characteristic has been achieved. In this manner, the present invention provides transformed seed (also referred to as "transgenic seed") having a polynucleotide of the invention, for example, an expression cassette of the invention, stably incorporated into their genome.

[00118] As used herein, the term plant also includes plant cells, plant protoplasts, plant cell tissue cultures from which plants can be regenerated, plant calli, plant clumps, and plant cells that are intact in plants or parts of plants such as embryos, pollen, ovules, seeds, leaves, flowers, branches, fruit, kernels, ears, cobs, husks, stalks, roots, root tips, anthers, and the like. Grain is intended to mean the mature seed produced by commercial growers for purposes other than growing or reproducing the species. Progeny, variants, and mutants of the regenerated plants are also included within the scope of the invention, provided that these parts comprise the introduced polynucleotides.

[00119] In some aspects, the composition is a seed of a plant comprising an expression construct that includes a promoter that drives expression in roots of a plant operably linked to a polynucleotide that encodes an RTH6 family member as described herein.

[00120] The expression constructs disclosed herein can be used for transformation of any plant species, including, but not limited to, monocots and dicots. Examples of plant species of interest include, but are not limited to, corn (*Zea mays*), *Brassica* sp. (e.g., *B. napus, B. rapa, B. juncea*), particularly those *Brassica* species useful as sources of seed oil, alfalfa (*Medicago sativa*), rice (*Oryza sativa*), rye (*Secale cereale*), sorghum (*Sorghum bicolor, Sorghum vulgare*), millet (e.g., pearl millet (*Pennisetum glaucum*), proso millet (*Panicum miliaceum*), foxtail millet (*Setaria italica*), finger millet (*Eleusine coracana*)), sunflower (*Helianthus annuus*), safflower (*Carthamus tinctorius*), wheat (*Triticum aestivum*), soybean (*Glycine max*), tobacco (*Nicotiana tabacum*), potato (*Solanum tuberosum*), peanuts (*Arachis hypogaea*), cotton (*Gossypium barbadense*, *Gossypium hirsutum*), sweet potato (*Ipomoea batatus*), cassava (*Manihot esculenta*), coffee (*Coffea* spp.), coconut (*Cocos nucifera*), pineapple (*Ananas comosus*), citrus trees (*Citrus* spp.), cocoa (*Theobroma cacao*), tea (*Camellia sinensis*), banana (*Musa* spp.), avocado (*Persea americana*), fig (*Ficus casica*), guava (*Psidium guajava*), mango (*Mangifera indica*), olive (*Olea europaea*), papaya (*Carica papaya*), cashew (*Anacardium occidentale*), macadamia (*Macadamia integrifolia*), almond (*Prunus amygdalus*), sugar beets (*Beta vulgaris*), sugarcane (*Saccharum* spp.), oats, barley, vegetables, ornamentals, and conifers.

[00121] Vegetables include tomatoes (*Lycopersicon esculentum*), lettuce (e.g., *Lactuca sativa*), green beans (*Phaseolus vulgaris*), lima beans (*Phaseolus limensis*), peas (*Lathyrus* spp.), and members of the genus *Cucumis* such as cucumber (*C. sativus*), cantaloupe (*C. cantalupensis*), and musk melon (*C. melo*). Ornamentals include azalea (*Rhododendron* spp.), hydrangea (*Macrophylla hydrangea*), hibiscus (*Hibiscus rosasanensis*), roses (*Rosa* spp.), tulips (*Tulipa* spp.), daffodils (*Narcissus* spp.), petunias (*Petunia hybrida*), carnation (*Dianthus caryophyllus*), poinsettia (*Euphorbia pulcherrima*), and chrysanthemum.

[00122] Conifers that can be employed in practicing the present invention include, for example, pines such as loblolly pine (*Pinus taeda*), slash pine (*Pinus elliotii*), ponderosa pine (*Pinus ponderosa*), lodgepole pine (*Pinus contorta*), and Monterey pine (*Pinus radiata*); Douglas-fir (*Pseudotsuga menziesii*); Western hemlock (*Tsuga canadensis*); Sitka spruce (*Picea glauca*); redwood (*Sequoia sempervirens*); true firs such as silver fir (*Abies amabilis*) and balsam fir (*Abies balsamea*); and cedars such as Western red cedar (*Thuja plicata*) and Alaska yellow-cedar (C*hamaecyparis nootkatensis*). In specific aspects, plants of the present invention are crop plants (for example, corn, alfalfa, sunflower, *Brassica*, soybean, cotton, safflower, peanut, sorghum, wheat, millet, tobacco, etc.). In various aspects, corn and soybean and sugarcane plants are optimal, and in yet various aspects, corn plants are optimal.

[00123] Other plants of interest include grain plants that provide seeds of interest, oil-seed plants, and leguminous plants. Seeds of interest include grain seeds, such as corn, wheat, barley, rice, sorghum, rye, etc. Oil-seed plants include cotton, soybean, safflower, sunflower, *Brassica*, maize, alfalfa, palm, coconut, etc. Leguminous plants include beans and peas. Beans include guar, locust bean, fenugreek, soybean, garden beans, cowpea, mungbean, lima bean, fava bean, lentils, chickpea, etc.

[00124] Another aspect is a method for transforming a cell (or microorganism) comprising transforming a cell (or microorganism) with any of the isolated polynucleotides or expression constructs of the present invention. The cell (or microorganism) transformed by this method is also included. In particular aspects, the cell is eukaryotic cell, e.g., a yeast, insect or plant cell, or prokaryotic, e.g., a bacterial cell. The microorganism can be Agrobacterium, e.g. Agrobacterium tumefaciens or Agrobacterium rhizogenes.

[00125] Another aspect is an expression construct (and corresponding cells, plants, seeds and methods) comprising a promoter (heterologous or native) that drives expression in a plant root operably linked to a polynucleotide, wherein said polynucleotide encodes a polypeptide that comprises the following regions: eight trans-membrane (TM) domains, an N-terminal RING/Ubox like Zinc-finger domain, and a C-terminal cellulose synthase-like protein

domain, wherein each of these regions have at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% amino acid sequence identity to the corresponding regions from SEQ ID NO: 3, wherein said polypeptide has cellulose synthase activity, e.g., ß-glycan synthase activity, .

V. STACKING OF TRAITS IN TRANSGENIC PLANTS

[00126] Disclosed transgenic plants may comprise a stack of one or more polynucleotides as set forth in SEQ ID NOs: 1 or 2, or variants or fragments thereof, or complements thereof, as disclosed herein with one or more additional polynucleotides resulting in the production or suppression of multiple polypeptide sequences. Transgenic plants comprising stacks of polynucleotide sequences can be obtained by either or both of traditional breeding methods or through genetic engineering methods. These methods include, but are not limited to, breeding individual lines each comprising a polynucleotide of interest, transforming a transgenic plant comprising an expression construct comprising various target polynucleotides as set forth in SEQ ID NOs: 1 or 2, or variants or fragments thereof, or complements thereof, as disclosed herein with a subsequent gene and co-transformation of genes into a single plant cell. As used herein, the term "stacked" includes having the multiple traits present in the same plant (i.e., both traits are incorporated into the nuclear genome, one trait is incorporated into the nuclear genome and one trait is incorporated into the genome of a plastid or both traits are incorporated into the genome of a plastid). In one non-limiting example, "stacked traits" comprise a molecular stack where the sequences are physically adjacent to each other. A trait, as used herein, refers to the phenotype derived from a particular sequence or groups of sequences. Co-transformation of polynucleotides can be carried out using single transformation vectors comprising multiple polynucleotides or polynucleotides carried separately on multiple vectors. If the sequences are stacked by genetically transforming the plants, the polynucleotide sequences of interest can be combined at any time and in any order. The traits can be introduced simultaneously in a co-transformation protocol with the polynucleotides of interest provided by any combination of transformation cassettes. For example, if two sequences will be introduced, the two sequences can be contained in separate transformation cassettes (trans) or contained on the same transformation cassette (cis). Expression of the sequences can be driven by the same promoter or by different promoters. It is further recognized that polynucleotide sequences can be stacked at a desired genomic location using a site-specific recombination system. *See*, for example, WO 1999/25821, WO 1999/25854, WO 1999/25840, WO 1999/25855 and WO 1999/25853, all of which are herein incorporated by reference.

[00127] In some aspects the various target polynucleotides as set forth in SEQ ID NOs: 1 or 2, variants or fragments thereof, or complements thereof, as disclosed herein, alone or stacked with one or more additional insect resistance traits can be stacked with one or more additional input traits (e.g., herbicide resistance, fungal resistance, virus resistance, stress tolerance, disease resistance, male sterility, stalk strength, and the like) or output traits (e.g., increased yield, modified starches, improved oil profile, balanced amino acids, high lysine or methionine, increased digestibility, improved fiber quality, drought resistance, and the like). Thus, the polynucleotide aspects can be used to provide a complete agronomic package of improved crop quality with the ability to flexibly and cost effectively control any number of agronomic pests.

[00128] Transgenes useful for stacking include, but are not limited to, to those as described herein, including, for example, conferring resistance to herbicide, insect and/or disease. Herbicide tolerance or resistance genes include, for example, genes encoding tolerance to glyphosate, glufosinate, dicamba, 2,4-D and HPPD inhibitors. Insect tolerance genes include *B. thuringiensis* (Bt) and non-Bt polypeptides, and RNAi polynucleotides targeting one or more insect pests.

a. Genes that affect abiotic stress resistance.

[00129] Including but not limited to flowering, ear and seed development, enhancement of nitrogen utilization efficiency, altered nitrogen responsiveness, drought resistance or tolerance, cold resistance or tolerance and salt resistance or tolerance and increased yield under stress. Non-limiting examples include: (A) For example, see: WO 2000/73475 where water use efficiency is altered through alteration of malate; U.S. Pat. Nos. 5,892,009, 5,965,705, 5,929,305, 5,891,859, 6,417,428, 6,664,446, 6,706,866, 6,717,034, 6,801,104, WO 2000/060089, WO 2001/026459, WO 2001/035725, WO 2001/034726, WO 2001/035727, WO 2001/036444, WO 2001/036597, WO 2001/036598, WO 2002/015675, WO 2002/017430, WO 2002/077185, WO 2002/079403, WO 2003/013227, WO 2003/013228, WO 2003/014327, WO 2004/031349, WO 2004/076638, WO 199809521; (B) WO 199938977 describing genes, including CBF genes and transcription factors effective in mitigating the negative effects of freezing, high salinity and drought on plants, as well as conferring other positive effects on plant phenotype; (C) US Patent Application Publication Number 2004/0148654 and WO 2001/36596 where abscisic acid is altered in plants resulting in improved plant phenotype such as increased yield and/or increased tolerance to abiotic stress; (D) WO 2000/006341, WO 2004/090143, U.S. Pat. Nos. 7,531,723 and 6,992,237 where cytokinin expression is modified resulting in plants with increased stress tolerance,

such as drought tolerance, and/or increased yield. Also see, WO 2002/02776, WO 2003/052063, JP 2002/281975, U.S. Pat. No. 6,084,153, WO 2001/64898, U.S. Pat. No. 6,177,275 and U.S. Pat. No. 6,107,547 (enhancement of nitrogen utilization and altered nitrogen responsiveness); (E) For ethylene alteration, see, US Patent Application Publication Number 2004/0128719, US Patent Application Publication Number 2003/0166197 and WO 2000/32761; (F) For plant transcription factors or transcriptional regulators of abiotic stress, see, e.g., U.S. Patent Application Publication Number 2004/0098764 or US Patent Application Publication Number 2004/0078852; (G) Genes that increase expression of vacuolar pyrophosphatase such as AVP1 (U.S. Pat. No. 8,058,515) for increased yield; nucleic acid encoding a HSFA4 or a HSFA5 (Heat Shock Factor of the class A4 or A5) polypeptides, an oligopeptide transporter protein (OPT4-like) polypeptide; a plastochron2-like (PLA2-like) polypeptide or a Wuschel related homeobox 1-like (WOX1-like) polypeptide (U.S. Patent Application Publication Number US 2011/0283420); (H) Down regulation of polynucleotides encoding poly (ADP-ribose) polymerase (PARP) proteins to modulate programmed cell death (U.S. Pat. No. 8,058,510) for increased vigor; (I) Polynucleotide encoding DTP21 polypeptides for conferring drought resistance (US Patent Application Publication Number US 2011/0277181); (J) Nucleotide sequences encoding ACC Synthase 3 (ACS3) proteins for modulating development, modulating response to stress, and modulating stress tolerance (U.S. Patent Application Publication Number US 2010/0287669); (K) Polynucleotides that encode proteins that confer a drought tolerance phenotype (DTP) for conferring drought resistance (WO 2012/058528); (L) Tocopherol cyclase (TC) genes for conferring drought and salt tolerance (US Patent Application Publication Number 2012/0272352); (M) CAAX amino terminal family proteins for stress tolerance (U.S. Pat. No. 8,338,661); (N) Mutations in the SAL1 encoding gene have increased stress tolerance, including increased drought resistant (U.S. Patent Application Publication Number 2010/0257633); (O) Expression of a nucleic acid sequence encoding a polypeptide selected from the group consisting of: GRF polypeptide, RAA1-like polypeptide, SYR polypeptide, ARKL polypeptide, and YTP polypeptide increasing yield-related traits (US Patent Application Publication Number 2011/0061133); and (P) Modulating expression in a plant of a nucleic acid encoding a Class III Trehalose Phosphate Phosphatase (TPP) polypeptide for enhancing yield-related traits in plants, particularly increasing seed yield (U.S. Patent Application Publication Number 2010/0024067).

[00130] Other genes and transcription factors that affect plant growth and agronomic traits such as yield, flowering, plant growth and/or plant structure, can be introduced or

introgressed into plants, see e.g., WO 1997/49811 (LHY), WO 1998/56918 (ESD4), WO 1997/10339 and U.S. Pat. No. 6,573,430 (TFL), U.S. Pat. No. 6,713,663 (FT), WO 1996/14414 (CON), WO 1996/38560, WO 2001/21822 (VRN1), WO 2000/44918 (VRN2), WO 1999/49064 (GI), WO 2000/46358 (FR1), WO 1997/29123, U.S. Pat. No. 6,794,560, U.S. Pat. No. 6,307,126 (GAI), WO 1999/09174 (D8 and Rht) and WO 2004/076638 and WO 2004/031349 (transcription factors).

b. Genes that Confer Increased Yield.

[00131] Non-limiting examples of genes that confer increased yield are: (A) A transgenic crop plant transformed by a 1-AminoCyclopropane-1-Carboxylate Deaminase-like Polypeptide (ACCDP) coding nucleic acid, wherein expression of the nucleic acid sequence in the crop plant results in the plant's increased root growth, and/or increased yield, and/or increased tolerance to environmental stress as compared to a wild type variety of the plant (U.S. Pat. No. 8,097,769); (B) Over-expression of maize zinc finger protein gene (Zm-ZFP1) using a seed preferred promoter has been shown to enhance plant growth, increase kernel number and total kernel weight per plant (US Patent Application Publication Number 2012/0079623); (C) Constitutive over-expression of maize lateral organ boundaries (LOB) domain protein (Zm-LOBDP1) has been shown to increase kernel number and total kernel weight per plant (US Patent Application Publication Number 2012/0079622); (D) Enhancing yield-related traits in plants by modulating expression in a plant of a nucleic acid encoding a VIM1 (Variant in Methylation 1)-like polypeptide or a VTC2-like (GDP-L-galactose phosphorylase) polypeptide or a DUF1685 polypeptide or an ARF6-like (Auxin Responsive Factor) polypeptide (WO 2012/038893); (E) Modulating expression in a plant of a nucleic acid encoding a Ste20-like polypeptide or a homologue thereof gives plants having increased yield relative to control plants (EP 2431472); and (F) Genes encoding nucleoside diphosphatase kinase (NDK) polypeptides and homologs thereof for modifying the plant's root architecture (US Patent Application Publication Number 2009/0064373).

VI. STACKING WITH OTHER RTH GENES

[00132] Disclosed transgenic plants may comprise a stack of one or more polynucleotides as set forth in SEQ ID NOs: 1 or 2 for RTH6, or variants or fragments thereof, or complements thereof, as disclosed herein with one or more additional polynucleotides for other members of the *roothairless* family of genes, including, but not limited to RTH5. Transgenic plants comprising stacks of *roothairless* polynucleotide sequences, *e.g.*, a stack comprising polynucleotide sequences for RTH5 and RTH6, can be obtained by either or both of traditional breeding methods or through genetic engineering methods. These methods

include, but are not limited to, breeding individual lines each comprising a polynucleotide of interest, transforming a transgenic plant comprising an expression construct comprising various target polynucleotides as set forth in SEQ ID NOs: 1 or 2, or variants or fragments thereof, or complements thereof, as disclosed herein with a subsequent *roothairless* gene and co-transformation of genes into a single plant cell. Plants comprising stacked *roothairless* genes, *e.g.*, RTH5 and RTH6 can be produced using methods disclosed herein.

[00133] In some aspects the various target polynucleotides as set forth in SEQ ID NOs: 1 or 2, variants or fragments thereof, or complements thereof, as disclosed herein, alone or stacked with one or more additional insect resistance traits can be stacked with one or more other *roothairless* genes. Thus, stacking of *roothairless* genes can be used to provide a complete agronomic package of improved drought tolerance in a plant, increased nutrient uptake in a plant, and reduced soil erosion in a crop field.

[00134] In various aspects, transgenic plants comprise a stack of one or more polynucleotides as set forth in SEQ ID NOs: 1 or 2 for RTH6 with one or more polynucleotides for RTH5, such as those set forth in WO2015/084969. Incorporated by reference are SEQ ID NO: 2 of WO2015/084969, and a nucleotide sequence encoding the amino acid sequence set forth in SEQ ID NO: 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, or 13 of WO2015/084969. Polynucleotides encoding RTH5 family members and include without limitation, the polynucleotides corresponding to SEQ ID NOs: 1 and 2 of WO2015/084969, as well as active fragments and variants thereof. RTH5 family members include polynucleotides having sequences with about 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.3%, 99.5% or more sequence identity to SEQ ID NO: 2 of WO2015/084969 and which encode RTH 5 polypeptides that have the ability to promote root hair formation and growth in a plant.

[00135] As referred to herein, the "RTH5 family" refers to a family of polynucleotides and polypeptides that have high sequence identity with and include the *Zea mays* L. (Maize) *roothairless5* (*rth5*) gene, cDNA, or polypeptide described in WO2015/084969 as SEQ ID NOs: 1, 2, and 3, respectively, and promote the formation and growth of root hairs in a plant. An "RTH5 family member" refers to a specific polynucleotide or polypeptide included in the RTH5 family (an RTH5 polynucleotide or RTH5 polypeptide, respectively). As an example, maize *roothairless5* is an RTH5 family member. The maize "*roothairless5*" and "*rth5*" genomic sequence, cDNA, and polypeptide are provided in SEQ ID NOs: 1, 2, and 3, respectively, of WO2015/084969.

[00136] The RTH5 family demonstrates nicotinamide adenine dinucleotide phosphate-oxidase (NADPH oxidase or NOX) activity. As referred to herein, "NADPH oxidase activity" or "NOX activity" refers to the enzymatic activity of the polypeptide that produces the reactive oxygen species (ROS) superoxide (O_2^-) using an electron donor. Without wishing to be limited to a particular mechanism of action, ROS and ROS-related proteins are thought to play an important role in root hair tip growth. The methods of the invention utilize RTH5 family polynucleotides to increase the expression of RTH5 family polypeptides, which advantageously increase root hair formation and growth.

VII. REDUCING EXPRESSION OF RTH6

[00137] Disclosed are methods for increasing resistance of a plant to a root pathogen, the method comprising the step of reducing the expression of an *rth6* gene family member or inhibiting the activity a *rth6* polypeptide. In various aspects, the expression of an *rth6* gene can be reduced by using silencing elements comprising SEQ ID NOs: 1 or 2, or variants or fragments thereof.

[00138] As used herein, a "target sequence" or "target polynucleotide" comprises any sequence in the pest that one desires to reduce the level of expression thereof. In certain aspects, modifying the level of the target sequence in the plant results in beneficial alteration of one or more agronomic characteristics. For instance, the target sequence may be essential for growth and development. Non-limiting examples of target sequences include a polynucleotide set forth in SEQ ID NOs:1 or 2, or variants and fragments thereof, and complements thereof.

[00139] By "silencing element" is intended a polynucleotide which when contacted by or ingested by a plant insect pest, is capable of reducing or eliminating the level or expression of a target polynucleotide or the polypeptide encoded thereby, and a silencing element may include a polynucleotide that encodes the polynucleotide which when contacted by or ingested by a pest, is capable of reducing or eliminating the level or expression of a target polynucleotide or the polypeptide encoded thereby. Accordingly, it is to be understood that "silencing element," as used herein, comprises polynucleotides such as RNA constructs, DNA constructs that encode the RNA constructs, expression constructs comprising the DNA constructs. In one aspect, the silencing element employed can reduce or eliminate the expression level of the target sequence by influencing the level of the target RNA transcript or, alternatively, by influencing translation and thereby affecting the level of the encoded polypeptide. Methods to assay for functional silencing elements that are capable of reducing or eliminating the level of a sequence of interest are disclosed elsewhere herein. A single

polynucleotide employed in the disclosed methods can comprise one or more silencing elements to the same or different target polynucleotides. The silencing element can be produced *in vivo* (i.e., in a host cell such as a plant or microorganism) or *in vitro*. It is to be understood that "silencing element," as used herein, is intended to comprise polynucleotides such as RNA constructs, DNA constructs that encode the RNA constructs, and/or expression constructs comprising the DNA constructs.

[00140] In certain aspects, a silencing element may comprise a chimeric construction molecule comprising two or more disclosed sequences. For example, the chimeric construction may be a hairpin or dsRNA comprising a disclosed sequence. A chimera may comprise two or more disclosed sequences. In one aspect, a chimera contemplates two complementary sequences set forth herein having some degree of mismatch between the complementary sequences such that the two sequences are not perfect complements of one another. Providing at least two different sequences in a single silencing element may allow for targeting multiple genes using one silencing element and/or for example, one expression cassette. Targeting multiple genes may allow for slowing or reducing the possibility of resistance by the pest. In addition, providing multiple targeting ability in one expressed molecule may reduce the expression burden of the transformed plant or plant product, or provide topical treatments that are capable of targeting multiple hosts with one application.

[00141] In certain aspects, the target sequence is not endogenous to the plant. In other aspects, while the silencing element controls pests, preferably the silencing element has no effect on the normal plant or plant part.

[00142] As discussed in further detail below, silencing elements can include, but are not limited to, a sense suppression element, an antisense suppression element, a double stranded RNA, a siRNA, an amiRNA, a miRNA, or a hairpin suppression element. In various aspects, silencing elements may comprise a chimera where two or more disclosed sequences or active fragments or variants, or complements thereof, are found in the same RNA molecule. In various aspects, a disclosed sequence or active fragment or variant, or complement thereof, may be present as more than one copy in a DNA construct, silencing element, DNA molecule or RNA molecule. In a hairpin or dsRNA molecule, the location of a sense or antisense sequence in the molecule, for example, in which sequence is transcribed first or is located on a particular terminus of the RNA molecule, is not limiting to the disclosed sequences, and the dsRNA is not to be limited by disclosures herein of a particular location for such a sequence. Non-limiting examples of silencing elements that can be employed to decrease expression of these target sequences comprise fragments or variants of the sense or antisense sequence, or

alternatively consists of the sense or antisense sequence of a sequence set forth in SEQ ID NOs:1 or 2, or variants and fragments thereof, and complements thereof. The silencing element can further comprise additional sequences that advantageously effect transcription and/or the stability of a resulting transcript. For example, the silencing elements can comprise at least one thymine residue at the 3' end. This can aid in stabilization. Thus, the silencing elements can have at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more thymine residues at the 3' end. As discussed in further detail below, enhancer suppressor elements can also be employed in conjunction with the silencing elements disclosed herein.

[00143] By "reduces" or "reducing" the expression level of a polynucleotide or a polypeptide encoded thereby is intended to mean, the polynucleotide or polypeptide level of the target sequence is statistically lower than the polynucleotide level or polypeptide level of the same target sequence in an appropriate control pest which is not exposed to (i.e., has not ingested or come into contact with) the silencing element. In particular aspects, methods and/or compositions disclosed herein reduce the polynucleotide level and/or the polypeptide level of the target sequence in a plant insect pest to less than 95%, less than 90%, less than 80%, less than 70%, less than 60%, less than 50%, less than 40%, less than 30%, less than 20%, less than 10%, or less than 5% of the polynucleotide level, or the level of the polypeptide encoded thereby, of the same target sequence in an appropriate control pest. Methods to assay for the level of the RNA transcript, the level of the encoded polypeptide, or the activity of the polynucleotide or polypeptide are discussed elsewhere herein.

[00144] A method is further provided for identifying a silencing element from the target polynucleotides set forth in SEQ ID NOs:1 or 2, or variants and fragments thereof, and complements thereof. Such methods comprise obtaining a candidate fragment of any one of SEQ ID NOs:1 or 2, or variants and fragments thereof, and complements thereof, which is of sufficient length to act as a silencing element and thereby reduce the expression of the target polynucleotide and/or control a desired pest; expressing said candidate polynucleotide fragment in an appropriate expression cassette to produce a candidate silencing element and determining is said candidate polynucleotide fragment has the activity of a silencing element and thereby reduce the expression of the target polynucleotide and/or controls a desired pest. Methods of identifying such candidate fragments based on the desired pathway for suppression, in light of the teachings provided herein, are known. For example, various bioinformatics programs can be employed to identify the region of the target polynucleotides that could be exploited to generate a silencing element. See, for example, Elbahir *et al.* (2001) *Genes and Development* 15:188-200, Schwartz *et al.* (2003) *Cell* 115:199-208,

Khvorova *et al.* (2003) *Cell* 115:209-216. See also, the siRNA tools available at the Whitehead Institute for Biomedical Research, which calculates the binding energies for both sense and antisense siRNAs. In various aspects, it is to be understand that the term "...SEQ ID NOs: 1 or 2, or variants or fragments thereof, or complements thereof..." is intended to mean that the disclosed sequences comprise SEQ ID NOs: 1 or 2, and/or fragments of SEQ ID NOs:1 or 2, and/or variants of SEQ ID NOs: 1 or 2, and/or the complements of SEQ ID NOs: 1 or 2, the variants of SEQ ID NOs: 1 or 2, and/or the fragments of SEQ ID NOs:1 or 2, individually (or) or inclusive of some or all listed sequences.

i. Sense Suppression Elements.

[00145] As used herein, a "sense suppression element" comprises a polynucleotide designed to express an RNA molecule corresponding to at least a part of a target messenger RNA in the "sense" orientation. Expression of the RNA molecule comprising the sense suppression element reduces or eliminates the level of the target polynucleotide or the polypeptide encoded thereby. The polynucleotide comprising the sense suppression element may correspond to all or part of the sequence of the target polynucleotide, all or part of the 5' and/or 3' untranslated region of the target polynucleotide, all or part of the coding sequence of the target polynucleotide, or all or part of both the coding sequence and the untranslated regions of the target polynucleotide.

[00146] Typically, a sense suppression element has substantial sequence identity to the target polynucleotide, typically greater than about 65% sequence identity, greater than about 85% sequence identity, about 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity. See, U.S. Patent Nos. 5,283,184 and 5,034,323; herein incorporated by reference. The sense suppression element can be any length so long as it allows for the suppression of the targeted sequence. The sense suppression element can be, for example, 15, 16, 17, 18, 19, 20, 22, 25, 30, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 900, 1000, 1100, 1200, 1300 nucleotides or longer of the target polynucleotides set forth in any of SEQ ID NOS.: 1-53, or variants and fragments thereof, and complements thereof. In other aspects, the sense suppression element can be, for example, about 15-25, 19-35, 19-50, 25-100, 100-150, 150-200, 200-250, 250-300, 300-350, 350-400, 450-500, 500-550, 550-600, 600-650, 650-700, 700-750, 750-800, 800-850, 850-900, 900-950, 950-1000, 1000-1050, 1050-1100, 1100-1200, 1200-1300, 1300-1400, 1400-1500, 1500-1600, 1600-1700, 1700-1800 nucleotides or longer of the target polynucleotides set forth in any of SEQ ID NOs:1 or 2, or variants and fragments thereof, and complements thereof.

ii. Antisense Suppression Elements.

[00147] As used herein, an "antisense suppression element" comprises a polynucleotide which is designed to express an RNA molecule complementary to all or part of a target messenger RNA. Expression of the antisense RNA suppression element reduces or eliminates the level of the target polynucleotide. The polynucleotide for use in antisense suppression may correspond to all or part of the complement of the sequence encoding the target polynucleotide, all or part of the complement of the 5' and/or 3' untranslated region of the target polynucleotide, all or part of the complement of the coding sequence of the target polynucleotide, or all or part of the complement of both the coding sequence and the untranslated regions of the target polynucleotide. In addition, the antisense suppression element may be fully complementary (i.e., 100% identical to the complement of the target sequence) or partially complementary (i.e., less than 100% identical to the complement of the target sequence) to the target polynucleotide. In certain aspects, the antisense suppression element comprises at least 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% sequence complementarity to the target polynucleotide. Antisense suppression may be used to inhibit the expression of multiple proteins in the same plant. See, for example, U.S. Patent No. 5,942,657. Furthermore, the antisense suppression element can be complementary to a portion of the target polynucleotide. Generally, sequences of at least 15, 16, 17, 18, 19, 20, 22, 25, 50, 100, 200, 300, 400, 450 nucleotides or greater of the sequence set forth in any of SEQ ID NOS.: 1-53, or variants and fragments thereof, and complements thereof may be used. Methods for using antisense suppression to inhibit the expression of endogenous genes in plants are described, for example, in Liu et al (2002) Plant Physiol. 129:1732-1743 and U.S. Patent No. 5,942,657, which is herein incorporated by reference.

iii. Double Stranded RNA Suppression Element.

[00148] A "double stranded RNA silencing element" or "dsRNA", which may also be referred to as "dsRNA construct", comprises at least one transcript that is capable of forming a dsRNA either before or after ingestion by a plant insect pest. Thus, a "dsRNA silencing element" includes a dsRNA, a transcript or polyribonucleotide capable of forming a dsRNA or more than one transcript or polyribonucleotide capable of forming a dsRNA. "Double stranded RNA" or "dsRNA" refers to a polyribonucleotide structure formed either by a single self-complementary RNA molecule or a polyribonucleotide structure formed by the expression of at least two distinct RNA strands. The dsRNA molecule(s) employed in the disclosed methods and compositions mediate the reduction of expression of a target sequence, for example, by mediating RNA interference "RNAi" or gene silencing in a

sequence-specific manner. In various aspects, the dsRNA is capable of reducing or eliminating the level or expression of a target polynucleotide or the polypeptide encoded thereby in a plant insect pest.

[00149] The dsRNA can reduce or eliminate the expression level of the target sequence by influencing the level of the target RNA transcript, by influencing translation and thereby affecting the level of the encoded polypeptide, or by influencing expression at the pre-transcriptional level (i.e., via the modulation of chromatin structure, methylation pattern, etc., to alter gene expression). For example, see Verdel *et al.* (2004) *Science* 303:672-676; Pal-Bhadra *et al.* (2004) *Science* 303:669-672; Allshire (2002) *Science* 297:1818-1819; Volpe *et al.* (2002) *Science* 297:1833-1837; Jenuwein (2002) *Science* 297:2215-2218; and Hall *et al.* (2002) *Science* 297:2232-2237. Methods to assay for functional dsRNA that are capable of reducing or eliminating the level of a sequence of interest are disclosed elsewhere herein. Accordingly, as used herein, the term "dsRNA" is meant to encompass other terms used to describe nucleic acid molecules that are capable of mediating RNA interference or gene silencing, including, for example, short-interfering RNA (siRNA), double-stranded RNA (dsRNA), micro-RNA (miRNA), hairpin RNA, short hairpin RNA (shRNA), post-transcriptional gene silencing RNA (ptgsRNA), and others.

[00150] In certain aspects, at least one strand of the duplex or double-stranded region of the dsRNA shares sufficient sequence identity or sequence complementarity to the target polynucleotide to allow the dsRNA to reduce the level of expression of the target sequence. As used herein, the strand that is complementary to the target polynucleotide is the "antisense strand" and the strand homologous to the target polynucleotide is the "sense strand."

[00151] In another aspect, the dsRNA comprises a hairpin RNA. A hairpin RNA comprises an RNA molecule that is capable of folding back onto itself to form a double stranded structure. Multiple structures can be employed as hairpin elements. In certain aspects, the dsRNA suppression element comprises a hairpin element which comprises in the following order, a first segment, a second segment, and a third segment, where the first and the third segment share sufficient complementarity to allow the transcribed RNA to form a double-stranded stem-loop structure.

[00152] The "second segment" of the hairpin comprises a "loop" or a "loop region." These terms are used synonymously herein and are to be construed broadly to comprise any nucleotide sequence that confers enough flexibility to allow self-pairing to occur between complementary regions of a polynucleotide (i.e., segments 1 and 3 which form the stem of the hairpin). For example, in some aspects, the loop region may be substantially single

stranded and act as a spacer between the self-complementary regions of the hairpin stem-loop. In some aspects, the loop region can comprise a random or nonsense nucleotide sequence and thus not share sequence identity to a target polynucleotide. In other aspects, the loop region comprises a sense or an antisense RNA sequence or fragment thereof that shares identity to a target polynucleotide. See, for example, International Patent Publication No. WO 02/00904, herein incorporated by reference. In certain aspects, the loop sequence can include an intron sequence, a sequence derived from an intron sequence, a sequence homologous to an intron sequence, or a modified intron sequence. The intron sequence can be one found in the same or a different species from which segments 1 and 3 are derived. In certain aspects, the loop region can be optimized to be as short as possible while still providing enough intramolecular flexibility to allow the formation of the base-paired stem region. Accordingly, the loop sequence is generally less than 1000, 900, 800, 700, 600, 500, 400, 300, 200, 100, 50, 25, 20, 19, 18, 17, 16, 15, 10 nucleotides or less.

[00153] The "first" and the "third" segment of the hairpin RNA molecule comprise the base-paired stem of the hairpin structure. The first and the third segments are inverted repeats of one another and share sufficient complementarity to allow the formation of the base-paired stem region. In certain aspects, the first and the third segments are fully complementary to one another. Alternatively, the first and the third segment may be partially complementary to each other so long as they are capable of hybridizing to one another to form a base-paired stem region. The amount of complementarity between the first and the third segment can be calculated as a percentage of the entire segment. Thus, the first and the third segment of the hairpin RNA generally share at least 50%, 60%, 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, up to and including 100% complementarity.

[00154] The first and the third segment are at least about 1000, 500, 475, 450, 425, 400, 375, 350, 325, 300, 250, 225, 200, 175, 150, 125, 100, 75, 60, 50, 40, 30, 25, 22, 20, 19, 18, 17, 16, 15 or 10 nucleotides in length. In certain aspects, the length of the first and/or the third segment is about 10-100 nucleotides, about 10 to about 75 nucleotides, about 10 to about 50 nucleotides, about 10 to about 40 nucleotides, about 10 to about 35 nucleotides, about 10 to about 30 nucleotides, about 10 to about 25 nucleotides, about 10 to about 19 nucleotides, about 10 to about 20 nucleotides, about 19 to about 50 nucleotides, about 50 nucleotides to about 100 nucleotides, about 100 nucleotides to about 150 nucleotides, about 100 nucleotides to about 300 nucleotides, about 150 nucleotides to about 200 nucleotides, about 200 nucleotides to about 250 nucleotides, about 250 nucleotides to about 300

nucleotides, about 300 nucleotides to about 350 nucleotides, about 350 nucleotides to about 400 nucleotides, about 400 nucleotide to about 500 nucleotides, about 600 nt, about 700 nt, about 800 nt, about 900 nt, about 1000 nt, about 1100 nt, about 1200 nt, 1300 nt, 1400 nt, 1500 nt, 1600 nt, 1700 nt, 1800 nt, 1900 nt, 2000 nt or longer. In other aspects, the length of the first and/or the third segment comprises at least 10-19 nucleotides, 10-20 nucleotides; 19-35 nucleotides, 20-35 nucleotides; 30-45 nucleotides; 40-50 nucleotides; 50-100 nucleotides; 100-300 nucleotides; about 500 -700 nucleotides; about 700-900 nucleotides; about 900-1100 nucleotides; about 1300 -1500 nucleotides; about 1500 - 1700 nucleotides; about 1700 - 1900 nucleotides; about 1900 - 2100 nucleotides; about 2100 - 2300 nucleotides; or about 2300 - 2500 nucleotides. See, for example, International Publication No. WO 0200904.

[00155] The disclosed hairpin molecules or double-stranded RNA molecules may have more than one disclosed sequence or active fragments or variants, or complements thereof, found in the same portion of the RNA molecule. For example, in a chimeric hairpin structure, the first segment of a hairpin molecule comprises two polynucleotide sections, each with a different disclosed sequence. For example, reading from one terminus of the hairpin, the first segment is composed of sequences from two separate genes (A followed by B). This first segment is followed by the second segment, the loop portion of the hairpin. The loop segment is followed by the third segment, where the complementary strands of the sequences in the first segment are found (B* followed by A*) in forming the stem-loop, hairpin structure, the stem contains SeqA-A* at the distal end of the stem and SeqB-B* proximal to the loop region.

[00156] In certain aspects, the first and the third segment comprise at least 20 nucleotides having at least 85% complementary to the first segment. In still other aspects, the first and the third segments which form the stem-loop structure of the hairpin comprises 3' or 5' overhang regions having unpaired nucleotide residues.

[00157] In certain aspects, the sequences used in the first, the second, and/or the third segments comprise domains that are designed to have sufficient sequence identity to a target polynucleotide of interest and thereby have the ability to decrease the level of expression of the target polynucleotide. The specificity of the inhibitory RNA transcripts is therefore generally conferred by these domains of the silencing element. Thus, in some aspects, the first, second and/or third segment of the silencing element comprise a domain having at least 10, at least 15, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 30, at least 40, at least 50, at least 100, at least 200, at least 300, at least 500, at least 1000, or more than 1000 nucleotides that share sufficient sequence identity to the target

polynucleotide to allow for a decrease in expression levels of the target polynucleotide when expressed in an appropriate cell. In other aspects, the domain is between about 15 to 50 nucleotides, about 19-35 nucleotides, about 20-35 nucleotides, about 25-50 nucleotides, about 19 to 75 nucleotides, about 20 to 75 nucleotides, about 40-90 nucleotides about 15-100 nucleotides, 10-100 nucleotides, about 10 to about 75 nucleotides, about 10 to about 50 nucleotides, about 10 to about 40 nucleotides, about 10 to about 35 nucleotides, about 10 to about 30 nucleotides, about 10 to about 25 nucleotides, about 10 to about 20 nucleotides, about 10 to about 19 nucleotides, about 50 nucleotides to about 100 nucleotides, about 100 nucleotides to about 150 nucleotides, about 150 nucleotides to about 200 nucleotides, about 200 nucleotides to about 250 nucleotides, about 250 nucleotides to about 300 nucleotides, about 300 nucleotides to about 350 nucleotides, about 350 nucleotides to about 400 nucleotides, about 400 nucleotide to about 500 nucleotides or longer. In other aspects, the length of the first and/or the third segment comprises at least 10-20 nucleotides, at least 10-19 nucleotides, 20-35 nucleotides, 30-45 nucleotides, 40-50 nucleotides, 50-100 nucleotides, or about 100-300 nucleotides.

[00158] In certain aspects, the domain of the first, the second, and/or the third segment has 100% sequence identity to the target polynucleotide. In other aspects, the domain of the first, the second and/or the third segment having homology to the target polypeptide have at least 50%, 60%, 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or greater sequence identity to a region of the target polynucleotide. The sequence identity of the domains of the first, the second and/or the third segments to the target polynucleotide need only be sufficient to decrease expression of the target polynucleotide of interest. See, for example, Chuang and Meyerowitz (2000) *Proc. Natl. Acad. Sci.* USA 97:4985-4990; Stoutjesdijk *et al.* (2002) *Plant Physiol.* 129:1723-1731; Waterhouse and Helliwell (2003) *Nat. Rev. Genet.* 4:29-38; Pandolfini *et al. BMC Biotechnology* 3:7, and U.S. Patent Publication No. 20030175965; each of which is herein incorporated by reference. A transient assay for the efficiency of hpRNA constructs to silence gene expression in vivo has been described by Panstruga *et al.* (2003) *Mol. Biol. Rep.* 30:135-140, herein incorporated by reference.

[00159] The amount of complementarity shared between the first, second, and/or third segment and the target polynucleotide or the amount of complementarity shared between the first segment and the third segment (i.e., the stem of the hairpin structure) may vary depending on the organism in which gene expression is to be controlled. Some organisms or cell types may require exact pairing or 100% identity, while other organisms or cell types

may tolerate some mismatching. In some cells, for example, a single nucleotide mismatch in the targeting sequence abrogates the ability to suppress gene expression. In these cells, the disclosed suppression cassettes can be used to target the suppression of mutant genes, for example, oncogenes whose transcripts comprise point mutations and therefore they can be specifically targeted using the methods and compositions disclosed herein without altering the expression of the remaining wild-type allele. In other organisms, holistic sequence variability may be tolerated as long as some 22 nt region of the sequence is represented in 100% homology between target polynucleotide and the suppression cassette.

[00160] Any region of the target polynucleotide can be used to design the domain of the silencing element that shares sufficient sequence identity to allow expression of the hairpin transcript to decrease the level of the target polynucleotide. For instance, the domain can be designed to share sequence identity to the 5' untranslated region of the target polynucleotide(s), the 3' untranslated region of the target polynucleotide(s), exonic regions of the target polynucleotide(s), intronic regions of the target polynucleotide(s), and any combination thereof. In certain aspects, a domain of the silencing element shares sufficient homology to at least about 15, 16, 17, 18, 19, 20, 22, 25 or 30 consecutive nucleotides from about nucleotides 1-50, 25-75, 75-125, 50-100, 125-175, 175-225, 100-150, 150-200, 200-250, 225-275, 275-325, 250-300, 325-375, 375-425, 300-350, 350-400, 425-475, 400-450, 475-525, 450-500, 525-575, 575-625, 550-600, 625-675, 675-725, 600-650, 625-675, 675-725, 650-700, 725-825, 825-875, 750-800, 875-925, 925-975, 850-900, 925-975, 975-1025, 950-1000, 1000-1050, 1025-1075, 1075-1125, 1050-1100, 1125-1175, 1100-1200, 1175-1225, 1225-1275, 1200-1300, 1325-1375, 1375-1425, 1300-1400, 1425-1475, 1475-1525, 1400-1500, 1525-1575, 1575-1625, 1625-1675, 1675-1725, 1725-1775, 1775-1825, 1825-1875, 1875-1925, 1925-1975, 1975-2025, 2025-2075, 2075-2125, 2125-2175, 2175-2225, 1500-1600, 1600-1700, 1700-1800, 1800-1900, 1900-2000 of the target sequence. In some instances to optimize the siRNA sequences employed in the hairpin, the synthetic oligodeoxyribonucleotide/RNAse H method can be used to determine sites on the target mRNA that are in a conformation that is susceptible to RNA silencing. See, for example, Vickers *et al.* (2003) *J. Biol. Chem.* 278:7108-7118 and Yang *et al.* (2002) *Proc. Natl. Acad. Sci.* USA 99:9442-9447, herein incorporated by reference. These studies indicate that there is a significant correlation between the RNase-H-sensitive sites and sites that promote efficient siRNA-directed mRNA degradation.

[00161] The hairpin silencing element may also be designed such that the sense sequence or the antisense sequence does not correspond to a target polynucleotide. In this aspect, the

sense and antisense sequence flank a loop sequence that comprises a nucleotide sequence corresponding to all or part of the target polynucleotide. Thus, it is the loop region that determines the specificity of the RNA interference. See, for example, WO 02/00904, herein incorporated by reference.

[00162] In addition, transcriptional gene silencing (TGS) may be accomplished through use of a hairpin suppression element where the inverted repeat of the hairpin shares sequence identity with the promoter region of a target polynucleotide to be silenced. See, for example, Aufsatz *et al.* (2002) *PNAS 99* (Suppl. 4):16499-16506 and Mette *et al.* (2000) *EMBO J.* 19(19):5194-5201.

[00163] In other aspects, the silencing element can comprise a small RNA (sRNA). sRNAs can comprise both micro RNA (miRNA) and short-interfering RNA (siRNA) (Meister and Tuschl (2004) *Nature* 431:343-349 and Bonetta *et al.* (2004) *Nature Methods* 1:79-86). miRNAs are regulatory agents comprising about 19 to about 24 ribonucleotides in length which are highly efficient at inhibiting the expression of target polynucleotides. See, for example Javier *et al.* (2003) *Nature* 425: 257-263, herein incorporated by reference. For miRNA interference, the silencing element can be designed to express a dsRNA molecule that forms a hairpin structure or partially base-paired structure containing 19, 20, 21, 22, 23, 24 or 25 -nucleotide sequence that is complementary to the target polynucleotide of interest. The miRNA can be synthetically made, or transcribed as a longer RNA which is subsequently cleaved to produce the active miRNA. Specifically, the miRNA can comprise 19 nucleotides of the sequence having homology to a target polynucleotide in sense orientation and 19 nucleotides of a corresponding antisense sequence that is complementary to the sense sequence. The miRNA can be an "artificial miRNA" or "amiRNA" which comprises a miRNA sequence that is synthetically designed to silence a target sequence.

[00164] When expressing an miRNA the final (mature) miRNA is present in a duplex in a precursor backbone structure, the two strands being referred to as the miRNA (the strand that will eventually base pair with the target) and miRNA*(star sequence). It has been demonstrated that miRNAs can be transgenically expressed and target genes of interest efficiently silenced (Highly specific gene silencing by artificial microRNAs in Arabidopsis Schwab R, Ossowski S, Riester M, Warthmann N, Weigel D. *Plant Cell.* 2006 May; 18(5):1121-33. *Epub* 2006 Mar 10; and Expression of artificial microRNAs in transgenic Arabidopsis thaliana confers virus resistance. Niu QW, Lin SS, Reyes JL, Chen KC, Wu HW, Yeh SD, Chua NH. *Nat. Biotechnol.* 2006 Nov; 24(11):1420-8. *Epub* 2006 Oct 22. Erratum in: *Nat. Biotechnol.* 2007 Feb; 25(2):254.)

[00165] The silencing element for miRNA interference comprises a miRNA primary sequence. The miRNA primary sequence comprises a DNA sequence having the miRNA and star sequences separated by a loop as well as additional sequences flanking this region that are important for processing. When expressed as an RNA, the structure of the primary miRNA is such as to allow for the formation of a hairpin RNA structure that can be processed into a mature miRNA. In some aspects, the miRNA backbone comprises a genomic or cDNA miRNA precursor sequence, wherein said sequence comprises a native primary in which a heterologous (artificial) mature miRNA and star sequence are inserted.

[00166] As used herein, a "star sequence" is the sequence within a miRNA precursor backbone that is complementary to the miRNA and forms a duplex with the miRNA to form the stem structure of a hairpin RNA. In some aspects, the star sequence can comprise less than 100% complementarity to the miRNA sequence. Alternatively, the star sequence can comprise at least 99%, 98%, 97%, 96%, 95%, 90%, 85%, 80% or lower sequence complementarity to the miRNA sequence as long as the star sequence has sufficient complementarity to the miRNA sequence to form a double stranded structure. In still further aspects, the star sequence comprises a sequence having 1, 2, 3, 4, 5 or more mismatches with the miRNA sequence and still has sufficient complementarity to form a double stranded structure with the miRNA sequence resulting in production of miRNA and suppression of the target sequence.

[00167] The miRNA precursor backbones can be from any plant. In some aspects, the miRNA precursor backbone is from a monocot. In other aspects, the miRNA precursor backbone is from a dicot. In further aspects, the backbone is from maize or soybean. MicroRNA precursor backbones have been described previously. For example, US20090155910A1 (WO 2009/079532) discloses the following soybean miRNA precursor backbones: 156c, 159, 166b, 168c, 396b and 398b, and US20090155909A1 (WO 2009/079548) discloses the following maize miRNA precursor backbones: 159c, 164h, 168a, 169r, and 396h. Each of these references is incorporated by reference in their entirety.

[00168] Thus, the primary miRNA can be altered to allow for efficient insertion of heterologous miRNA and star sequences within the miRNA precursor backbone. In such instances, the miRNA segment and the star segment of the miRNA precursor backbone are replaced with the heterologous miRNA and the heterologous star sequences, designed to target any sequence of interest, using a PCR technique and cloned into an expression construct. It is recognized that there could be alterations to the position at which the artificial miRNA and star sequences are inserted into the backbone. Detailed methods for inserting the

miRNA and star sequence into the miRNA precursor backbone are described in, for example, US Patent Applications 20090155909A1 and US20090155910A1, herein incorporated by reference in their entirety.

[00169] When designing a miRNA sequence and star sequence, various design choices can be made. See, for example, Schwab R, *et al.* (2005) *Dev. Cell* 8: 517-27. In non-limiting aspects, the miRNA sequences disclosed herein can have a "U" at the 5'-end, a "C" or "G" at the 19th nucleotide position, and an "A" or "U" at the 10th nucleotide position. In other aspects, the miRNA design is such that the miRNA have a high free delta-G as calculated using the ZipFold algorithm (Markham, N. R. & Zuker, M. (2005) *Nucleic Acids Res.* 33: W577-W581.) Optionally, a one base pair change can be added within the 5' portion of the miRNA so that the sequence differs from the target sequence by one nucleotide.

[00170] The methods and compositions disclosed herein employ silencing elements that when transcribed "form" a dsRNA molecule. Accordingly, the heterologous polynucleotide being expressed need not form the dsRNA by itself, but can interact with other sequences in the plant cell or in the pest gut after ingestion to allow the formation of the dsRNA. For example, a chimeric polynucleotide that can selectively silence the target polynucleotide can be generated by expressing a chimeric construct comprising the target sequence for a miRNA or siRNA to a sequence corresponding to all or part of the gene or genes to be silenced. In this aspect, the dsRNA is "formed" when the target for the miRNA or siRNA interacts with the miRNA present in the cell. The resulting dsRNA can then reduce the level of expression of the gene or genes to be silenced. See, for example, US Application Publication 2007-0130653, entitled "Methods and Compositions for Gene Silencing", herein incorporated by reference. The construct can be designed to have a target for an endogenous miRNA or alternatively, a target for a heterologous and/or synthetic miRNA can be employed in the construct. If a heterologous and/or synthetic miRNA is employed, it can be introduced into the cell on the same nucleotide construct as the chimeric polynucleotide or on a separate construct. As discussed elsewhere herein, any method can be used to introduce the construct comprising the heterologous miRNA.

VIII. GENOME EDITING OF TARGET GENES USING CAS/CRISPR

[00171] In one aspect, one or more polynucleotides as set forth in SEQ ID NOs: 1 and 2, an expression construct comprising a sequence as set forth in SEQ ID NOs: 1 and 2, or an expression construct encoding the polypeptide as set forth in SEQ ID NO:3, and compositions comprising said sequences, can be edited or inserted by genome editing using a double-stranded break-inducing agent such as a CRISPR/Cas9 system.

[00172] CRISPR loci (Clustered Regularly Interspaced Short Palindromic Repeats) (also known as SPIDRs--SPacer Interspersed Direct Repeats) constitute a family of recently described DNA loci. CRISPR loci consist of short and highly conserved DNA repeats (typically 24 to 40 bp, repeated from 1 to 140 times-also referred to as CRISPR-repeats) which are partially palindromic. The repeated sequences (usually specific to a species) are interspaced by variable sequences of constant length (typically 20 to 58 by depending on the CRISPR locus (WO2007/025097 published March 1, 2007).

[00173] Cas endonuclease relates to a Cas protein encoded by a Cas gene, wherein said Cas protein is capable of introducing a double strand break into a DNA target sequence. The Cas endonuclease is guided by a guide polynucleotide to recognize and optionally introduce a double strand break at a specific target site into the genome of a cell (*See* U.S. Patent Application Publication No. 2015/0082478). The guide polynucleotide/Cas endonuclease system includes a complex of a Cas endonuclease and a guide polynucleotide that is capable of introducing a double strand break into a DNA target sequence. The Cas endonuclease unwinds the DNA duplex in close proximity of the genomic target site and cleaves both DNA strands upon recognition of a target sequence by a guide RNA if a correct protospacer-adjacent motif (PAM) is approximately oriented at the 3' end of the target sequence.

[00174] In one aspect, the methods comprise creating a plant, wherein the target gene is edited so that it is no longer function, thereby creating a null mutant. The polynucleotide sequence of the target gene can be used to knockout the target gene polynucleotide in a plant by means known to those skilled in the art, including, but not limited to use of a Cas9/CRISPR system, TALENs, homologous recombination, and viral transformation. *See* Ma et al (2014), *Scientific Reports*, 4: 4489; Daimon et al (2013), *Development, Growth, and Differentiation*, 56(1): 14-25; and Eggleston et al (2001) *BMC Genetics*, 2:11.

IX. ALTERED EXPRESSION OF TARGET GENES USING DOUBLE-STRAND-BREAK TECHNOLOGIES

[00175] Sequence changes also can be introduced at specific selected sites using double-strand-break technologies such as ZNFs, custom designed homing endonucleases, TALENs, CRISPR/CAS (also referred to as guide RNA/Cas endonuclease systems (US patent application 14/463687 filed on August 20, 2014), or other protein and/or nucleic acid based mutagenesis technologies. The resultant variants can be screened for altered activity. It will be appreciated that these techniques are often not mutually exclusive. Indeed, the various methods can be used singly or in combination, in parallel or in series, to create or access diverse sequence variants.

[00176] A method is provided of making a plant, wherein the method comprises: (a) introducing into a plant cell a double-stranded break-inducing agent and at least one heterologous regulatory element; (b) regenerating a plant from the plant cell of step (a); (c) selecting a plant from step (b) that comprises the endogenous polynucleotide operably linked to the at least one heterologous regulatory element, wherein said endogenous polynucleotide encodes a polypeptide having an amino acid sequence of at least 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity, when compared to SEQ ID NO: 3; and (d) selecting the plant of step (c) that exhibits at least one trait selected from the group consisting of: increased root hair formation and growth, increased drought tolerance, and enhanced nutrient uptake, when compared to a control plant not comprising the at least one heterologous regulatory element. The at least one heterologous regulatory element may be selected from the group comprising: a promoter, an intron, an enhancer and a multimer of identical or different enhancers. The at least one heterologous regulatory element may comprise one, two, three or four copies of the CaMV 35S enhancer. The polypeptide may be over-expressed in at least one tissue of the plant (e.g., a root tissue), or during at least one condition of abiotic stress, or both. The plant may be selected from the group consisting of: maize, soybean, sunflower, sorghum, canola, wheat, alfalfa, cotton, rice, barley, millet, sugar cane and switchgrass.

X. QTL CHROMOSOME INTERVALS

[00177] In some aspects, the invention provides QTL chromosome intervals, where a QTL (or multiple QTLs) that segregate with at least one agronomic characteristic selected from the group consisting of increased root hair formation, increased root growth, increased drought tolerance, enhanced nutrient uptake, increased yield, and combinations thereof, are contained in those intervals. A variety of methods well known in the art are available for identifying chromosome intervals, including those methods described herein. The boundaries of such chromosome intervals are drawn to encompass markers that will be linked to one or more QTL. In other words, the chromosome interval is drawn such that any marker that lies within that interval (including the terminal markers that define the boundaries of the interval) can be used as markers for disease tolerance. Each interval comprises at least one QTL, and furthermore, may indeed comprise more than one QTL. Close proximity of multiple QTL in the same interval may obfuscate the correlation of a particular marker with a particular QTL, as one marker may demonstrate linkage to more than one QTL. Conversely, e.g., if two markers in close proximity show co-segregation with the desired phenotypic trait, it is sometimes unclear if each of those markers identifying the same QTL or two different QTL.

Regardless, knowledge of how many QTL are in a particular interval is not necessary to make or practice the invention.

[00178] The present invention provides soybean chromosome intervals, where the markers within that interval demonstrate co-segregation with at least one agronomic characteristic selected from the group consisting of increased root hair formation, increased root growth, increased drought tolerance, enhanced nutrient uptake, increased yield, and combinations thereof, in the chromosomal interval at 104,616,335 - 104,622,459 of chromosome 1 (SEQ ID NO:1), or comprising a sequence complementary to a sequence therein.

XI. TECHNIQUES FOR MARKER DETECTION

[00179] The invention provides molecular markers that have a significant probability of co-segregation with QTL that co-segregate with at least one agronomic characteristic selected from the group consisting of increased root hair formation, increased root growth, increased drought tolerance, enhanced nutrient uptake, increased yield, and combinations thereof,. These QTL markers find use in marker assisted selection for desired traits (tolerance or improved tolerance), and also have other uses. It is not intended that the invention be limited to any particular method for the detection of these markers.

[00180] Markers corresponding to genetic polymorphisms between members of a population can be detected by numerous methods well-established in the art (e.g., PCR-based sequence specific amplification, restriction fragment length polymorphisms (RFLPs), isozyme markers, allele specific hybridization (ASH), amplified variable sequences of the plant genome, self-sustained sequence replication, simple sequence repeat (SSR), single nucleotide polymorphism (SNP), random amplified polymorphic DNA ("RAPD") or amplified fragment length polymorphisms (AFLP)). In one additional embodiment, the presence or absence of a molecular marker is determined simply through nucleotide sequencing of the polymorphic marker region. This method is readily adapted to high throughput analysis as are the other methods noted above, e.g., using available high throughput sequencing methods such as sequencing by hybridization.

[00181] In general, the majority of genetic markers rely on one or more property of nucleic acids for their detection. For example, some techniques for detecting genetic markers utilize hybridization of a probe nucleic acid to nucleic acids corresponding to the genetic marker (e.g., amplified nucleic acids produced using genomic soybean DNA as a template). Hybridization formats, including but not limited to solution phase, solid phase, mixed phase, or in situ hybridization assays are useful for allele detection. An extensive guide to the hybridization of nucleic acids is found in Tijssen (1993) Laboratory Techniques in

Biochemistry and Molecular Biology—Hybridization with Nucleic Acid Probes Elsevier, New York; Berger and Kimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology volume 152 Academic Press, Inc., San Diego, Calif. ("Berger"); as well as in Sambrook and Ausubel (herein).

[00182] For example, markers that comprise restriction fragment length polymorphisms (RFLP) are detected, e.g., by hybridizing a probe which is typically a sub-fragment (or a synthetic oligonucleotide corresponding to a sub-fragment) of the nucleic acid to be detected to restriction digested genomic DNA. The restriction enzyme is selected to provide restriction fragments of at least two alternative (or polymorphic) lengths in different individuals or populations. Determining one or more restriction enzyme that produces informative fragments for each cross is a simple procedure, well known in the art. After separation by length in an appropriate matrix (e.g., agarose or polyacrylamide) and transfer to a membrane (e.g., nitrocellulose, nylon, etc.), the labeled probe is hybridized under conditions which result in equilibrium binding of the probe to the target followed by removal of excess probe by washing.

[00183] Nucleic acid probes to the marker loci can be cloned and/or synthesized. Any suitable label can be used with a probe of the invention. Detectable labels suitable for use with nucleic acid probes include, for example, any composition detectable by spectroscopic, radioisotopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means. Useful labels include biotin for staining with labeled streptavidin conjugate, magnetic beads, fluorescent dyes, radiolabels, enzymes, and colorimetric labels. Other labels include ligands which bind to antibodies labeled with fluorophores, chemiluminescent agents, and enzymes. A probe can also constitute radiolabelled PCR primers that are used to generate a radiolabelled amplicon. Labeling strategies for labeling nucleic acids and corresponding detection strategies can be found, e.g., in Haugland (1996) Handbook of Fluorescent Probes and Research Chemicals Sixth Edition by Molecular Probes, Inc. (Eugene Oreg.); or Haugland (2001) Handbook of Fluorescent Probes and Research Chemicals Eighth Edition by Molecular Probes, Inc. (Eugene Oreg.) (Available on CD ROM).

[00184] PCR, RT-PCR and LCR are in particularly broad use as amplification and amplification-detection methods for amplifying nucleic acids of interest (e.g., those comprising marker loci), facilitating detection of the markers. Details regarding the use of these and other amplification methods can be found in any of a variety of standard texts, including, e.g., Sambrook, Ausubel, Berger and Croy, herein. Many available biology texts also have extended discussions regarding PCR and related amplification methods. One of

skill will appreciate that essentially any RNA can be converted into a double stranded DNA suitable for restriction digestion, PCR expansion and sequencing using reverse transcriptase and a polymerase ("Reverse Transcription-PCR, or "RT-PCR"). See also Ausubel, Sambrook and Berger, above.

[00185] In one aspect, real time PCR or LCR is performed on the amplification mixtures described herein, e.g., using molecular beacons or TaqMan™ probes. A molecular beacon (MB) is an oligonucleotide or PNA which, under appropriate hybridization conditions, self-hybridizes to form a stem and loop structure. The MB has a label and a quencher at the termini of the oligonucleotide or PNA; thus, under conditions that permit intra-molecular hybridization, the label is typically quenched (or at least altered in its fluorescence) by the quencher. Under conditions where the MB does not display intra-molecular hybridization (e.g., when bound to a target nucleic acid, e.g., to a region of an amplicon during amplification), the MB label is unquenched. Details regarding standard methods of making and using MBs are well established in the literature and MBs are available from a number of commercial reagent sources. See also, e.g., Leone, *et al.*, (1995) "Molecular beacon probes combined with amplification by NASBA enable homogenous real-time detection of RNA" *Nucleic Acids Res.* 26:2150-2155; Tyagi and Kramer, (1996) "Molecular beacons: probes that fluoresce upon hybridization" *Nature Biotechnology* 14:303-308; Blok and Kramer, (1997) "Amplifiable hybridization probes containing a molecular switch" *Mol. Cell Probes* 11:187-194; Hsuih, *et al.*, (1997) "Novel, ligation-dependent PCR assay for detection of hepatitis C in serum" *J. Clin. Microbiol.* 34:501-507; Kostrikis, *et al.*, (1998) "Molecular beacons: spectral genotyping of human alleles" *Science* 279:1228-1229; Sokol, *et al.*, (1998) "Real time detection of DNA:RNA hybridization in living cells" *Proc. Natl. Acad. Sci.* USA 95:11538-11543; Tyagi, *et al.*, (1998) "Multicolor molecular beacons for allele discrimination" *Nature Biotechnology* 16:49-53; Bonnet, *et al.*, (1999) "Thermodynamic basis of the chemical specificity of structured DNA probes" *Proc. Natl. Acad. Sci.* USA 96:6171-6176; Fang, *et al.*, (1999) "Designing a novel molecular beacon for surface-immobilized DNA hybridization studies" *J. Am. Chem. Soc.* 121:2921-2922; Marras, *et al.*, (1999) "Multiplex detection of single-nucleotide variation using molecular beacons" Genet *Anal. Biomol. Eng.* 14:151-156; and Vet, *et al.*, (1999) "Multiplex detection of four pathogenic retroviruses using molecular beacons" *Proc. Natl. Acad. Sci.* USA 96:6394-6399. Additional details regarding MB construction and use is found in the patent literature, e.g., U.S. Pat. No. 5,925,517 (Jul. 20, 1999) to Tyagi, *et al.*, entitled "Detectably labeled dual conformation oligonucleotide probes, assays and kits;" U.S. Pat. No. 6,150,097 (November

21, 2000) to Tyagi, *et al.*, entitled "Nucleic acid detection probes having non-FRET fluorescence quenching and kits and assays including such probes" and U.S. Pat. No. 6,037,130 (Mar. 14, 2000) to Tyagi, *et al.*, entitled "Wavelength-shifting probes and primers and their use in assays and kits."

[00186] PCR detection and quantification using dual-labeled fluorogenic oligonucleotide probes, commonly referred to as "TagMan™" probes, can also be performed according to the present invention. These probes are composed of short (e.g., 20-25 base) oligodeoxynucleotides that are labeled with two different fluorescent dyes. On the 5′ terminus of each probe is a reporter dye, and on the 3′ terminus of each probe a quenching dye is found. The oligonucleotide probe sequence is complementary to an internal target sequence present in a PCR amplicon. When the probe is intact, energy transfer occurs between the two fluorophores and emission from the reporter is quenched by the quencher by FRET. During the extension phase of PCR, the probe is cleaved by 5′ nuclease activity of the polymerase used in the reaction, thereby releasing the reporter from the oligonucleotide-quencher and producing an increase in reporter emission intensity. Accordingly, TagMan™ probes are oligonucleotides that have a label and a quencher, where the label is released during amplification by the exonuclease action of the polymerase used in amplification. This provides a real time measure of amplification during synthesis. A variety of TagMan™ reagents are commercially available, e.g., from Applied Biosystems (Division Headquarters in Foster City, Calif.) as well as from a variety of specialty vendors such as Biosearch Technologies (e.g., black hole quencher probes).

[00187] Amplified variable sequences refer to amplified sequences of the plant genome which exhibit high nucleic acid residue variability between members of the same species. All organisms have variable genomic sequences and each organism (with the exception of a clone) has a different set of variable sequences. Once identified, the presence of specific variable sequence can be used to predict phenotypic traits. Preferably, DNA from the plant serves as a template for amplification with primers that flank a variable sequence of DNA. The variable sequence is amplified and then sequenced.

[00188] Alternatively, self-sustained sequence replication can be used to identify genetic markers. Self-sustained sequence replication refers to a method of nucleic acid amplification using target nucleic acid sequences which are replicated exponentially in vitro under substantially isothermal conditions by using three enzymatic activities involved in retroviral replication: (1) reverse transcriptase, (2) RNase H, and (3) a DNA-dependent RNA polymerase (Guatelli, *et al.*, (1990) *Proc. Natl. Acad. Sci.* USA 87:1874). By mimicking the

retroviral strategy of RNA replication by means of cDNA intermediates, this reaction accumulates cDNA and RNA copies of the original target.

[00189] Amplified fragment length polymorphisms (AFLP) can also be used as genetic markers (Vos, *et al.*, (1995) *Nucleic Acids Res.* 23:4407). The phrase "amplified fragment length polymorphism" refers to selected restriction fragments which are amplified before or after cleavage by a restriction endonuclease. The amplification step allows easier detection of specific restriction fragments. AFLP allows the detection large numbers of polymorphic markers and has been used for genetic mapping of plants (Becker, *et al.*, (1995) *Mol. Gen. Genet.* 249:65; and Meksem, *et al.*, (1995) *Mol. Gen. Genet.* 249:74).

[00190] Allele-specific hybridization (ASH) can be used to identify the genetic markers of the invention. ASH technology is based on the stable annealing of a short, single-stranded, oligonucleotide probe to a completely complementary single-strand target nucleic acid. Detection is via an isotopic or non-isotopic label attached to the probe.

[00191] For each polymorphism, two or more different ASH probes are designed to have identical DNA sequences except at the polymorphic nucleotides. Each probe will have exact homology with one allele sequence so that the range of probes can distinguish all the known alternative allele sequences. Each probe is hybridized to the target DNA. With appropriate probe design and hybridization conditions, a single-base mismatch between the probe and target DNA will prevent hybridization. In this manner, only one of the alternative probes will hybridize to a target sample that is homozygous or homogenous for an allele. Samples that are heterozygous or heterogeneous for two alleles will hybridize to both of two alternative probes.

[00192] ASH markers are used as dominant markers where the presence or absence of only one allele is determined from hybridization or lack of hybridization by only one probe. The alternative allele may be inferred from the lack of hybridization. ASH probe and target molecules are optionally RNA or DNA; the target molecules are any length of nucleotides beyond the sequence that is complementary to the probe; the probe is designed to hybridize with either strand of a DNA target; the probe ranges in size to conform to variously stringent hybridization conditions, etc.

[00193] PCR allows the target sequence for ASH to be amplified from low concentrations of nucleic acid in relatively small volumes. Otherwise, the target sequence from genomic DNA is digested with a restriction endonuclease and size separated by gel electrophoresis. Hybridizations typically occur with the target sequence bound to the surface of a membrane or, as described in U.S. Pat. No. 5,468,613, the ASH probe sequence may be bound to a

membrane.

[00194] In one embodiment, ASH data are typically obtained by amplifying nucleic acid fragments (amplicons) from genomic DNA using PCR, transferring the amplicon target DNA to a membrane in a dot-blot format, hybridizing a labeled oligonucleotide probe to the amplicon target, and observing the hybridization dots by autoradiography.

[00195] Single nucleotide polymorphisms (SNP) are markers that consist of a shared sequence differentiated on the basis of a single nucleotide. Typically, this distinction is detected by differential migration patterns of an amplicon comprising the SNP on e.g., an acrylamide gel. However, alternative modes of detection, such as hybridization, e.g., ASH, or RFLP analysis are also appropriate.

[00196] Isozyme markers can be employed as genetic markers, e.g., to track markers other than the tolerance markers herein, or to track isozyme markers linked to the markers herein. Isozymes are multiple forms of enzymes that differ from one another in their amino acid, and therefore their nucleic acid sequences. Some isozymes are multimeric enzymes containing slightly different subunits. Other isozymes are either multimeric or monomeric but have been cleaved from the proenzyme at different sites in the amino acid sequence. Isozymes can be characterized and analyzed at the protein level, or alternatively, isozymes which differ at the nucleic acid level can be determined. In such cases any of the nucleic acid based methods described herein can be used to analyze isozyme markers.

[00197] As herein, nucleic acid amplification techniques such as PCR and LCR are well known in the art and can be applied to the present invention to amplify and/or detect nucleic acids of interest, such as nucleic acids comprising marker loci. Examples of techniques sufficient to direct persons of skill through such in vitro methods, including the polymerase chain reaction (PCR), the ligase chain reaction (LCR), Qββ-replicase amplification and other RNA polymerase mediated techniques (e.g., NASBA), are found in the references noted above, e.g., Innis, Sambrook, Ausubel, Berger and Croy. Additional details are found in Mullis, *et al.*, (1987) U.S. Pat. No. 4,683,202; Arnheim and Levinson, (Oct. 1, 1990) C&EN 36-47; *The Journal Of NIH Research* (1991) 3:81-94; Kwoh, *et al.*, (1989) *Proc. Natl. Acad. Sci.* USA 86:1173; Guatelli, *et al.*, (1990) *Proc. Natl. Acad. Sci.* USA 87:1874; Lomeli, *et al.*, (1989) *J. Clin. Chem.* 35:1826; Landegren, *et al.*, (1988) *Science* 241:1077-1080; Van Brunt, (1990) *Biotechnology* 8:291-294; Wu and Wallace, (1989) *Gene* 4:560; Barringer, *et al.*, (1990) *Gene* 89:117, and Sooknanan and Malek, (1995) *Biotechnology* 13:563-564. Improved methods of amplifying large nucleic acids by PCR, which is useful in the context of positional cloning, are further summarized in Cheng, *et al.*, (1994) *Nature* 369:684, and

the references therein, in which PCR amplicons of up to 40 kb are generated.

[00198] In general, synthetic methods for making oligonucleotides, including probes, primers, molecular beacons, PNAs, LNAs (locked nucleic acids), etc., are well known. For example, oligonucleotides can be synthesized chemically according to the solid phase phosphoramidite triester method described by Beaucage and Caruthers, (1981) *Tetrahedron Letts* 22(20):1859-1862, e.g., using a commercially available automated synthesizer, e.g., as described in Needham-VanDevanter, *et al.*, (1984) *Nucleic Acids Res.* 12:6159-6168. Oligonucleotides, including modified oligonucleotides can also be ordered from a variety of commercial sources known to persons of skill. There are many commercial providers of oligo synthesis services, and thus this is a broadly accessible technology. Any nucleic acid can be custom ordered from any of a variety of commercial sources, such as The Midland Certified Reagent Company, The Great American Gene Company, ExpressGen Inc., Operon Technologies Inc. (Alameda, Calif.) and many others. Similarly, PNAs can be custom ordered from any of a variety of sources, such as PeptidoGenic, HTI Bio-Products, Inc., BMA Biomedicals Ltd. (U.K.), Bio•Synthesis, Inc., and many others.

[00199] In alternative embodiments, in silico methods can be used to detect the marker loci of interest. For example, the sequence of a nucleic acid comprising the marker locus of interest can be stored in a computer. The desired marker locus sequence or its homolog can be identified using an appropriate nucleic acid search algorithm as provided by, for example, in such readily available programs as BLAST, or even simple word processors.

[00200] In some preferred embodiments, the molecular markers of the invention are detected using a suitable PCR-based detection method, where the size or sequence of the PCR amplicon is indicative of the absence or presence of the marker (e.g., a particular marker allele). In these types of methods, PCR primers are hybridized to the conserved regions flanking the polymorphic marker region. As used in the art, PCR primers used to amplify a molecular marker are sometimes termed “PCR markers” or simply “markers”.

[00201] It will be appreciated that, although many specific examples of primers are provided herein (see, FIG. 2), suitable primers to be used with the invention can be designed using any suitable method. It is not intended that the invention be limited to any particular primer or primer pair. For example, primers can be designed using any suitable software program, such as LASERGENE®.

[00202] In some embodiments, the primers of the invention are radiolabelled, or labeled by any suitable means (e.g., using a non-radioactive fluorescent tag), to allow for rapid visualization of the different size amplicons following an amplification reaction without any

additional labeling step or visualization step. In some embodiments, the primers are not labeled, and the amplicons are visualized following their size resolution, e.g., following agarose gel electrophoresis. In some embodiments, ethidium bromide staining of the PCR amplicons following size resolution allows visualization of the different size amplicons.

[00203] It is not intended that the primers of the invention be limited to generating an amplicon of any particular size. For example, the primers used to amplify the marker loci and alleles herein are not limited to amplifying the entire region of the relevant locus. In some embodiments, marker amplification produces an amplicon at least 20 nucleotides in length, or alternatively, at least 50 nucleotides in length, or alternatively, at least 100 nucleotides in length, or alternatively, at least 200 nucleotides in length.

[00204] The following examples are offered by way of illustration and not by way of limitation.

EXAMPLES

[00205] Efforts have been made to ensure accuracy with respect to numbers (*e.g.*, amounts, temperature, etc.), but some errors and deviations should be accounted for. Unless indicated otherwise, temperature is in °C or is at ambient temperature, and pressure is at or near atmospheric.

EXAMPLE 1: ISOLATION OF THE REFERENCE ALLELE *RTH6-1*

[00206] The mutant *rth6* was initially identified by screening segregating F_2-families derived from *Mu* active lines of the collection at Pioneer Hi-Bred International, Inc. (Johnston, Iowa, US). The reference allele was designated *rth6-1* (Proprietary database, Schnable Laboratory Accession Number: *1351*) and backcrossed into the inbred line B73 >8 times.

EXAMPLE 2: CSEM IMAGING

[00207] Wild-type and *rth6* mutant seeds for cSEM (cryo scanning electron microscopy) were germinated for 3-4 days. Seedling roots were cut into 1-2 cm pieces with a razor blade and mounted on a specimen holder with a mixture of Tissue-Tek® O.C.T.™Compound (Sakura Finetek Europe B.V., Alphen aan den Rijn, The Netherlands) and colloidal graphite (Agar Scientific Ltd., Stansted, United Kingdom) and were immediately frozen in a nitrogen slush. Specimens were then transferred into a Quorum PP3010T cryo preparation chamber (Quorum Technologies Ltd., Laughton, United Kingdom) at -140 °C and water was sublimated at -80 °C for 40 min. Subsequently, specimens were platinum sputtered at 10 mA for 60 sec and imaged in a Zeiss SIGMA VP cryo scanning electron microscope (Carl Zeiss

AG, Oberkochen, Germany). Images were taken at a magnification of 100x, 1,600x and 16,000x using the SE2 detector with electron emission at 3 kV.

EXAMPLE 3: SEQUENCE WALKING

[00208] A Seq-Walking library was generated from genomic DNA extracted from 24 homozygous *rth6* mutant seedlings. The isolated DNA was sheared using a BioRuptor-UCD-200 sonication system (Diagenode, Inc., Denville, New Jersey, US) with 15 s/ 30 s on/ off cycles at low speed. The library was prepared as described previously (Li *et al.*, 2013) and sequenced on an Ion Proton system for next generation sequencing (Life Technologies, Grand Island, New York, US). As a control, Seq-Walking was also conducted on DNA isolated from the inbred line B73. The Seq-Walking library was prepared using the itp-Mu8 (GTCGAT) barcode, while the itp-Mu31 (CTGCTA) barcode was used for B73 library generation. Other primers used for Seq-Walking library preparation are listed in Li *et al.*, (2013) (see Li L, *et al.* (2013) *PLoS ONE*, 8(12): e82333).

EXAMPLE 4: BSR-SEQUENCE MAPPING

[00209] A line carrying the *rth6-1* allele backcrossed for five generations into the inbred line B73 was crossed with the inbred line Mo17 and subsequently self-pollinated to generate multiple F_2-families: [*Mo17/Mo17 x rth6-1/B73*]@. Kernels from an F_2-family segregating for the *rth6-1* allele were rolled up in water-soaked germination paper (Anchor Paper Company, Roseville, Minnesota, US) and placed in a 10 liter bucket filled with ca. 2 liters of distilled water. After 6 days in the incubator at 25 °C under constant light, 3 cm primary roots were harvested from both the *rth6* mutants and their wild-type siblings. In total, 123 mutant and 123 wild-type individuals were collected and separately pooled according to their phenotype. RNA was extracted separately from each pool with the RNeasy Mini Kit (QIAGEN GmbH, Hilden, Germany) and subsequently subjected to a DNaseI treatment as per the manufacturer's protocol. RNA quality (RIN >8) was checked on a 2100 Bioanalyzer (Agilent Technologies, Inc., Santa Clara, California, US) using a RNA 6000 NanoChip. RNA-Seq libraries were constructed using the Illumina RNA-Seq sample preparation kit according to the manufacturer's protocol. Two different Illumina TruSeq barcodes were used for the two libraries (Li et al., 2013; *ibid*). The libraries were pooled and sequenced on an Illumina HiSeq2000 (Illumina, Inc., San Diego, California, US), generating 99 bp single-end reads, which were analyzed as previously described (Li et al., 2013; *ibid*). Briefly, the two alleles of a given SNP site should be detected in approximately equal numbers of RNA-Seq reads when considering both pools of RNA-Seq reads. In contrast, only one allele of a SNP that is completely linked to the causal gene should be present in the RNA-Seq reads from the

mutant pool. Hence, linkage probabilities of each SNP with the causal gene were calculated and plotted versus the physical position of the gene. Subsequently, chromosome 1 was scanned by using a chromosome window containing 100 SNPs and the median linkage probability was plotted versus the middle physical position of the window. The window was slid with a step size of 20 SNPs. Sequence reads from the BSR-Seq experiment have been deposited in the SRA ("sequence read archive"; available via the National Center for Biotechnology Information website) under AC: SRP044758.

EXAMPLE 5: TRANSPOSON INSERTION ALLELES

[00210] Confirmation of the rth6 candidate gene was carried out by independent transposon insertion allele methods. Briefly, novel *Mutator* insertions in the *rth6* gene to independently confirm the *rth6* candidate gene were generated by a direct transposon tagging experiment in which plants homozygous for the *rth6-1* allele were crossed as males with *Mu*-active stocks. Three novel alleles *rth6-2* (*rth6-Mu 10B-668*), *rth6-3* (*rth6-Mu 11B-451*) and *rth6-4* (*rth6-Mu 11B-453*) were identified in a forward genetic screen for roothairless phenotypes among ~94,000 seeds of the resulting progeny. Mutant progeny were self-pollinated and in parallel crossed with plants homozygous for *rth6-1* to validate the new mutant alleles. Moreover, two *Ac/Ds* insertion alleles *rth6-5* (Plant GDB Accession No.: *Ac*.mon00102; Schnable Lab Accession No.: 5972) and *rth6-6* (Plant GDB Accession No.: *Ds* I.S07.1244A; Schnable Lab Accession No.: 5971) were obtained by reverse genetics from Plant GDB website. All *Mutator* and *Ac/Ds* insertion sites were mapped by PCR and subsequent sequencing using the transposon-specific oligonucleotide primers in combination with the *rth6* specific oligonucleotide primers listed in Table 1 below.

Table 1.

Name	Seq ID No.	Sequence
IDP6931-Fw	18	5´ TTGCATGTAGTGGGAAGTGC 3´
IDP6931-Rv	19	5´ TTAAATGGGCGAACATAGCC 3´
IDP525-Fw	20	5´ TATGGGAGTGAAAGCCATCC 3´
IDP525-Rv	21	5´ GCGTGCTGAGACTTACCTGG 3´
MuTIR	22	5´ AGAGAAGCCAACGCCA(AT)CGCCTC(CT) ATTTCGTC 3´
rth6-2 (rth6C1R4)	23	5´ GCATGTTGACGTCGAAGAAA 3´
rth6-4 (rth6C1R2)	24	5´ ACTCCAGGGAGATCCGACTT 3´
Ac/Ds specific (JSR05)	25	5´ CGTCCCGCAAGTTAAATATGA 3´

Name	Seq ID No.	Sequence
rth6-5 (rth6C1R1)	26	5´ CTCCCACTCGGTGTTCTTGT 3´
rth6-6 (I.S07.1244A-R)	27	5´ CCCACTCGGTGTTCTTGTAAGG 3´
ZmCslD1-Fw	28	5´ CGGTGGAGCTTCTTGATG 3´
ZmCslD1-Rev	29	5´ AGGAGGGCTGATGTAGAC 3´
ZmCslD2-Fw	30	5´ CAGCCGCACCATCTACAG 3´
ZmCslD2-Rev	31	5´ GTTGATGGCCACCCACAG 3´
ZmCslD3-Fw	32	5´ CCGTTGGCAGTGATCGGTA 3´
ZmCslD3-Rev	33	5´ CCAGCACCCAGAAGCTGAA 3´
ZmCslD4-Fw	34	5´ TCGTTTGGTCGGGACTCATC 3´
ZmCslD4-Rev	35	5´ AAGCTAGGCGGTCGTTGTC 3´
ZmCslD5 (*rth*6)-Fw	36	5´ CTGGGTGCTTGCTCATCTGT 3´
ZmCslD5 (*rth*6)-Rev	37	5´ GTTGGCAGCTTGTGATGGAG 3´
Myosin-Fw	38	5´ AGAAGGCCGTACAGGATCTTACC 3´
Myosin-Rev	39	5´ CAAGGAGAGACTCTGTGAGCTTCA 3´

EXAMPLE 6: QRT-PCR ASSAY

[00211] For quantitative real time PCR, seedlings of the maize inbred line B73 plants were grown for three to ten days in germination paper rolls (Anchor Paper Company) as previously described (Hetz W., *et al.* (1996) *Plant J.* 10: 845-857) under a 16 h light/ 8 h dark photoperiod at 28 °C and 24 °C, respectively. Subsequently, roots of ten plants were collected for each of the four biological replicates per tissue or developmental stage and immediately frozen in liquid nitrogen. Furthermore, four different tissues of young primary roots were collected. The meristematic zone comprising the first two mm of the root tip and the proximal elongation zone were cut with a razor blade under a stereo microscope (Zeiss Stemi 2000, Carl Zeiss AG). The differentiation zone which was distinguished from the neighboring elongation zone by the presence of root hairs was separated into cortical parenchyma and stele tissues as previously described (Saleem M., *et al.* (2009): *J. Proteome Res.* 8: 2285-2297). For each root hair sample, 100 - 150 primary roots of three-day-old seedlings were dipped into liquid nitrogen. Subsequently, frozen root hairs were broken off the primary root using a pre-cooled spatula and collected in a mortar where they were pulverized in liquid nitrogen with a pestle. Total RNA was extracted with the RNeasy Kit (QIAGEN GmbH).

RNA integrity was measured using an Agilent 2100 Bioanalyzer and a RNA 6000 Nano chip (Agilent Technologies Inc., Santa Clara, USA). RIN values for all samples were >9.5. cDNA was synthesized from 1 µg of total RNA using the Quanta qScript™ cDNA SuperMix (Quanta BioSciences, Inc., Gaithersburg, Maryland, US). The cDNA was 1:2 diluted with water and a dilution series was prepared up to 1:128. Each biological replicate was measured in a BioRad CFX 384 Real-Time System (Bio-Rad Laboratories, Inc., Hercules, California, US) in three technical replicates using the Quanta PerfeCTa® SYBR® Green SuperMix (Quanta BioSciences, Inc.). Primer efficiencies were calculated using the following formula: PCR amplification efficiency= $10^{-1/slope}$ – 1 (Bustin SA, *et al. Clin Chem.* 2009 Apr;55(4):611-22). Primer efficiencies were between 85% and 105% and R^2 was >0.995. Expression levels were calculated relative to a homolog of a *myosin* heavy-chain gene (GenBank Accession No.: AI941656) previously used as a reference for expression in maize roots (Hoecker N., *et al.* (2008) *Genetics* 179: 1275-83.).

EXAMPLE 7: PHYLOGENETIC ANALYSIS

[00212] The predicted amino acid sequence of RTH6 was compared (using tblastn) to the translated nucleotide databases of *C. reinhardii*, *P. patens*, *S. moellendorffii*, *A. thaliana*, *G.raimondii*, *M. domestica*, *P. trichocarpa*, *C. sativus*, *S. bicolor*, *B. distachyon*, *Z. mays*, and *O. sativa* from the Phytozome 10.1 plant genomics web portal (US Department of Energy). Phytozome is the Plant Comparative Genomics portal of the Department of Energy's Joint Genome Institute. Homologous sequences were downloaded and aligned using ClustalW with a gap opening penalty of 10 and gap extension penalty of 0.1 using MEGA 6 (Tamura *et al. Mol. Biol. Evol.* 2013 Dec;30(12):2725-9). Alignments were exported into a NEXUS file and trees were generated by MrBayes (Huelsenbeck and Ronquist (2001) *Bioinformatics* 17(8):754-5; Ronquist and Huelsenbeck (2003) *Bioinformatics* 19(12):1572-4). The tree was calculated using a Markov Chain Monte Carlo approach and three hot chains and 5 million generations until the standard deviation of split frequencies dropped below 0.01. The phylogenetic tree was built using FigTree software (available via the website of Andrew Rambaut, Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, United Kingdom).

EXAMPLE 8: *ROOTHAIRLESS 6 (RTH6)* GENE CONTROLS ROOT HAIR LENGTH

[00213] A mutant specifically affected in root hair elongation but otherwise normal was identified in a phenotypic screen of an F_2-population resulting from a *Mutator* transposon mutagenesis. Genetic crosses established that this mutant phenotype is controlled by a single recessive allele. Moreover, genetic crosses of homozygous mutants of this new mutant allele

with the previously isolated mutants *rth1 to rth5* resulted in complementation thus demonstrating that these mutants are not allelic. Therefore, this new mutant was designated *roothairless 6* (*rth6*).

[00214] Stereo microscopy demonstrated that *rth6* mutants display significantly shorter root hairs than their wild-type siblings in all major root-types including primary-, seminal- and crown roots (FIG. 1A). Quantification of root hair length established that on average *rth6* root hairs developed only 4-5% of the length of wild-type root hairs in primary, seminal and crown roots (FIG. 1B). Root hair morphology of wild-type and *rth6* were analyzed via cSEM (cryo scanning electron microscopy). High resolution cSEM pictures revealed that *rth6* forms only root hair bulges (see FIGs. 1F-1H) which did not elongate compared to wild-type root hairs (FIGs. 1C-1E). The basal bulges of *rth6* were characteristically swollen (FIG. 1G) like the basal region of wild-type root hairs (FIG. 1D). In contrast to wild-type root hairs, the bulges of *rth6* were arrested at this developmental stage and unable to elongate. High resolution pictures of root hair tips illustrated that wild-type root hairs had a smooth surface (FIG. 1E) while the surface of mutant *rth6* root hairs was rough as a consequence of defects during root hair formation (FIG. 1H). In addition, mutant *rth6* root hair tips (FIG. 1H) displayed small outgrowths with a very rough surface at the tip. Without wishing to be bound by a particular theory, it is possible that these are initiation sites which represent the switch from bulge formation to root hair elongation which is defective in *rth6*.

[00215] The maize mutants *rth1* (Wen *et al.*, (2005) *Plant Physiol.* 138: 1637-1643), *rth2* (Wen and Schnable (1994) *Am. J. Bot.* 81:833-843), *rth3* (Hochholdinger *et al.*, (2008) *Plant J.* 54: 888-98) and *rth5* (Nestler *et al.*, (2014) *Plant J.*, 79: 729-740) also display defects in root hair formation constitutively in all root types. In contrast, other aspects of maize root development such as lateral root formation are controlled by root-type specific genetic mechanisms. This is illustrated by the mutants *rum1* (Woll *et al.*, (2005) *Plant Physiol.* 139: 1255-1267; von Behrens *et al.*, (2011) *Plant J.*, 66: 341-353) and *lrt1* (Hochholdinger and Feix, (1998) *Plant J.* 16:247-255). In both mutants only embryonic primary and seminal roots are defective in lateral root initiation while postembryonic shoot-borne roots are unaffected by these mutations.

[00216] The maize roothairless mutants identified thus far are affected at different stages of the three step process of root hair development (Dolan *et al.*, (1994) *Development* 120:2465-2474). Both, the mutant *rth3* (Hochholdinger et al., (2008) *Plant J.* 54: 888-98) and the mutant *rth6*, as disclosed herein, can form a bulge but are unable to initiate root hair elongation. In contrast, the mutant *rth5* forms a bulge and initiates tip growth but ceases

elongation soon after (Nestler *et al.*, (2014) *Plant J.*, 79: 729-740). Finally, the mutants *rth1* (Wen *et al.*, (2005) *Plant Physiol.* 138: 1637-1643) and *rth2* (Wen and Schnable (1994) *Am. J. Bot.* 81:833-843) display the longest root hairs of all maize roothairless mutants. In both mutants root hairs initiate polar tip growth and elongate to a certain extent, but their overall final length lacks significantly behind their wild-type siblings. This diversity of root hair mutant phenotypes in maize illustrates the subtle multi-step genetic regulation of this process in this species.

[00217] To date three genes that control root hair elongation in maize have been cloned. The *roothairless1* (*rth1*) gene encodes a SEC3 subunit (Wen *et al.*, (2005) *Plant Physiol.* 138: 1637-1643) of the exocyst complex which tethers exocytotic vesicles prior to their fusion thus mediating exocytotic tip growth of root hairs (Hala *et al.*, (2008) *Plant Cell* 20: 1330-45). Moreover, the *roothairless3* (*rth3*) gene is translated into a monocot-specific COBRA like cell wall protein (Hochholdinger *et al.*, (2008) *Plant J.* 54: 888-98). Most recently, the *roothairless5* (*rth5*) gene was demonstrated to give rise to a monocot-specific NADPH oxidase, which is involved in the tip growth of root hairs (Nestler *et al.*, (2014) *Plant J.*, 79: 729-740) *Plant J.*, 79: 729-740). In maize the *roothairless5* (*rth5*) gene is yet the only gene that is known to control root hair density and thus the differentiation of epidermis cells into trichoblasts (Nestler *et al.*, (2014) *Plant J.*, 79: 729-740). However, *rth5* such as all other maize roothairless mutants (Wen and Schnable (1994) *Am. J. Bot.* 81:833-843) was initially identified based on its root hair elongation defect.

EXAMPLE 9: CLONING OF *RTH6*

[00218] The *rth6* gene was genetically mapped to the short arm of chromosome 1 by B-A translocation stocks. Subsequently, 123 mutant and 123 wild-type individuals of a F_2-mapping population segregating for *rth6* mutant phenotypes were subjected to BSR-Seq (see Example 4 for method). This analysis demonstrated that the causative gene maps to the centromeric region of chromosome 1 (FIG. 2A). Chromosome 1 was then scanned by using a window containing 100 SNPs and a step size of 20 SNPs. Within each window, the median linkage probability obtained from a Bayesian BSA analysis across all 100 SNPs was determined and plotted against the middle physical position of the window. As a result the *rth6* gene was mapped to a 15.7 cM interval flanked by IDP (insertion deletion polymorphism) markers IDP6931 (130.8 cM) and IDP525 (146.5 cM) of ISU IBM Integrated Map (Liu, S. (2009) *PLoS Genet.* 5(11): e1000733. doi:10.1371/journal.pgen.1000733; see FIG. 2B herein) which corresponded to ~50 Mb on the physical map of chromosome 1 (maize B73 reference genome sequence, Maize Genome AGP version 1, release 4a53;

available via ftp at the web portal for the Maize Genome Sequencing Project). Correlation of physical and genetic distances on chromosome 1 calculated via molecular markers from an IBM genetic map (Li et al., 2013) indicated that the centrometic region denoted by a vertical line exhibited low rates of recombination per Mb (FIG. 2C).

[00219] Genomic DNA extracted from the inbred line B73 and a pool of seedlings derived from a self of a plant with the genotype *rth6-1/rth6-2* was used to construct Seq-Walking libraries, which were then sequenced on an Ion Proton instrument (see Example 3 above). In total, 16,680,521 raw sequence reads were obtained, 7,175,209 from the *rth6* library and 9,505,312 from the B73 library. Reads were sorted and decoded according to their barcode sequences. Subsequently, barcodes, adapters and *Mu*-related sequences were removed. The remaining *Mu* flanking sequences that were longer than 90 bp (3,054,970 and 2,703,957, respectively) were aligned to the B73 reference genome.

[00220] *Mu* insertions recovered from the B73 Seq-Walking library were assumed not to be responsible for an *rth6* mutation. This subtractive process resulted in the identification of 29 *Mu* insertion sites within the *rth6* mapping interval (Chr1: from 102,196,700 bp to 144,956,100 bp) that had been defined by the BSR-Seq experiment (see Table 2 below). Table 2 shows the number of rth6-specific Seq-Walking reads recovered per Mu insertion site in the mapped interval on chromosome 1.

Table 2.

***Mu* sites (bp)**	**No. of reads**
102,196,700	3
102,580,900	5
102,626,800	1
103,419,600	21
104,610,900	22,945
105,325,800	10
105,765,800	3
106,045,500	14
111,201,400	1
111,201,500	11
112,896,100	2
113,445,900	1
115,219,500	31
116,723,900	1
119,214,500	1
124,691,700	44

Mu sites (bp)	No. of reads
126,516,200	3
126,775,900	1
129,262,600	2
131,559,500	10
133,526,300	1
138,592,600	4
138,614,400	30
138,952,700	37
140,962,500	43

Many more reads (N=22,945) were obtained for a single insertion site (Chr1: 104,610,900 bp) within the GRMZM2G436299 gene than for any of other of the 28 sites within the mapping interval. Confirmative PCR experiments on genomic DNA of the *rth6-1* and *rth6-2* alleles demonstrated that the observed *Mu* insertion derived from allele *rth6-2* whereas the reference allele *rth6-1* did not contain a *Mu* transposon insertion. Based on these results we declared GRMZM2G436299 an *rth6* candidate gene.

EXAMPLE 10: CONFIRMATION OF RTH6 IDENTITY

[00221] To confirm that GRMZM2G436299 represents indeed the *rth6* gene, independent mutant alleles were generated. Three novel *Mu* induced alleles were generated by direct transposon tagging and identified in a forward genetic screen (see methods). These alleles were designated *rth6-2*, *rth6-3* and *rth6-4*. Moreover, two *Ac/Ds* insertion alleles designated *rth6-5* and *rth6-6* were obtained in a reverse genetic approach from Plant GDB after the *rth6* candidate gene was identified (see Example 5 methods). All novel alleles displayed the roothairless phenotype and transposon insertion sites were mapped by PCR and subsequent sequencing. The results pertaining to the mutant alleles of *rth6* are summarized in Table 3 below. These experiments confirmed that GRMZM2G436299 is the *rth6* gene by four independent mutant alleles which contained transposon insertions at different positions of exon 1 (*rth6-4*, *rth6-5*, *rth6-6*) and exon 3 (*rth6-2*) of the candidate gene which all conditioned a roothairless phenotype. These results were further substantiated by crosses of the four transposon induced alleles to the reference allele *rth6-1*. All progeny displayed the roothairless phenotype demonstrating that these mutants are allelic.

[00222]

Table 3.

Allele name	Allele ID	Type of mutation	Mutation site in bp on chr. 1 (AGPv2)
rth6-1	*rth6-ref*	n. d.[a]	n. d.
rth6-2	*rth6-Mu 10B-668*	*Mu1*	104,610,915
rth6-3	*rth6-Mu 11B-451*	n. d.	n. d.
rth6-4	*rth6-Mu 11B-453*	*Mu1*	104,609,772
rth6-5	*rth6-Ac*.mon00102	*Ac*	104,609,104
rth6-6	*rth6-Ds* l.S07.1244A	*Ds*	104,609,449

[a] "n.d." indicates parameter was not determined.

EXAMPLE 11: GENE STRUCTURE AND FUNCTION OF *RTH6*

[00223] The *rth6* gene contains four exons and three introns encoding a 3,866 bp open reading frame which translates into an 1,159 aa protein (FIG. 2D) with a predicted molecular weight of 129 kD and an isoelectric point of 7.9 (based on the primary amino acid sequence; determined using the web-based tools at Swiss Institute of Bioinformatics ("SIB") Bioinformatics Resource Portal, ExPASy). RTH6 is predicted to be a membrane protein with 8 transmembrane domains that contains an N-terminal RING/Ubox like Zinc-finger domain, and a C-terminal cellulose synthase-like protein domain (using the web-based tool, HMMER, available at the web site of the European Bioinformatics Institute of the European Molecular Biology Laboratory; described in Finn, R.D., *et al.*, *Nucl. Acids Res.* (2011) Web Server Issue 39:W29-W37).

EXAMPLE 12: HIGHLY CONSERVED RTH6 HOMOLOGS IN PLANTS

[00224] Sequence similarity searches and phylogenetic analyses revealed that *rth6* encodes a Cellulose Synthase-like D (CSLD) protein which belongs to the D-type subfamily of the Cellulose Synthase superfamily. The CSLD proteins which are illustrated in FIG. 3 cluster into monocot, dicot, and non-seed-plant (see FIG. 3) subclades which are strictly separated. Each of the analyzed monocot species maize, rice, sorghum and brachypodium encodes five *CslD* genes. The five maize *CslD* genes have been previously designated *CslD1* to *CslD1* (Penning *et al.*, (2009) *Plant Physiology* 151: 1703-1728). The *rth6* genes correspond to

CslD5. Proteins encoded by monocot *CslD* genes form five groups in which each of the four species is represented by one protein. Typically, clusters of monocot CSLD proteins group with clusters of dicot CSLD proteins. Even evolutionary distantly related plant species such as the moss *Physcomitrella patens* or the clubmoss *Selaginella moellendorffii* encode for CSLD proteins with a surprisingly high degree of sequence identity of 67%-68% with RTH6 of maize. No homologs were identified in green algae such as *Chlamydomonas reinhardii*.

[00225] Without wishing to be bound by a particular theory, it is believed that *rth6* encodes CSLD5 of the Cellulose synthase like D (CSLD) protein family (Penning *et al.*, 2009, *ibid*). RTH6 is predicted to contain a C-terminal cellulose synthase domain and eight transmembrane domains. Based on shared sequence identity, all members of the cellulose synthase superfamily are predicted to be membrane bound processive glycosyltransferases that synthesize b-linked glycan polymers such as cellulose or hemicellulose backbones found in cell walls (Richmond and Sommerville, (2000) *Plant Physiol.* 124(2):495-498). Evidence that CSLD proteins indeed act as cellulose synthases was provided by the complementation of the Arabidopsis *csld3* mutant by a chimeric CSLD3 protein containing a CESA domain known to catalyze cellulose synthesis (Park, S., et al. (2011) *Nature Cell Biology* 13(8):973-980). The N-terminal RING/IUbox Zinc-finger-like domain predicted in RTH6 is specific for the CSLD and CESA subfamilies of Cellulose Synthases (Richmond and Sommerville, 2000, *ibid*). This domain was suggested to mediate protein-protein interactions (Kurek *et al.*, (2002) *Proc. Natl. Acad. Sci.* USA 99(17):11109-11114; Gamsjaeger *et al.*, (2007) *Trends Biochem. Sci.* 32: 63–70).

[00226] Several members of the CSLD gene family have been functionally characterized and it has been demonstrated that these genes are involved tip growing processes of specific cells such as pollen tubes (*AtCslD1* and *AtCslD4*, Bernal *et al.*, (2008) *Plant Physiol.* 148: 1238–1253) and root hairs (*AtCslD3/KJK*, Favery *et al.*, 2001; Wang *et al.*, 2001; *OsCslD1*, Kim *et al.*, (2007) *Genes Dev.* 15: 79–89; *AtCslD2*, Bernal *et al.*, 2008, *ibid*; *ZmCslD5*, Penning *et al.*, (2009) *Physiol.* 151: 1703–1728) but also the establishment of new cross walls during cell division (*ZmCSLD1*; Hunter *et al.*, (2012) *Plant Physiol.* 2012 Feb; 158(2): 708–724).

[00227] Despite the high degree of sequence conservation on the amino acid level, mutations in different members of the CSLD family result in morphologically different root hair phenotypes. For instance, the maize mutant *rth6* (*csld5*) and Arabidopsis *csld3/kjk* display very short root hairs which are arrested shortly after initiation and are unable to elongate after bulge formation. A potential role for CSLD3 during cell-wall synthesis in

apical plasma membranes of tip-growing root-hair cells has been previously described (Park, S., *et al.* (2011) *Nature Cell Biology* 13(8):973-980). In contrast, root hairs in the rice mutant *csld1* are initiated normally but their elongation is impaired relatively late in development and root hairs display kinks and swellings along their length (Kim *et al.*, (2007) *Plant Physiol.* 143: 1220–1230). Similarly, distinct root hair surface morphologies have been observed. While maize *rth6* mutants display a ruptured root hair surface these defects are even more pronounced in the Arabidopsis mutants *csld2* and *csld3/kjk* which often even results in the loss of material from inside the cell (Bernal *et al.*, (2008) *Plant Physiol.* 148: 1238–1253; Galway *et al.*, 2013). Hence, despite the fundamental structural differences in type I (Arabidopsis) and type II (maize and rice) cell wall compositions (Carpita and Gibeaut (1993) *Plant J.* 3: 1–30) these closely related CSLD proteins function in root hair formation although the morphology and surface structure of mutant root hairs in the studied species is different.

[00228] CSLD proteins belong to one of ten subfamilies of the highly conserved Cellulase Synthase superfamily (Penning *et al.*, (2009) *Plant Physiology* 151: 1703-1728). The high degree of evolutionary conservation of this subfamily is supported by the observation that early land plants such as the club moss *Selaginella* and the moss *Physcomitrella* already contain *CslD* genes (FIG. 3). Selaginella belongs to the oldest living vascular plant division Lycopodiophyta which emerged ~410 million years ago. Phylogenetic reconstruction revealed that monocot, dicot and non-seed-plant subclades are strictly separated which might illustrate that diversification of the gene family mainly occurred along the boundary of monocot, dicot and non-seed plant species (FIG. 3). Therefore, a one to one correlation of monocot and dicot CSLD proteins is difficult although it has been demonstrated that monocot and dicot members of closely related clades can have similar functions as illustrated above for the monocot genes *rth6* (*ZmCsld5*) and *OsCSLD1* (Kim *et al.*, (2007) *Plant Physiol.* 143: 1220–1230) and the dicot genes *AtCSLD2* (Bernal *et al.*, (2008) *Plant Physiol.* 148: 1238–1253) and *AtCSLD3/KJK* (Favery *et al.*, (2001) *Genes Dev.* 15: 79–89; Wang *et al.*, (2001) *Plant Physiol.* 126: 575–586) in root hair development.

[00229] Based on their low (*CslD2*) to absent (*CslD1*, *CslD3*, *CslD4*) expression in root hairs, the other four members of the maize Cellulose Synthase-like D subfamily have likely other functions than controlling root hair formation. To date except for *rth6* (*CslD5*), only *CslD1* has been characterized by a mutant which is defective in cell division and expansion and therefore displays a significantly reduced width of several organs including leaves (Hunter *et al.*, (2012) *Plant Physiol.* 158 (2):708-24).

EXAMPLE 13: ROOT-TYPE AND TISSUE-SPECIFIC EXPRESSION OF *RTH6*

[00230] Expression of *rth6* was surveyed in a wide range of tissues and root types by qRT-PCR (FIGs. 4A-4C). In primary roots of different length *rth6* displayed the highest expression in young primary roots of 1-2 and 2-4 cm while significantly less expression was detected in older primary roots of 4-8 and 10-14 cm length (FIG. 4A). Lateral roots displayed similarly high expression levels as young primary roots, whereas seminal and crown roots showed lower expression levels similar to older primary roots (FIG. 4A). Expression of *rth6* in leaves as an example for non-root tissues was significantly lower than in most root tissues (FIG. 4A).

[00231] Primary roots of 2-4 cm length were then dissected longitudinally into the meristematic zone (MZ), the elongation zone (EZ) and the differentiation zone. The differentiation zone was further dissected into cortex (C), stele (S), and root hairs (RH). Root hairs displayed significantly higher expression of *rth6* than all other tissues. Expression in elongation zone and cortex could be attributed to trichoblasts present in these tissues (FIG. 4B).

[00232] Root hair specific expression of *rth6* and the visible mutant phenotype implies that *rth6* has only limited functional redundancy with other members of the maize *CslD* family. Similar root hair specific expression and visible root hair defective mutant phenotypes have been observed for the rice *OsCslD1* (Kim *et al.*, (2007) *Plant Physiol.* 143: 1220–1230) and maize *rth5* which encodes a NADPH oxidase (Nestler *et al.*, (2014) *Plant J.*, 79: 729-740) also implying limited functional redundancy with other members of these gene families. For the Arabidopsis AtCSLD2 and AtCSLD3 proteins partial redundancy has been observed because double mutants of these genes display even shorter root hairs than the single mutants (Bernal *et al.*, (2008) *Plant Physiol.* 148: 1238–1253).

[00233] Finally, *rth6* expression was monitored in primary roots of 2-4 cm length in wild-type and mutant *rth1*, *rth2*, *rth3*, *rth5*, and *rth6* seedlings (FIG. 4C). Expression of *rth6* in mutant primary roots was on average repressed to 6% of the expression levels in wild-type primary roots. Moreover, *rth6* expression was repressed in roots of the mutants *rth3* and *rth5* and significantly enhanced in the mutant *rth2*. Expression of *rth6* was unaffected in the mutant *rth1*. Among the five closely related *CslD* genes of maize only *CslD1*, *CslD2* and *CslD5* (*rth6*) were expressed in roots (FIG. 5A). While for instance *CslD2* displayed similar expression in older primary roots, seminal roots and crown roots (FIG. 5A), *rth6* was the only *CslD* gene that displayed highly specific expression in root hairs (FIG. 5B).

[00234] The foregoing comparative expression survey of *rth6* transcripts in wildtype versus known maize roothairless mutants suggests that RTH6 directly or indirectly interacts with *rth2, rth3* and *rth5.* While *rth6* expression is upregulated in the mutant *rth2*, its expression is reduced in the mutants *rth3 and rth5.* While the molecular function of *rth2* has not yet been elucidated, *rth3* (Hochholdinger *et al.*, (2008) *Plant J.* 54: 888-98) encodes a COBRA-like cell wall protein, while *rth5* encodes a NADPH-oxidase. COBRA (COB) has been identified as a putative regulator of cellulose synthesis (Schindelmann et al., (2001) Genes Dev. 15:1115–1127). Recently, it has been suggested in Arabidopsis that COBRA and the cellulose synthase complex reside in close proximity on the plasma membrane and that COBRA facilitates cellulose crystallization from the emerging β1-4-glucan chains by acting as a "polysaccharide chaperone" (Sorek *et al.*, (2014) *J. Biol. Chem.* 2014 Dec 12;289(50):34911-20). A functional link between *rth5* and *rth6* has been established by a comparative RNA-Seq analysis of wild-type versus mutant *rth5* roots. Only two GO terms were enriched in this comparative analysis (Nestler *et al.*, (2014) *Plant J.*, 79: 729-740). In addition to "oxidation/reduction" which included some peroxidases which likely act downstream of RTH5 and facilitate the loosening of the elongating root hairs, the GO subgroup "cellulose biosynthesis" was overrepresented. Cellulose biosynthesis follows peroxidase mediated cell wall loosening (Nestler *et al.*, (2014) *Plant J.*, 79: 729-740). The genes *rth3*, *rth5* and *rth6* which are all functionally linked to cellulose biosynthesis are all affected at a very early stage of root hair formation around the formation of the bulge.

EXAMPLE 14: SINGLE-NUCLEOTIDE POLYMORPHISMS ASSOCIATED WITH *RTH6*

[00235] Single nucleotide polymorphisms were identified in the Nested Association Mapping (NAM) Recombinant Inbred Lines B97, CML103, CML228, CML247, CML277, CML322, CML333, CML52, CML69, Hp301, IL14H, Ki11, Ki3, Ky21, M162W, M37W, Mo17, Mo18W, Ms71, NC350, NC358, Oh43, Oh7B, P39, Tx303, and Tzi8. The data are given in Table 4. The position for the start codon of *rth6* is 104,608,915.

Table 4.

Position*	Distance from Start Codon (bp)	REF‡	SNP†
104,608,874	-41	G	G/A
104,608,896	-19	G	G/T
104,608,935	20	T	T/C
104,609,040	125	T	T/C
104,609,061	146	T	T/G
104,609,085	170	C	C/T

Position*	Distance from Start Codon (bp)	REF‡	SNP†
104,609,088	173	C	C/T
104,609,124	209	C	C/A
104,609,184	269	C	C/G
104,609,244	329	T	T/C
104,609,364	449	T	T/C
104,609,388	473	T	T/C
104,609,502	587	G	G/T
104,609,560	645	C	C/A
104,609,583	668	G	G/T
104,609,628	713	T	T/C
104,609,696	781	C	C/T
104,609,707	792	A	A/G
104,610,183	1,268	G	G/A
104,610,394	1,479	C	C/T
104,610,489	1,574	G	G/T
104,610,555	1,640	C	C/G
104,610,646	1,731	G	G/A
104,610,764	1,849	A	A/G
104,610,891	1,976	T	T/C
104,610,913	1,998	A	A/G
104,611,203	2,288	C	C/G
104,611,221	2,306	G	G/T
104,611,224	2,309	G	G/A
104,611,260	2,345	G	G/C
104,611,412	2,497	C	C/G
104,611,560	2,645	T	T/C
104,611,563	2,648	G	G/C
104,611,698	2,783	A	A/G
104,611,815	2,900	C	C/A
104,611,940	3,025	T	T/G
104,611,984	3,069	T	T/C
104,612,293	3,378	T	T/C
104,612,335	3,420	A	A/G
104,612,692	3,777	C	C/A
104,612,698	3,783	G	G/A
104,612,723	3,808	G	G/A
104,612,747	3,832	T	T/A
104,612,890	3,975	T	T/C

* Position relative Maize B73 RefGen_V2 (AGPv2) for chromosome 1.

‡ "REF" indicates the nucleotide found in at the indicated position in the reference sequence.

† "SNP" indicates the single nucleotide polymorphism at the indicated position, with the first nucleotide indicating the reference nucleotide and the second nucleotide indicating the polymorphism.

EXAMPLE 15: REGIONS SUITABLE FOR MUTATION IN THE *RTH6* POLYPEPTIDE

[00236] *Rth6* homologues were identified by Blastp (protein-protein BLAST analysis) using as the query sequence polypeptide sequence of SEQ ID NO: 3 against the non-redundant protein sequence database of NCBI. The Blastp software used was that hosted by NCBI as a web-based tool. The Blastp was run using the default parameters of max target sequences: 100; Expect threshold: 10; Matrix: BLOSUM62; Gap Costs: Existence: 11, Extension: 1; Compositional adjustments: Conditional compositional score matrix adjustment. The data from the Blastp analysis was used to construct a multiple sequence alignment used ClusterW2, as hosted by EMBL-EBI as a web-based tool.

[00237]

[00238] The disclosed embodiments comprise mutant forms *RTH6*. For example, it is contemplated that mutations of SEQ ID NO: 3 could be undertaken using the numeroud homolog sequences of SEQ ID NO: 3 provided herein. Sites for potential mutation include the nonconserved sites. Alternatively, conserved sites can be substituted with alternative conserved amino acids.

[00239] Table 7 below provides additional information on the sequences disclosed herein, along with further information pertaining to a multiple sequence alignment.

[00240] Representative sequences of *RTH6* homologues are disclosed herein. The following RTH6 homologues were aligned using standard multiple sequence alignment tool: GmCs1D2x1 (SEQ ID NO: 70); GmCs1D2 (SEQ ID NO: 69); Pv003G040200 (SEQ ID NO: 89); MtCsl (SEQ ID NO: 77); CaCslD3 (SEQ ID NO: 57); BrCslD3 (SEQ ID NO: 54); BrCslD3x1 (SEQ ID NO: 55); EsHypo (SEQ ID NO: 66); CrHypo (SEQ ID NO: 62); CasCslD3 (SEQ ID NO: 58); AtCslD3 (SEQ ID NO: 51); AlHypo (SEQ ID NO: 50); *Th*CslD3 (SEQ ID NO: 98); CcCslD3x1 (SEQ ID NO: 59); PeCslD2 (SEQ ID NO: 85); PtCs3 (SEQ ID NO: 88); JcCslD3 (SEQ ID NO: 73); RcCsAsub3 (SEQ ID NO: 91); GrCslD3 (SEQ ID NO: 71); TcCslD3 (SEQ ID NO: 97); GaCslD3 (SEQ ID NO: 68); MnCslD3 (SEQ ID NO: 76); VvCslD3 (SEQ ID NO: 100); MdCslD3 (SEQ ID NO: 75); PxbCslD3 (SEQ ID NO: 90); PmCslD3 (SEQ ID NO: 86); PpHypo (SEQ ID NO: 87); FvCslD3 (SEQ ID NO: 67); CmCslD3 (SEQ ID NO: 60); CsCslD3 (SEQ ID NO: 63); CmCslD3x1 (SEQ ID NO: 61); NnCslD3 (SEQ ID NO: 78); BvCslD3 (SEQ ID NO: 56); NtCslD3 (SEQ ID NO: 79); SeiCslD3x1 (SEQ ID NO: 94); *Pd*CSLD2x2 (SEQ ID NO: 84);

*Pd*CSLD2x1 (SEQ ID NO: 84); *Eg*CslD2 (SEQ ID NO: 64); *Pd*CslD2 (SEQ ID NO: 82); SiCslD2 (SEQ ID NO: 96); Sb10g000980 (SEQ ID NO: 93); BdCslD2 (SEQ ID NO: 53); HvHypo (SEQ ID NO: 72) ; MaCs1D2 (SEQ ID NO: 75); *Sb*01g027880 (SEQ ID NO: 93); *Si*CslD1 (SEQ ID NO: 95); *Zm*CslD1 (SEQ ID NO: 101); *Aet*CslD1 (SEQ ID NO: 49); TuCs1LD1 (SEQ ID NO: 99); *Bd*CslD1 (SEQ ID NO: 52); *Os*CslD1 (SEQ ID NO: 81); and *Ob*CslD1 (SEQ ID NO: 80). Additional advantages of the invention will be set forth in part in the description which follows, and in part will be obvious from the description, or can be learned by practice of the invention. The advantages of the invention will be realized and attained by means of the elements and combinations particularly pointed out in the appended claims. It is to be understood that both the foregoing general description and the following detailed description are exemplary and explanatory only and are not restrictive of the invention, as claimed.

[00241]

Table 7.

SEQ ID NO.	Name in the figure	Species	Description
49	*Aet*CslD1	Aegilops tauschii	cellulose synthase-like protein D1
50	AlHypo	Arabidopsis lyrata subsp. lyrata	hypothetical protein ARALYDRAFT_477551
51	AtCslD3	Arabidopsis thaliana	cellulose synthase-like protein
52	*Bd*CslD1	Brachpodium distachyon	cellulose synthase-like protein D1
53	BdCslD2	Brachypodium distachyon	cellulose synthase-like protein D2
54	BrCslD3	Brassica rapa	cellulose synthase-like protein D3
55	BrCslD3x1	Brassica rapa	cellulose synthase-like protein D3, isoform x1
56	BvCslD3	Beta vulgaris subsp. vulgaris	cellulose synthase-like protein D3
57	CaCslD3	Cicer arietinum	cellulose synthase-like protein D3
58	CasCslD3	Camelina sativa	cellulose synthase-like protein D3
59	CcCslD3x1	Citrus clementina	hypothetical protein CICLE_v10023400mg
60	CmCslD3	Cucumis melo	cellulose synthase-like protein D3
61	CmCslD3x1	Cucumis melo	cellulose synthase-like protein D3, isoform x1
62	CrHypo	Capsella rubella	hypothetical protein CARUB_v10012847mg
63	CsCslD3	Cucumis sativus	cellulose synthase-like protein D3
64	*Eg*CslD2	Elaeis guineensus	cellulose synthase-like protein D2
65	EgCslD2x1	Elaeis guineensis	cellulose synthase-like protein D2 isoform X1

SEQ ID NO.	Name in the figure	Species	Description
66	EsHypo	Eutrema salsugineum	hypothetical protein EUTSA_v10019935mg
67	FvCslD3	Fragaria vesca subsp. vesca	cellulose synthase-like protein D3
68	GaCslD3	Gossypium arboreum	cellulose synthase-like protein D3
69	GmCslD2	Glycine max	cellulose synthase-like protein D2-like
70	GmCslD2x1	Glycine max	cellulose synthase-like protein D2-like isoform X1
71	GrCslD3	Gossypium raimondii	cellulose synthase-like protein D3
72	HvHypo	Hordeum vulgare subsp. vulgare	predicted protein
73	JcCslD3	Jatropha curcas	cellulose synthase-like protein D3
74	MaCslD2	Musa acuminata subsp. malaccensis	cellulose synthase-like protein D2
75	MdCslD3	Malus domestica	cellulose synthase-like protein D3
76	MnCslD3	Morus notabilis	cellulose synthase-like protein D3
77	MtCsl	Medicago truncatula	cellulose synthase-like protein
78	NnCslD3	Nelumbo nucifera	cellulose synthase-like protein D3
79	NtCslD3	Nicotiana tomentosiformis	cellulose synthase-like protein D3
80	*Ob*CslD1	Oryza brachyantha	cellulose synthase-like protein D1
81	*Os*CslD1	Oryza sativa	cellulose synthase-like protein D1
82	*Pd*CslD2	Phoenix dactylifera	cellulose synthase-like protein D2
83	*Pd*CSLD2x1	Phoenix dactylifera	cellulose synthase-like protein D2, isoform X1
84	*Pd*CSLD2x2	Phoenix dactylifera	cellulose synthase-like protein D2, isoform X2
85	PeCslD2	Populus euphratica	cellulose synthase-like protein D2
86	PmCslD3	Prunus mume	cellulose synthase-like protein D3
87	PpHypo	Prunus persica	hypothetical protein PRUPE_ppa000473mg
88	PtCs3	Populus trichocarpa	cellulase synthase 3 family protein
89	Pv003G040200	Phaseolus vulgaris	hypothetical protein PHAVU_002G040200g
90	PxbCslD3	Pyrus x bretschneideri	cellulose synthase-like protein D3
91	RcCsAsub3	Ricinus communis	cellulose synthase A catalytic subunit 3 [UDP-forming]
92	*Sb*01g027880	Sorghum bicolor	hypothetical protein
93	Sb10g000980	Sorghum bicolor	hypothetical protein SORBIDRAFT_10g000980
94	SeiCslD3x1	Sesamum indicum	cellulose synthase-like protein D3 isoform X1
95	*Si*CslD1	Setaria italica	cellulose synthase-like protein D1
96	SiCslD2	Setaria italica	cellulose synthase-like protein D2

SEQ ID NO.	Name in the figure	Species	Description
97	TcCslD3	Theobroma cacao	Cellulose synthase-like D3
98	*Th*CslD3	Tarenaya hassleriana	cellulose synthase-like protein D3
99	*Tu*CslLD1	Triticum urartu	cellulose synthase-like protein D1
100	VvCslD3	Vitis vinifera	cellulose synthase-like protein D3
101	*Zm*CslD1	Zea mays	cellulose synthase-like protein D1

Table 7, continued.

SEQ ID NO.	Length	Identities	Positives	Gaps
49	1141	1023/1159	1073/1159	23/1159
50	1145	886/1166	989/1166	28/1166
51	1145	888/1166	988/1166	28/1166
52	1151	1035/1163	1085/1163	16/1163
53	1182	874/1150	978/1150	33/1150
54	1146	879/1134	984/1134	26/1134
55	1151	888/1168	992/1168	26/1168
56	1143	880/1166	989/1166	30/1166
57	1141	876/1169	995/1169	38/1169
58	1149	886/1139	986/1139	25/1139
59	1142	894/1166	998/1166	31/1166
60	1146	898/1168	999/1168	31/1168
61	1148	881/1137	985/1137	28/1137
62	1147	891/1166	991/1166	26/1166
63	1146	896/1168	998/1168	31/1168
64	1154	907/1171	1016/1171	31/1171
65	1149	906/117	1016/1171	34/1171
66	1151	883/1149	989/1149	26/1149
67	1149	883/1172	986/1172	45/1172
68	1143	899/1169	997/1169	36/1169
69	1143	889/1144	996/1144	37/1144
70	1143	891/1151	1000/1151	41/1151
71	1144	896/1170	996/1170	37/1170
72	1188	875/1156	980/1156	40/1156
73	1141	891/1171	993/1171	42/1171
74	1159	920/1178	1023/1178	38/1178
75	1146	879/1138	981/1138	23/1138
76	1146	892/1147	990/1147	29/1147
77	1142	866/1137	971/1137	24/1137
78	1147	867/1149	978/1149	36/1149
79	1139	869/1142	979/1142	33/1142
80	1125	987/1165	1037/1165	46/1165
81	1127	988/1166	1045/1166	46/1166

SEQ ID NO.	Length	Identities	Positives	Gaps
82	1148	914/1167	1017/1167	27/1167
83	1148	908/1149	1009/1149	31/1149
84	1134	910/1166	1015/1166	39/1166
85	1143	891/1169	991/1169	36/1169
86	1145	884/1143	979/1143	34/1143
87	1145	886/1153	984/1153	42/1153
88	1143	895/1169	992/1169	36/1169
89	1144	887/1151	993/1151	41/1151
90	1146	879/1138	980/1138	23/1138
91	1143	887/1172	991/1172	42/1172
92	1164	1099/1168	1119/1168	13/1168
93	1179	898/1191	999/1191	44/1191
94	1152	874/1174	991/1174	43/1174
95	1157	1090/1164	1115/1164	12/1164
96	1175	898/1178	998/1178	34/1178
97	1175	897/1197	996/1197	60/1197
98	1141	896/1164	997/1164	28/1164
99	1150	1005/1141	1055/1141	26/1141
100	1149	884/1145	992/1145	35/1145
101	1159	1159/1159	1159/1159	0/1159

[00242] All publications and patent applications mentioned in the specification are indicative of the level of those skilled in the art to which this invention pertains. All publications and patent applications are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

[00243] Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, certain changes and modifications may be practiced within the scope of the appended claims.

[00244] The present invention comprises the sequences referred to herein, SEQ. ID NOs: 1-101, and the full sequences are provided herein below.

CLAIMS

What is claimed is:

1. A method for improving at least one agronomic characteristic in a plant, the method comprising introducing into the plant an expression construct, the construct comprising a promoter operably linked to a polynucleotide, wherein the polynucleotide comprises a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence comprising SEQ ID NO:2;

(b) a nucleotide sequence encoding the amino acid sequence set forth in SEQ ID NO:3;

(c) a nucleotide sequence having at least 90% sequence identity to SEQ ID NO:2, wherein the polynucleotide encodes a polypeptide having ß-glycan synthase activity; and

(d) a nucleotide sequence encoding an amino acid sequence having at least 90% sequence identity to SEQ ID NO:3, wherein the polynucleotide encodes a polypeptide having ß-glycan synthase activity;

wherein the at least one agronomic characteristic is selected from the group consisting of increased root hair formation, increased root growth, increased drought tolerance, enhanced nutrient uptake, increased yield, enhanced nutrient uptake, and combinations thereof.

2. The method of claim 1, wherein the polynucleotide is stably integrated into the genome of the plant.

3. The method of any one of claims 1 or 2, wherein promoter is a heterologous root-preferred promoter.

4. The method of any one of claims 1, 2, or 3, wherein the plant is a plant cell.

5. The method of any one of claims 1, 2, or 3, wherein the plant is selected from the group consisting of rice, corn, soybean, canola, potato, wheat, mung bean, alfalfa, barley, rye, cotton, sunflower, peanut, sweet potato, bean, pea, chicory, lettuce, endive, cabbage, brussel sprout, beet, parsnip, turnip, cauliflower, broccoli, radish, spinach, onion, garlic, eggplant, pepper, celery, carrot, squash, pumpkin, zucchini, cucumber, apple, pear, melon, citrus, strawberry, grape, raspberry, pineapple, tobacco, tomato, sorghum, sugarcane, banana, *Arabidopsis thaliana*, African violet, petunia, pelargonium, poinsettia, chrysanthemum, carnation, crocus, marigold, daffodil, pine, *Medicago truncatula*, *Sandersonia aurantiaca*, and zinnia.

6. The method of any one of claims 1, 2, or 3, wherein the plant is a dicot.

7. The method of claim 6, wherein the dicot is soybean, *Brassica*, sunflower, cotton, or alfalfa.

8. The method of any one of claims 1, 2, or 3, wherein the plant is a monocot.

9. The method of claim 8, wherein the monocot is a cereal grain.

10. The method of claim 9, wherein the cereal grain is *Eleusine coracana*, *Panicum miliaceum*, *Pennisetum glaucum*, *Setaria italica*, *Digitaria exilis*, *Digitaria iburua*, *Paspalum scrobiculatum*, *Echinochloa esculenta*, *Coix lacryma-jobi*, *Zea mays*, *Panicum miliaceum*, *Sorghum* spp., *Hordeum vulgare*, *Avena sativa*, *Oryza sativa*, *Secale cereale*, *Eragrostis tef*, triticale, *Triticum* spp., or *Zizania* spp.

11. The method of claim 8, wherein the monocot is *Poa* spp., *Agrostis* spp., *Lolium* spp., *Festuca* spp., *Calamogrostis* spp., *Deschampsia* spp., *Zoysia* spp., *Cynodon* spp., *Stenotaphrum* spp., *Paspalum* spp., *Eremachloa* spp., *Axonopus* spp., or *Bouteloua* spp.

12. The method of claim 8, wherein the monocot is maize, sugarcane, wheat, rice, barley, sorghum, or rye.

13. The method of claim 1, wherein the enhanced agronomic characteristic is enhanced nutrient uptake.

14. The method of claim 1, wherein the enhanced agronomic characteristic is improved drought tolerance.

15. The method of claim 1, wherein the enhanced agronomic characteristic is increased yield.

16. The method of claim 15, wherein the increased yield is under water-limiting conditions.

17. A plant produced by the method of any one of claims 1-16.

18. The plant of claim 17, wherein the plant comprises the expression construct.

19. A seed of the plant produced by the method of any one of claims 1-16.

20. The seed of claim 19, wherein the seed comprises the expression construct.

21. A plant comprising an expression construct comprising a promoter operably linked to a polynucleotide, wherein the polynucleotide comprises a nucleotide sequence selected from

the group consisting of:

(a) a nucleotide sequence comprising SEQ ID NO: 2;

(b) a nucleotide sequence encoding the amino acid sequence set forth in SEQ ID NO:3;

(c) a nucleotide sequence having at least 90% sequence identity to SEQ ID NO: 2, wherein the polynucleotide encodes a polypeptide having ß-glycan synthase activity; and

(d) a nucleotide sequence encoding an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 3, wherein the polynucleotide encodes a polypeptide having ß-glycan synthase activity.

22. The plant of claim 21, wherein the plant is a cell.

23. The plant of claim 21, wherein the plant is selected from the group consisting of rice, corn, soybean, canola, potato, wheat, mung bean, alfalfa, barley, rye, cotton, sunflower, peanut, sweet potato, bean, pea, chicory, lettuce, endive, cabbage, brussel sprout, beet, parsnip, turnip, cauliflower, broccoli, radish, spinach, onion, garlic, eggplant, pepper, celery, carrot, squash, pumpkin, zucchini, cucumber, apple, pear, melon, citrus, strawberry, grape, raspberry, pineapple, tobacco, tomato, sorghum, sugarcane, banana, *Arabidopsis thaliana*, African violet, petunia, pelargonium, poinsettia, chrysanthemum, carnation, crocus, marigold, daffodil, pine, *Medicago truncatula*, *Sandersonia aurantiaca*, and zinnia.

24. The plant of claim 21, wherein the plant is a monocot.

25. The plant of claim 22, wherein the monocot is a cereal grain.

26. The plant of claim 25, wherein the cereal grain is *Eleusine coracana*, *Panicum miliaceum*, *Pennisetum glaucum*, *Setaria italica*, *Digitaria exilis*, *Digitaria iburua*, *Paspalum scrobiculatum*, *Echinochloa esculenta*, *Coix lacryma-jobi*, *Zea mays*, *Panicum miliaceum*, *Sorghum* spp., *Hordeum vulgare*, *Avena sativa*, *Oryza sativa*, *Secale cereale*, *Eragrostis tef*, triticale, *Triticum* spp., or *Zizania* spp.

27. The plant of claim 24, wherein the monocot is *Poa* spp., *Agrostis* spp., *Lolium* spp., *Festuca* spp., *Calamogrostis* spp., *Deschampsia* spp., *Zoysia* spp., *Cynodon* spp., *Stenotaphrum* spp., *Paspalum* spp., *Eremachloa* spp., *Axonopus* spp., or *Bouteloua* spp.

28. The plant of claim 24, wherein the monocot is maize, sugarcane, wheat, rice, barley, sorghum, or rye.

29. The plant of claim 21, wherein the plant is a dicot.

30. The plant of claim 29, wherein the dicot is soybean, *Brassica*, sunflower, cotton, or alfalfa.

31. The plant of any one of claims 21 to 30 wherein the expression construct is stably incorporated into the genome of the plant.

32. The plant of any one of claims 21 to 31, wherein the promoter is a heterologous root-preferred promoter.

33. A seed of the plant of claim 32.

34. The seed of claim 33, wherein the seed comprises the expression construct.

35. An expression cassette comprising a polynucleotide operably linked to a promoter that drives expression of the polynucleotide in a root of a plant, wherein the polynucleotide comprises a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence comprising SEQ ID NO: 2;

(b) a nucleotide sequence encoding the amino acid sequence set forth in SEQ ID NO: 3;

(c) a nucleotide sequence having at least 90% sequence identity to SEQ ID NO: 2, wherein the polynucleotide encodes a polypeptide having ß-glycan synthase activity; and

(d) a nucleotide sequence encoding an amino acid sequence having at least 90% sequence identity to SEQ ID NO: 3, wherein the polynucleotide encodes a polypeptide having ß-glycan synthase activity.

36. The expression cassette of claim 35, wherein the plant is selected from the group consisting of rice, corn, soybean, canola, potato, wheat, mung bean, alfalfa, barley, rye, cotton, sunflower, peanut, sweet potato, bean, pea, chicory, lettuce, endive, cabbage, brussel sprout, beet, parsnip, turnip, cauliflower, broccoli, radish, spinach, onion, garlic, eggplant, pepper, celery, carrot, squash, pumpkin, zucchini, cucumber, apple, pear, melon, citrus, strawberry, grape, raspberry, pineapple, tobacco, tomato, sorghum, sugarcane, banana, *Arabidopsis thaliana*, African violet, petunia, pelargonium, poinsettia, chrysanthemum, carnation, crocus, marigold, daffodil, pine, *Medicago truncatula*, *Sandersonia aurantiaca*, and zinnia.

37. The expression cassette of claim 35, wherein the plant is a monocot.

38. The expression cassette of claim 37, wherein the monocot is a cereal grain.

39. The expression cassette of claim 38, wherein the cereal grain is *Eleusine coracana*,

Panicum miliaceum, *Pennisetum glaucum*, *Setaria italica*, *Digitaria exilis*, *Digitaria iburua*, *Paspalum scrobiculatum*, *Echinochloa esculenta*, *Coix lacryma-jobi*, *Zea mays*, *Panicum miliaceum*, *Sorghum* spp., *Hordeum vulgare*, *Avena sativa*, *Oryza sativa*, *Secale cereale*, *Eragrostis tef*, triticale, *Triticum* spp., or *Zizania* spp.

40. The expression cassette of claim 37, wherein the monocot is *Poa* spp., *Agrostis* spp., *Lolium* spp., *Festuca* spp., *Calamogrostis* spp., *Deschampsia* spp., *Zoysia* spp., *Cynodon* spp., *Stenotaphrum* spp., *Paspalum* spp., *Eremachloa* spp., *Axonopus* spp., or *Bouteloua* spp.

41. The expression cassette of claim 37, wherein the monocot is maize, sugarcane, wheat, rice, barley, sorghum, or rye.

42. The expression cassette of claim 35, wherein the plant is a dicot.

43. The expression cassette of claim 42, wherein the dicot is soybean, *Brassica*, sunflower, cotton, or alfalfa.

44. The expression cassette of claim 35, wherein the promoter is a heterologous root-preferred promoter.

45. A method of selecting an allelic variant of *rth6* in a maize plant, the method comprising the steps of:

(a) obtaining a population of maize plants, wherein the maize plants exhibit an alteration of at least one agronomic characteristic; wherein the at least one agronomic characteristic is selected from the group consisting of: increased root hair formation, root hair growth, increased drought tolerance, enhanced nutrient uptake, increased yield and combinations thereof;

(b) evaluating allelic variations with respect to the polynucleotide sequence encoding a protein comprising SEQ ID NO: 3, or in the genomic region that regulates the expression of the polynucleotide encoding the protein;

(c) associating allelic variations with the alteration of at least one agronomic characteristic; and

(d) selecting an allelic variant that is associated with the alteration of at least one agronomic characteristic.

46. A method of selecting a first maize plant or a first maize germplasm that has one or more beneficial alleles of *rth6*, the method comprising:

(a) screening a plurality of maize plants or a plurality of maize germplasm for

at least one polymorphism within a marker locus, wherein the marker locus is:

(i) a first polynucleotide having at least 90% and less than 100% nucleotide sequence identity with SEQ ID NO: 1 or 2; or

(ii) a second polynucleotide encoding a polypeptide comprising an amino acid sequence having at least 90% and less than 100% sequence identity to SEQ ID NO: 3, wherein expression of the first or second polynucleotide in a maize plant results in a phenotype comprising an alteration of at least one agronomic characteristic when compared to a control maize plant; wherein the at least one agronomic characteristic is selected from the group consisting of increased root hair formation, increased root hair growth, increased drought tolerance, or enhanced nutrient uptake, and combinations thereof; and wherein the control maize plant comprises:

a polynucleotide having the nucleotide sequence of SEQ ID NO: 1 or 2; or

a polynucleotide encoding the amino acid sequence of SEQ ID NO: 3;

(b) identifying a first maize plant or a first maize germplasm comprising the at least one polymorphism of the marker locus; and

(c) selecting the first maize plant or first maize germplasm of step (b).

47. A method of reducing soil erosion in a crop field, the method comprising the step of planting a crop field with a plant comprising an expression construct, the construct comprising a promoter operably linked to a polynucleotide, wherein the polynucleotide comprises a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence comprising SEQ ID NO: 2;

(b) a nucleotide sequence encoding the amino acid sequence set forth in SEQ ID NO:3;

(c) a nucleotide sequence having at least 90% sequence identity to SEQ ID NO: 2, wherein the polynucleotide encodes a polypeptide having ß-glycan synthase activity; and

(d) a nucleotide sequence encoding an amino acid sequence having at least 90% sequence identity to SEQ ID NO:3, wherein the polynucleotide encodes a polypeptide having ß-glycan synthase activity.

48. A method of increasing infection of a plant with a beneficial microorganism, said method comprising introducing into said plant an expression construct, said construct comprising a promoter that drives expression in a plant root operably linked to a polynucleotide, wherein said polynucleotide comprises a nucleotide sequence selected from the group consisting of:

(a) a nucleotide sequence comprising SEQ ID NO: 2;

(b) a nucleotide sequence encoding the amino acid sequence set forth in SEQ ID NO: 3;

(c) a nucleotide sequence having at least 90% sequence identity to SEQ ID NO: 2, wherein the polynucleotide encodes a polypeptide having ß-glycan synthase activity; and

(d) a nucleotide sequence encoding an amino acid sequence having at least 90% sequence identity to SEQ ID NO:3, wherein the polynucleotide encodes a polypeptide having ß-glycan synthase activity.

49. The method of claim 48, further comprising:

(e) obtaining seed of the plant, wherein the seed comprises the expression construct; and

(f) contacting the seed with the beneficial microorganism.

50. The method of claim 48 or 49, wherein the beneficial microorganism is *Pseudomonas putida* KT2440.

51. The method of claim 48 or 49, wherein the promoter is a heterologous promoter.

52. The method of claim 51, wherein the promoter is a heterologous root-preferred promoter.

53. A method of increasing resistance of a plant to a root pathogen, the method comprising the step of reducing the expression of a *rth6* gene family member or inhibiting the activity a *rth6* polypeptide.

54. The method of claim 53, wherein the root pathogen is *Fusarium* spp.

55. The method of claim 54, wherein the *Fusarium* spp. is *Fusarium verticilloides* or *Fusarium graminearum*.

56. A method of identifying a first corn plant or germplasm that displays tolerance, improved tolerance, or susceptibility to drought; the method comprising detecting in the first corn plant or germplasm at least one allele of a quantitative trait locus that is associated with the tolerance, improved tolerance, or susceptibility; wherein the quantitative locus is a chromosomal interval located within about 10 kbp of an interval at 104,616,335 - 104,622,459 of chromosome 1 (SEQ ID NO: 1).

57. The method of claim 56, wherein quantitative trait locus is localized within about 5

kbp of an interval at 104,616,335 - 104,622,459 of chromosome 1 (SEQ ID NO: 1).

58. The method of claim 56, wherein quantitative trait locus is localized to the chromosomal interval at 104,616,335 - 104,622,459 of chromosome 1 (SEQ ID NO: 1).

59. The method of claim 56, wherein the quantitative trait locus is associated with at least one marker allele; and wherein marker allele comprises at least one polymorphism set forth in Table 4.

60. The method of claim 56, wherein detecting comprises a PCR method, hybridization method, next-genome sequence method, comparative mapping method, or a bulk segregation analysis method.

61. The method of claim 56, wherein detecting comprises quantitative PCR.

62. The method of claim 61, wherein the PCR amplifies an amplicon within the chromosomal interval at 104,616,335 - 104,622,459 of chromosome 1 (SEQ ID NO: 1).

63. The method of claim 61, wherein PCR comprises a labeled primer pair.

64. The method of claim 56, wherein detecting comprises detection of a labeled probe.

65. The method of claim 64, wherein the labeled probe comprises an oligonucleotide complementary to a sequence in the chromosomal interval at 104,616,335 - 104,622,459 of chromosome 1 (SEQ ID NO: 1).

66. The method of claim 56, wherein the quantitative trait locus is further associated with enhanced nutrient uptake in a plant.

67. The method of claim 56, further comprising the steps of:

a) selecting the first corn plant or germplasm, or selecting a progeny of the first corn plant or germplasm; and

b) crossing the selected first corn plant or germplasm with a second corn plant or germplasm to introgress the quantitative trait locus into progeny corn germplasm.

68. The method of claim 67, wherein the second corn plant or germplasm displays less tolerance to drought as compared to the first corn plant or germplasm, and wherein the introgressed corn plant or germplasm displays an increased tolerance to drought as compared to the second plant or germplasm.

69. The method of claim 67, further comprising the steps of:

(c) analyzing progeny corn germplasm to determine the presence of tolerance to drought; and

(d) selecting progeny corn germplasm that tests positive for the presence of tolerance to drought as being drought germplasm into which germplasm having said quantitative trait locus has been introgressed.

70. An introgressed corn plant or germplasm produced by the method of any of claims 56-69.

71. A kit for selecting at least one corn plant by marker assisted selection of a quantitative trait locus associated with at least one improved agronomic characteristic:

(a) labeled primers or probes for detecting at least one nucleic acid sequence comprising a sequence in the chromosomal interval at 104,616,335 - 104,622,459 of chromosome 1 (SEQ ID NO: 1), or comprising a sequence complementary to a sequence therein; and

(b) instructions for using the primers or probes to detect the marker loci and correlating the loci with predicted improved agronomic characteristic;

wherein the at least one agronomic characteristic selected from the group consisting of increased root hair formation, increased root growth, increased drought tolerance, enhanced nutrient uptake, increased yield, and combinations thereof.

72. A seed preparation comprising a seed and a beneficial microorganism, wherein the seed comprises the expression construct of claim 35.

73. A plant comprising in its genome an endogenous polynucleotide operably linked to at least one heterologous regulatory element, wherein said endogenous polynucleotide encodes a polypeptide having an amino acid sequence of at least 90% sequence identity, when compared to SEQ ID NO: 3, and wherein said plant exhibits at least one trait selected from the group consisting of: increased root hair formation and growth, increased drought tolerance, enhanced nutrient uptake, or combinations thereof, when compared to a control plant not comprising the heterologous regulatory element.

74. The plant of claim 73, wherein the at least one heterologous regulatory element is selected from the group consisting of: a promoter, an enhancer and an intron.

75. A method of making the plant of claim 73, wherein the method comprises:

(a) introducing into a plant cell a double-stranded break-inducing agent and at least one heterologous regulatory element;

(b) regenerating a plant from the plant cell of step (a);

(c) selecting a plant from step (b) that comprises an endogenous polynucleotide operably linked to the at least one heterologous regulatory element, wherein said endogenous polynucleotide encodes a polypeptide having an amino acid sequence of at least 90% sequence identity, when compared to SEQ ID NO: 3; and

(d) selecting the plant of step (c) that exhibits at least one trait selected from the group consisting of: increased root hair formation and growth, increased drought tolerance, and enhanced nutrient uptake, when compared to a control plant not comprising the heterologous regulatory element.

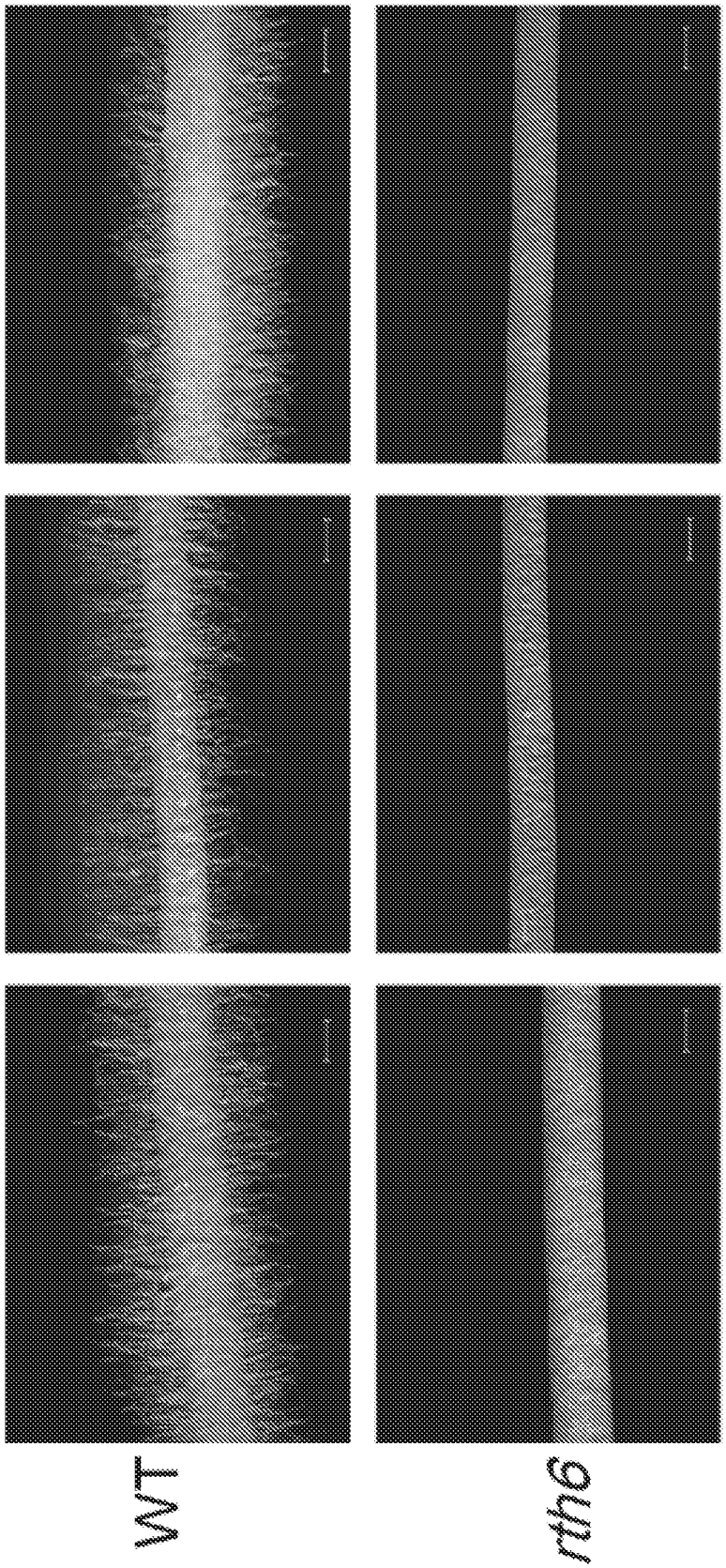


FIG. 1A

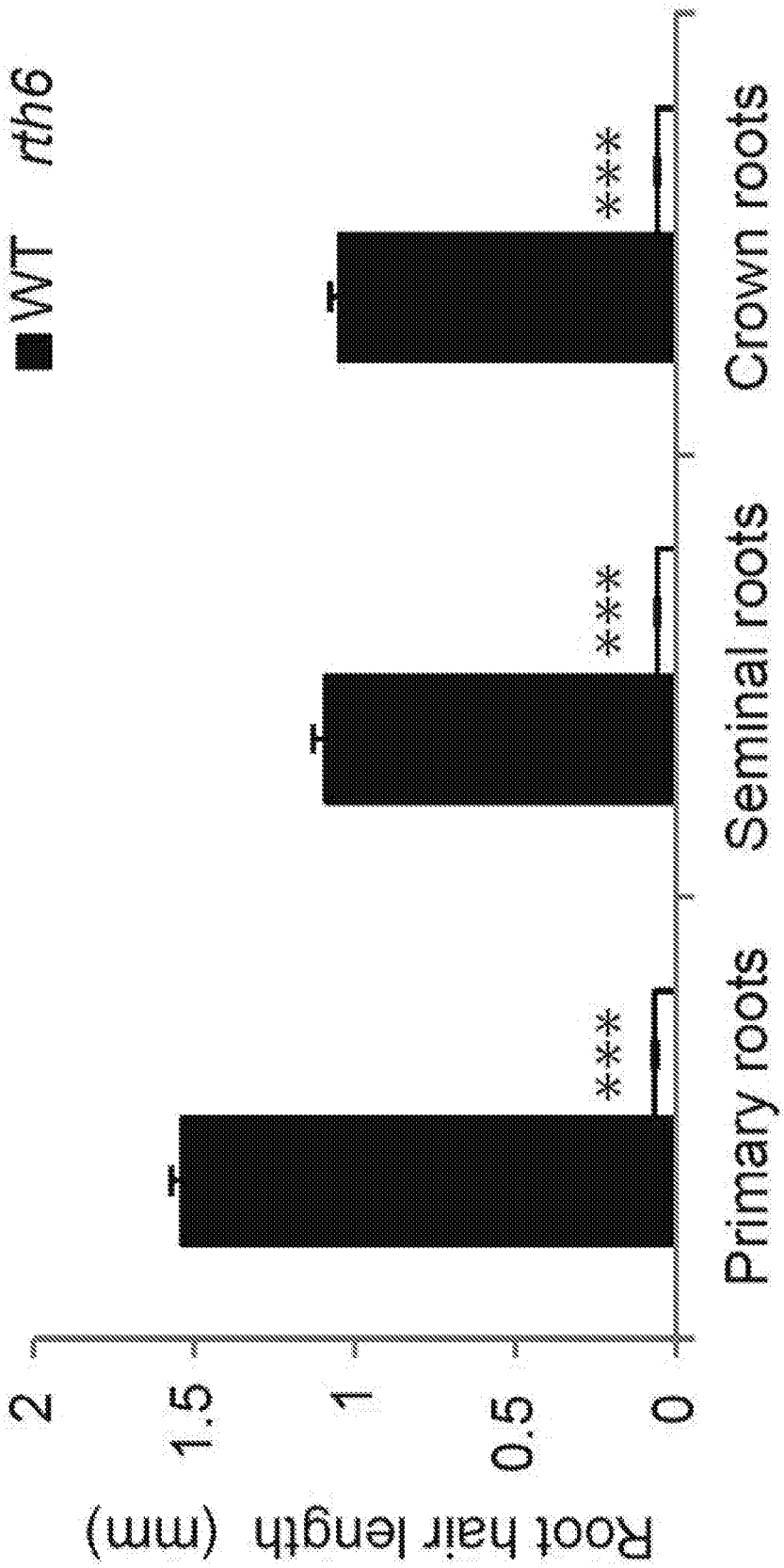
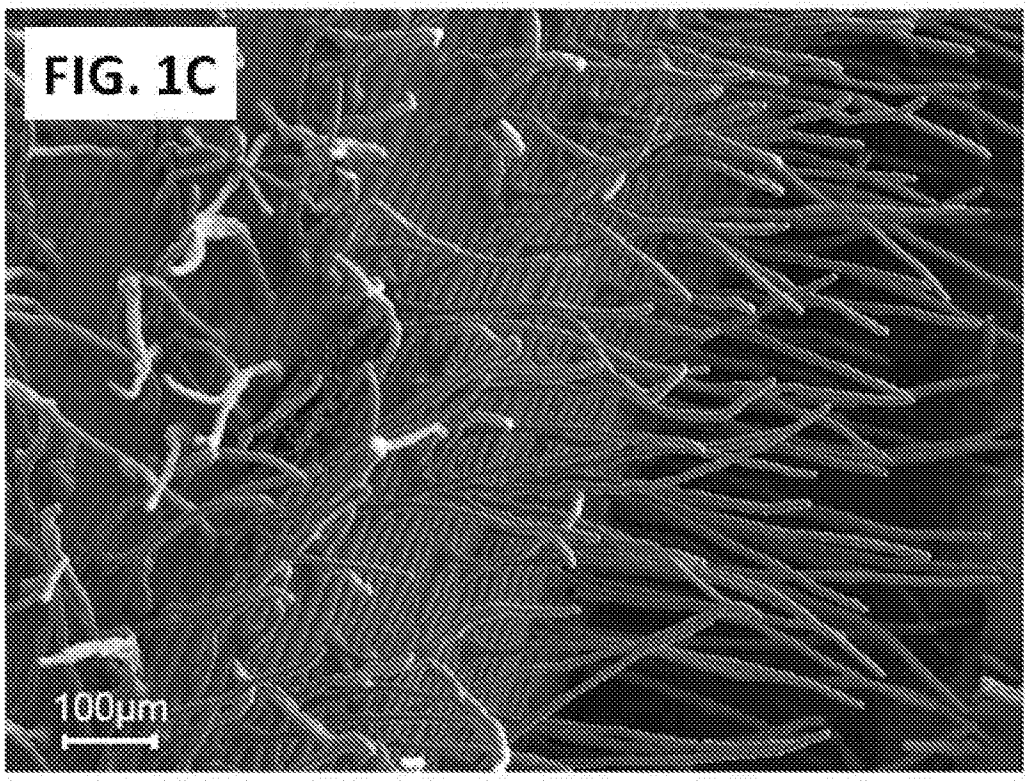
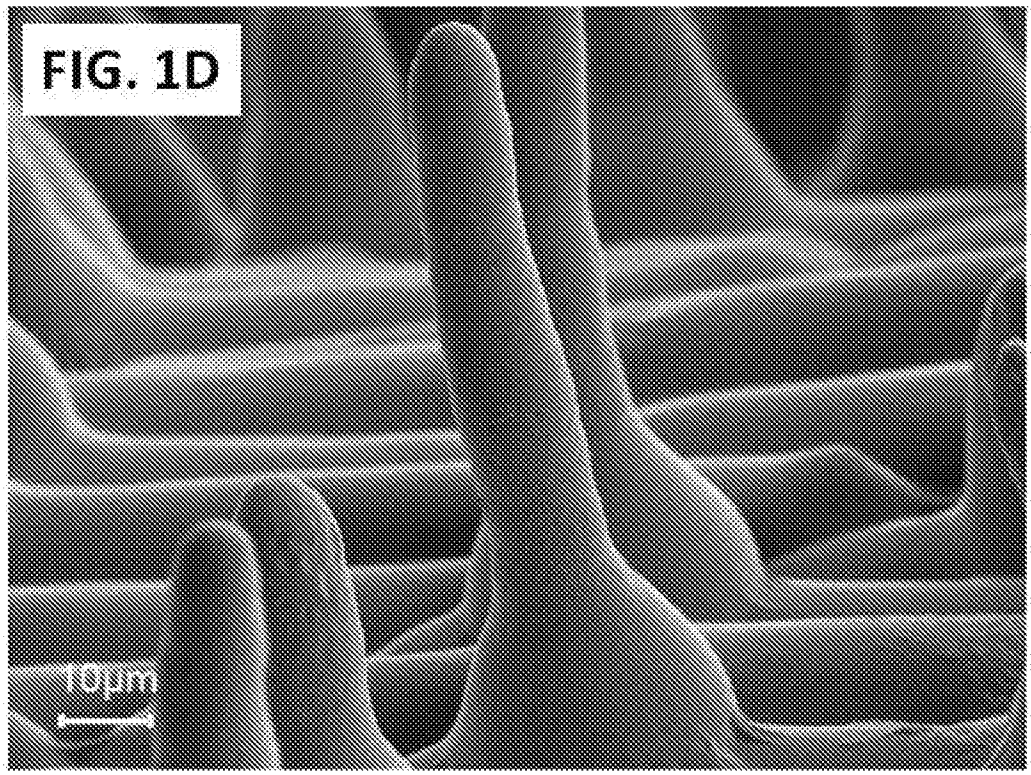
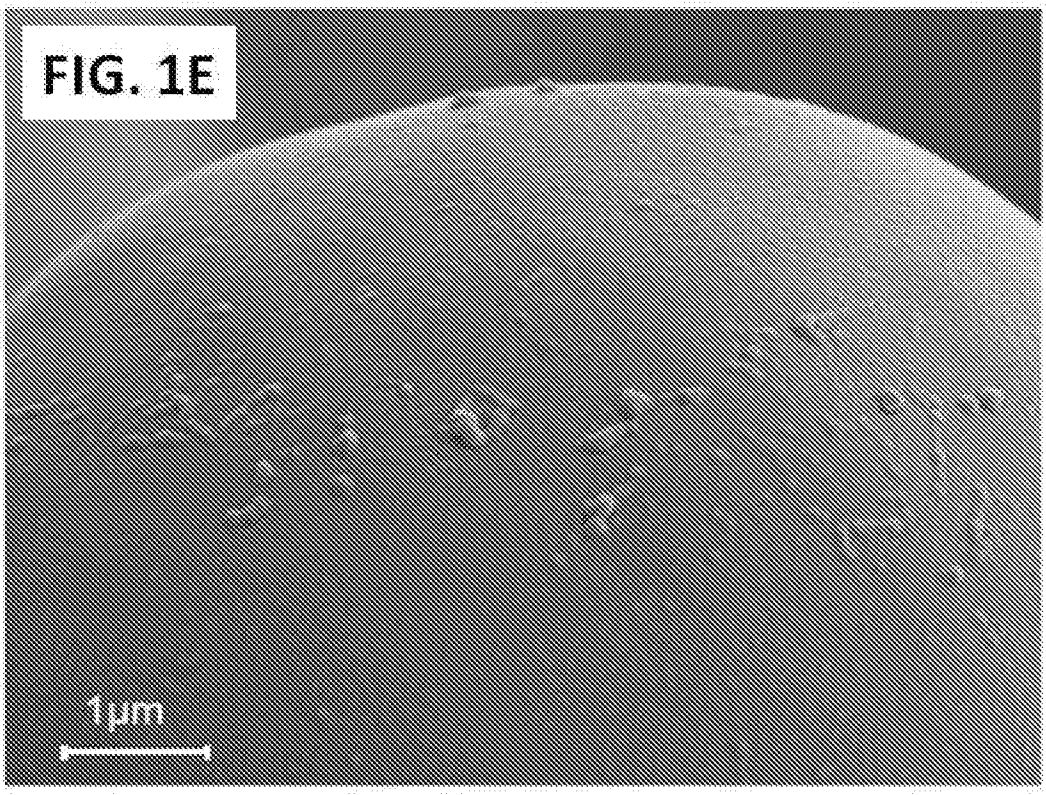

Root hair length (mm)
2
1.5
1
0.5
0
WT
rth6
***
***
***
Primary roots
Seminal roots
Crown roots

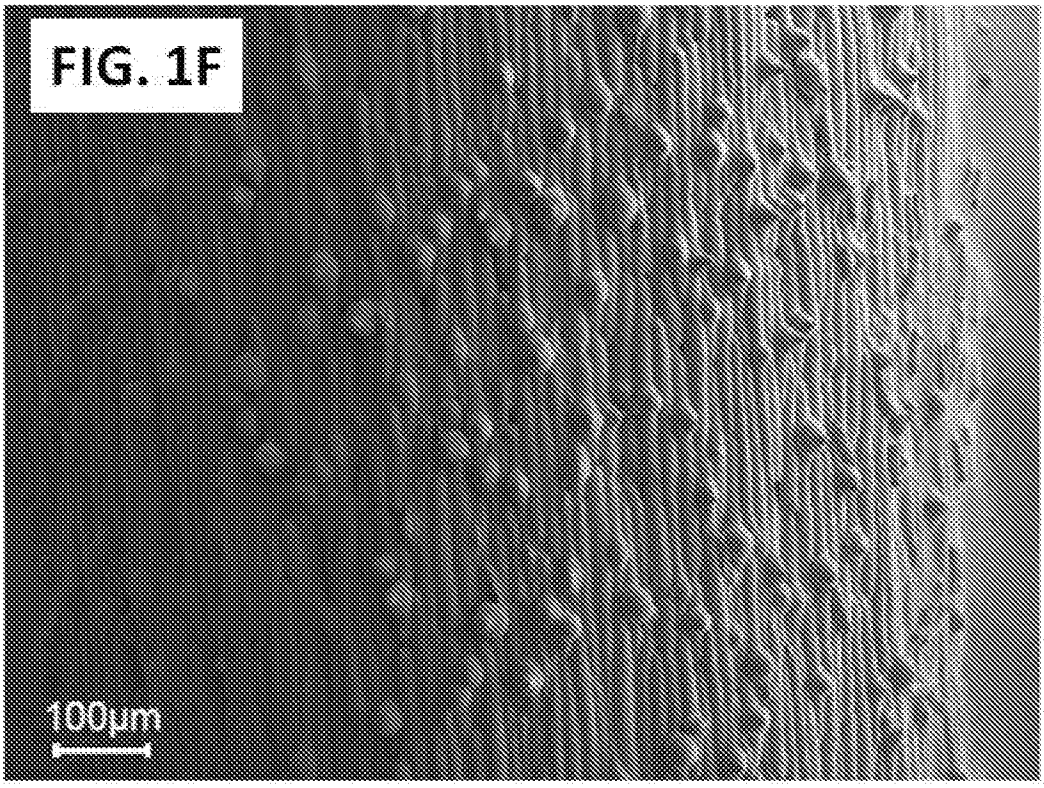
FIG. 1B



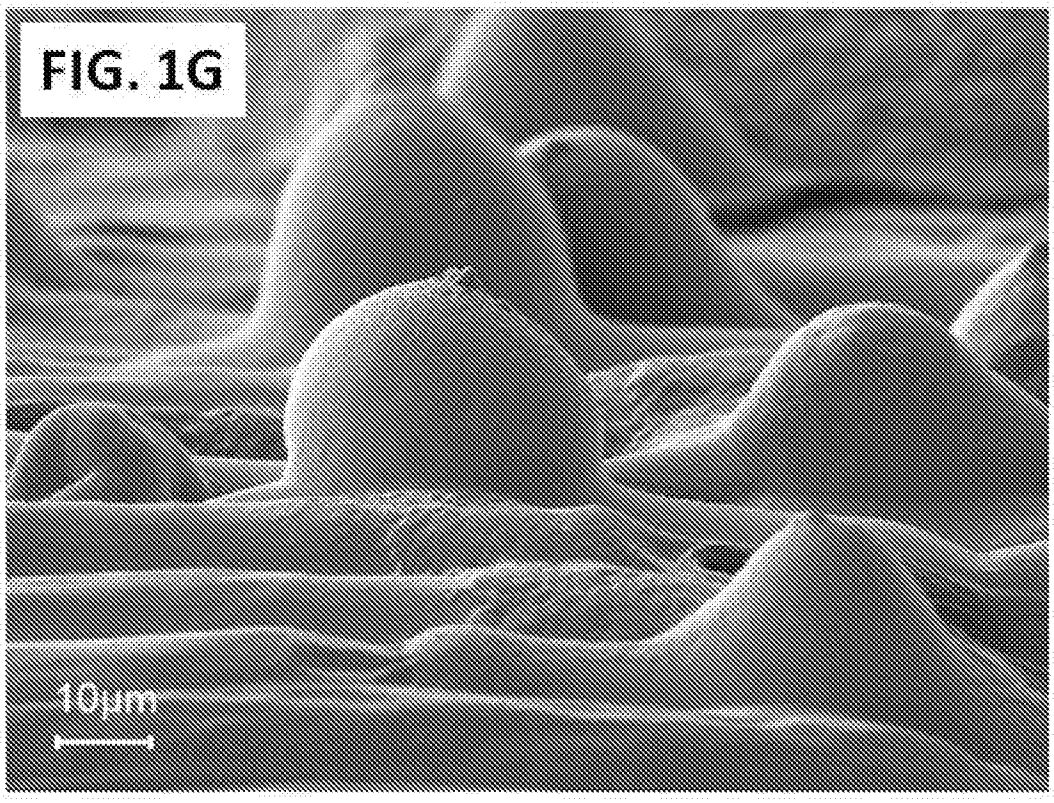


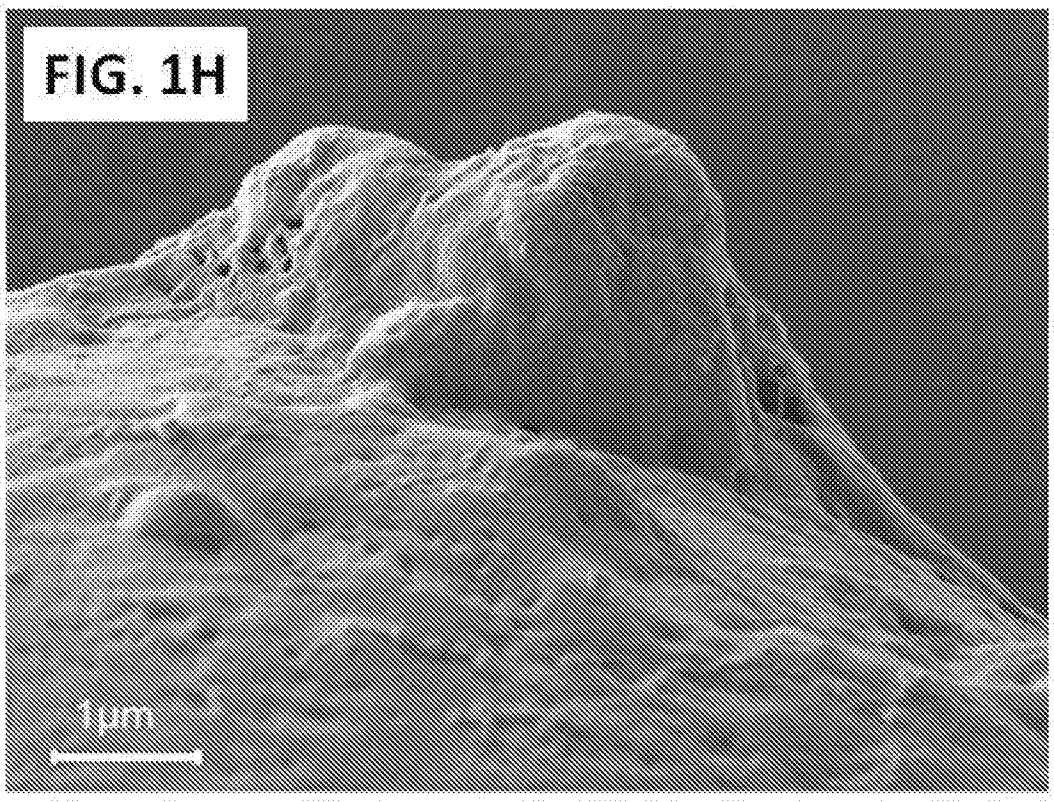
FIGS 1C-1D





FIGS 1E-1F





FIGS. 1G-1H

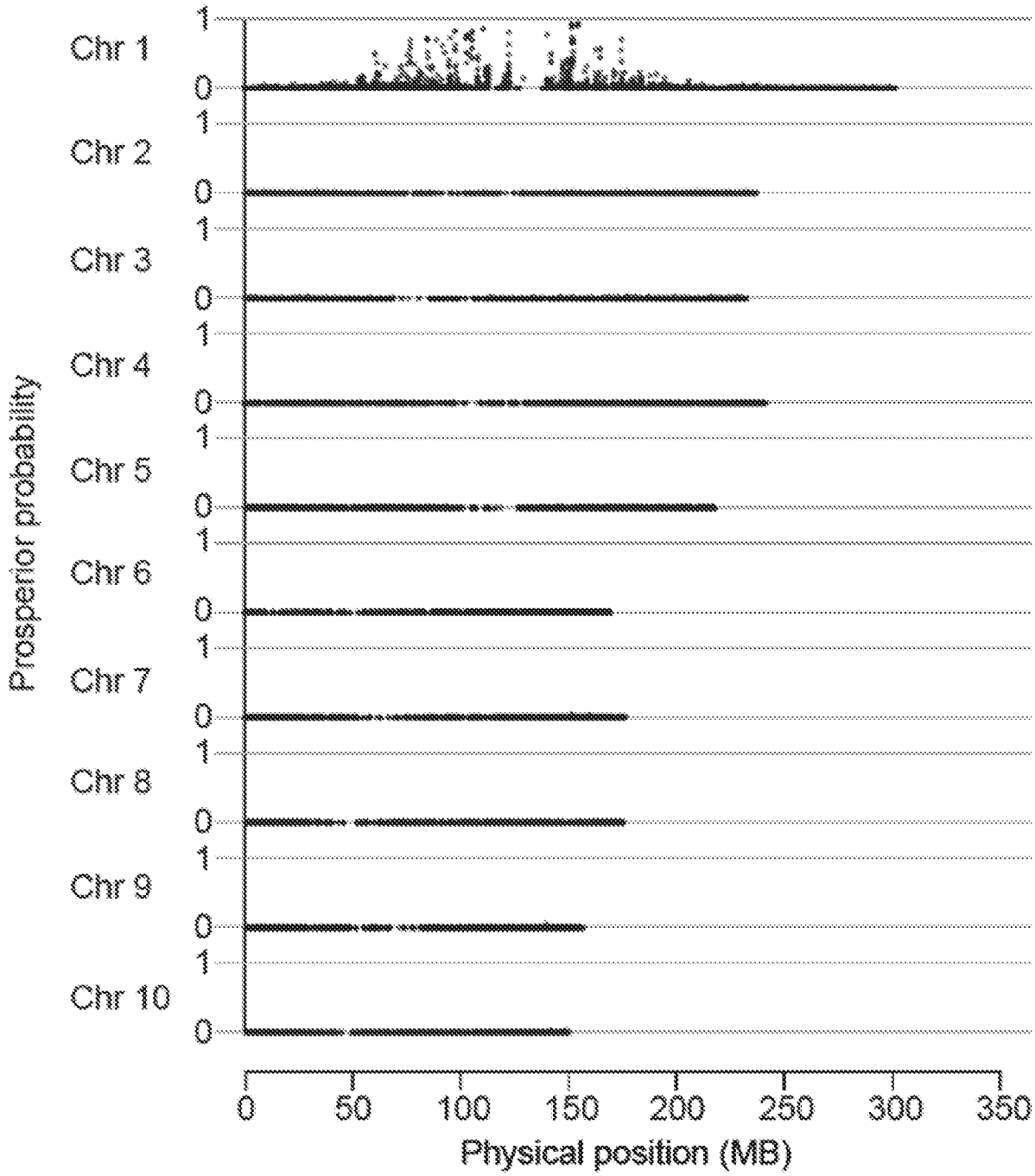

Chr 1
Chr 2
Chr 3
Chr 4
Chr 5
Chr 6
Chr 7
Chr 8
Chr 9
Chr 10
1
0
Prosperior probability
0
50
100
150
200
250
300
350
Physical position (MB)

FIG. 2A

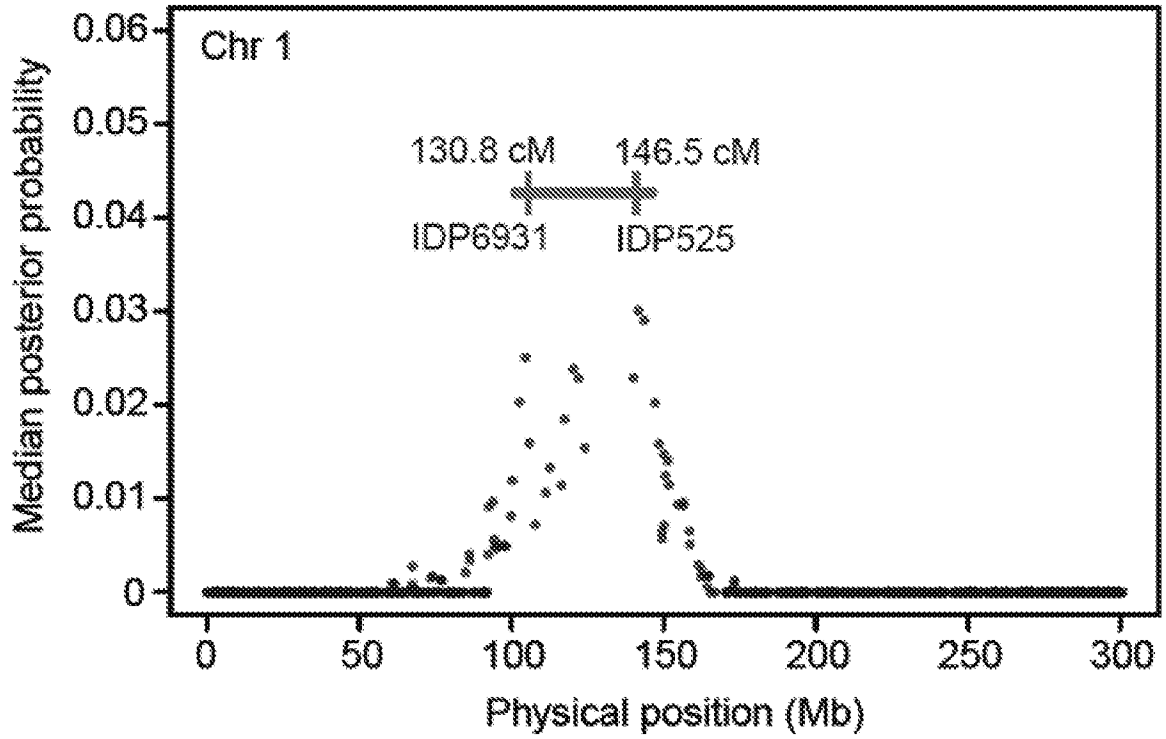

Chr 1
130.8 cM
146.5 cM
IDP6931
IDP525
Median posterior probability
0.06
0.05
0.04
0.03
0.02
0.01
0
0
50
100
150
200
250
300
Physical position (Mb)

FIG. 2B

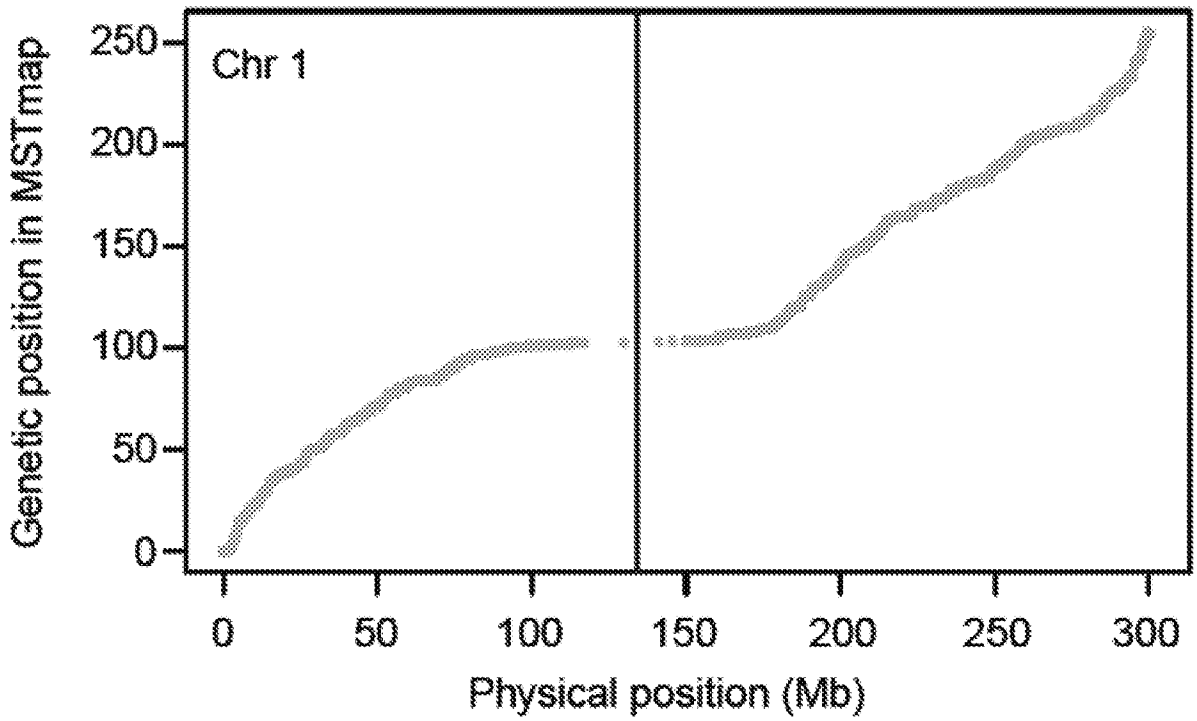

Chr 1
Genetic position in MSTmap
250
200
150
100
50
0
0
50
100
150
200
250
300
Physical position (Mb)

FIG. 2C

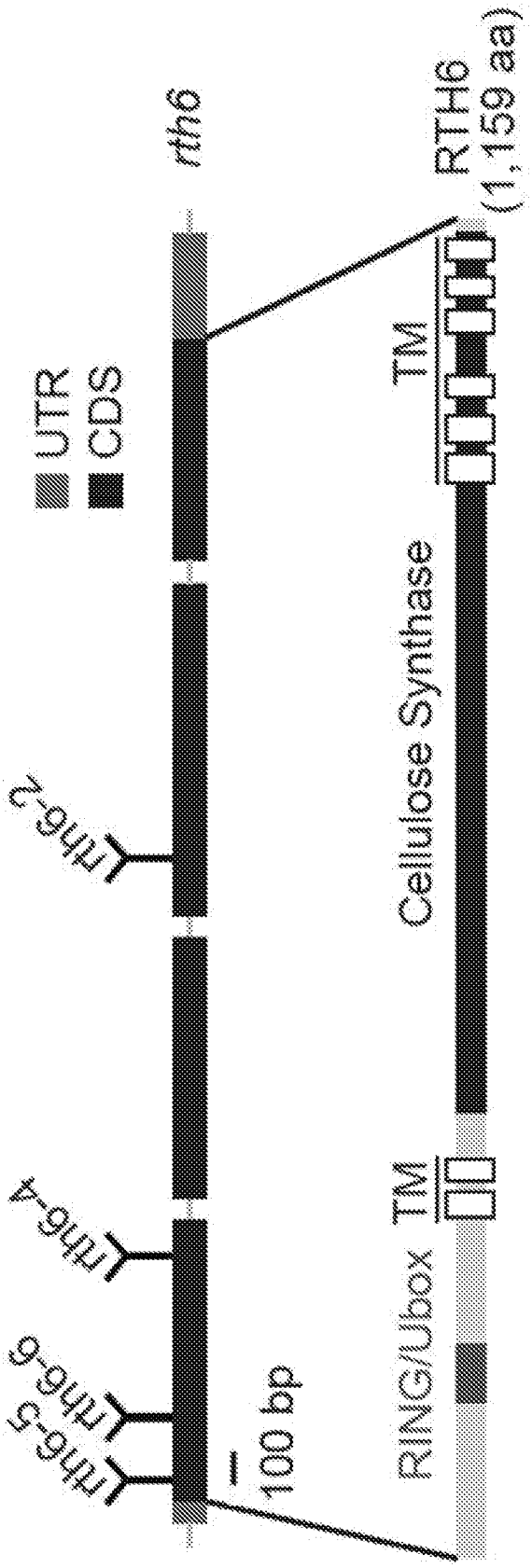

rth6
RTH6
(1,159 aa)
UTR
CDS
TM
Cellulose Synthase
rth6-2
rth6-4
rth6-6
rth6-5
RING/Ubox
TM
100 bp

FIG. 2D

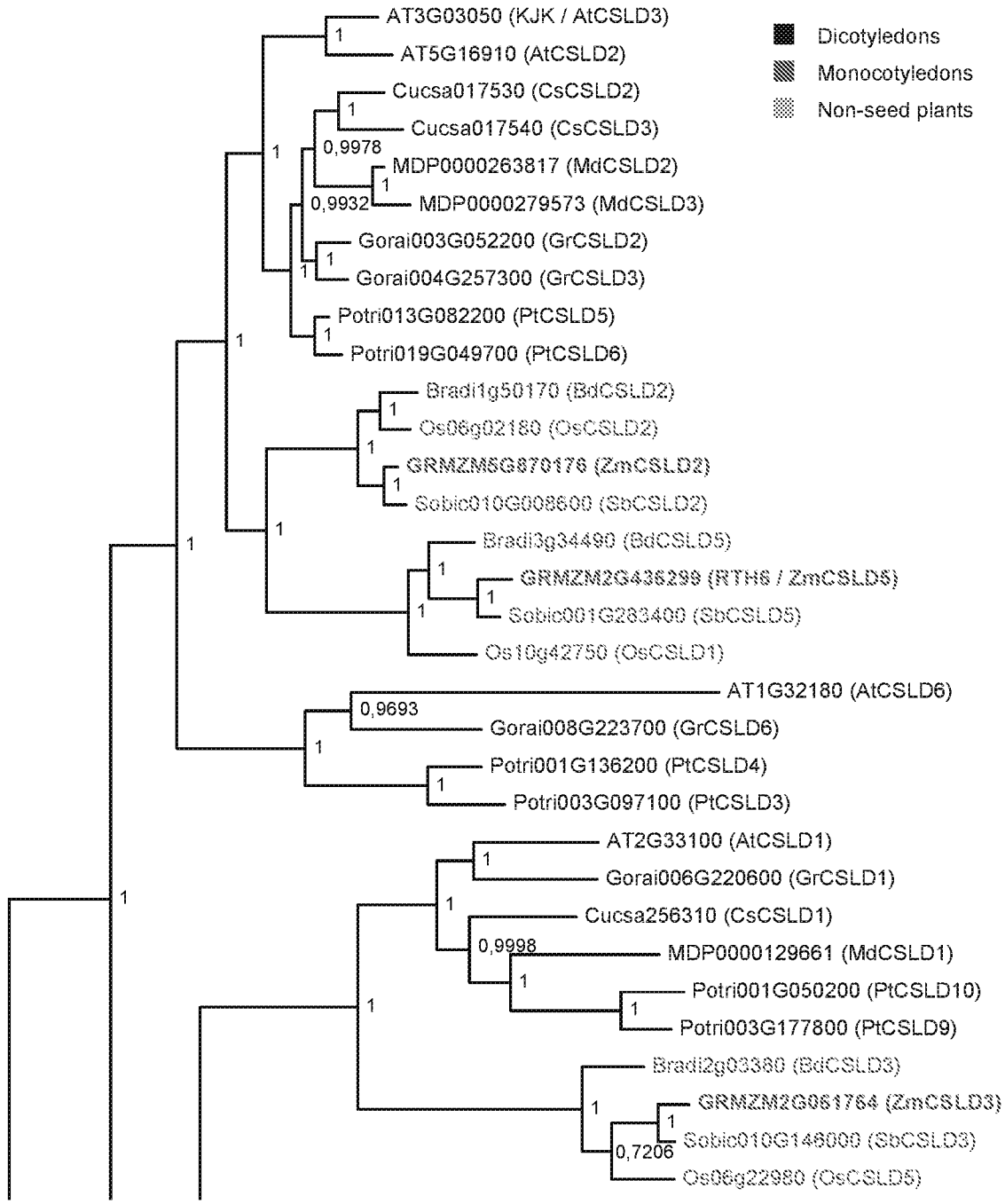

Dicotyledons
Monocotyledons
Non-seed plants
AT3G03050 (KJK / AtCSLD3)
AT5G16910 (AtCSLD2)
Cucsa017530 (CsCSLD2)
Cucsa017540 (CsCSLD3)
MDP0000263817 (MdCSLD2)
MDP0000279573 (MdCSLD3)
Gorai003G052200 (GrCSLD2)
Gorai004G257300 (GrCSLD3)
Potri013G082200 (PtCSLD5)
Potri019G049700 (PtCSLD6)
Bradi1g50170 (BdCSLD2)
Os06g02180 (OsCSLD2)
GRMZM5G870176 (ZmCSLD2)
Sobic010G008600 (SbCSLD2)
Bradi3g34490 (BdCSLD5)
GRMZM2G436299 (RTH6 / ZmCSLD5)
Sobic001G283400 (SbCSLD5)
Os10g42750 (OsCSLD1)
AT1G32180 (AtCSLD6)
Gorai008G223700 (GrCSLD6)
Potri001G136200 (PtCSLD4)
Potri003G097100 (PtCSLD3)
AT2G33100 (AtCSLD1)
Gorai006G220600 (GrCSLD1)
Cucsa256310 (CsCSLD1)
MDP0000129661 (MdCSLD1)
Potri001G050200 (PtCSLD10)
Potri003G177800 (PtCSLD9)
Bradi2g03380 (BdCSLD3)
GRMZM2G061764 (ZmCSLD3)
Sobic010G146000 (SbCSLD3)
Os06g22980 (OsCSLD5)
0,9978
0,9932
0,9693
0,9998
0,7206

FIG. 3

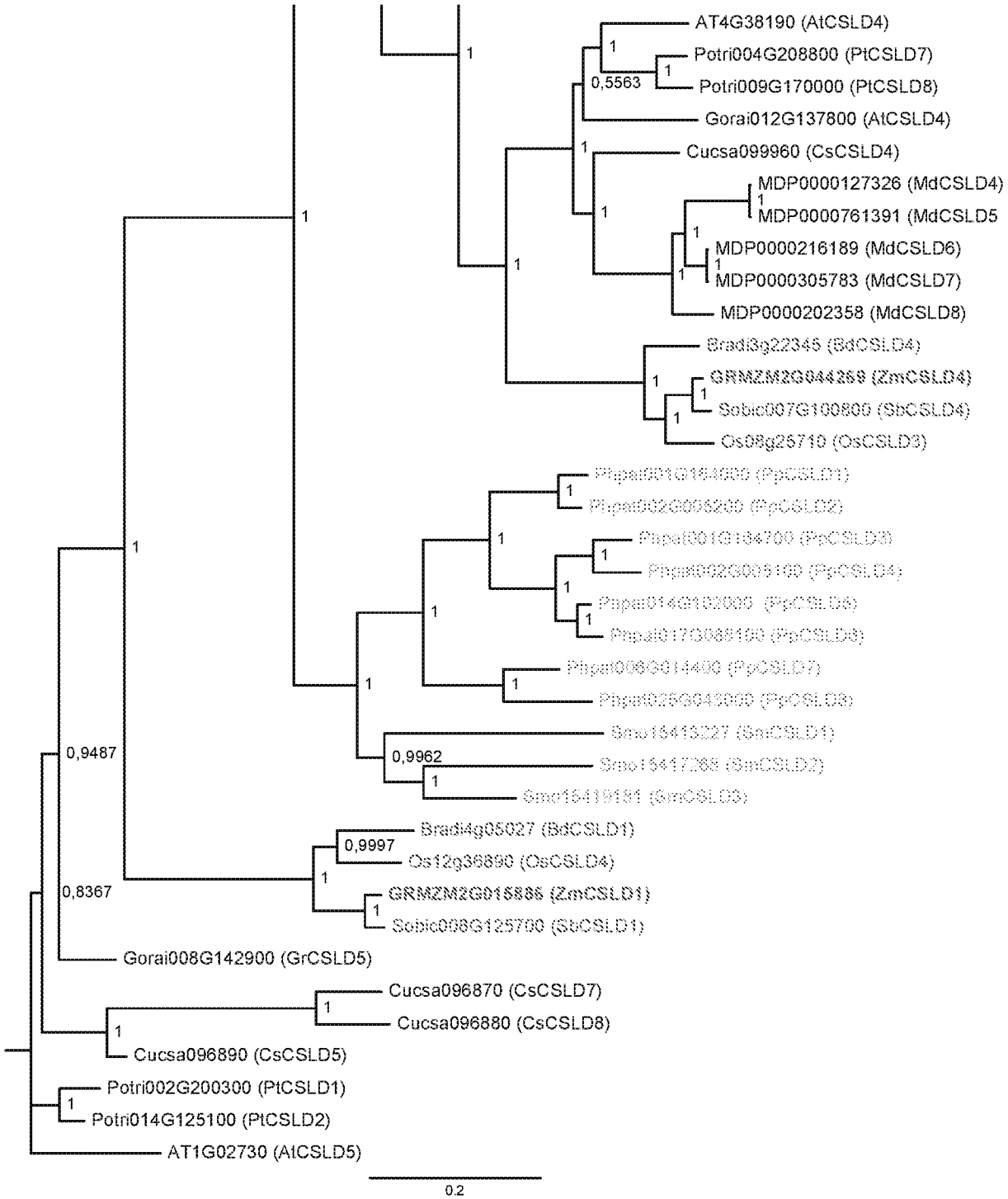


FIG. 3 CONTINUED

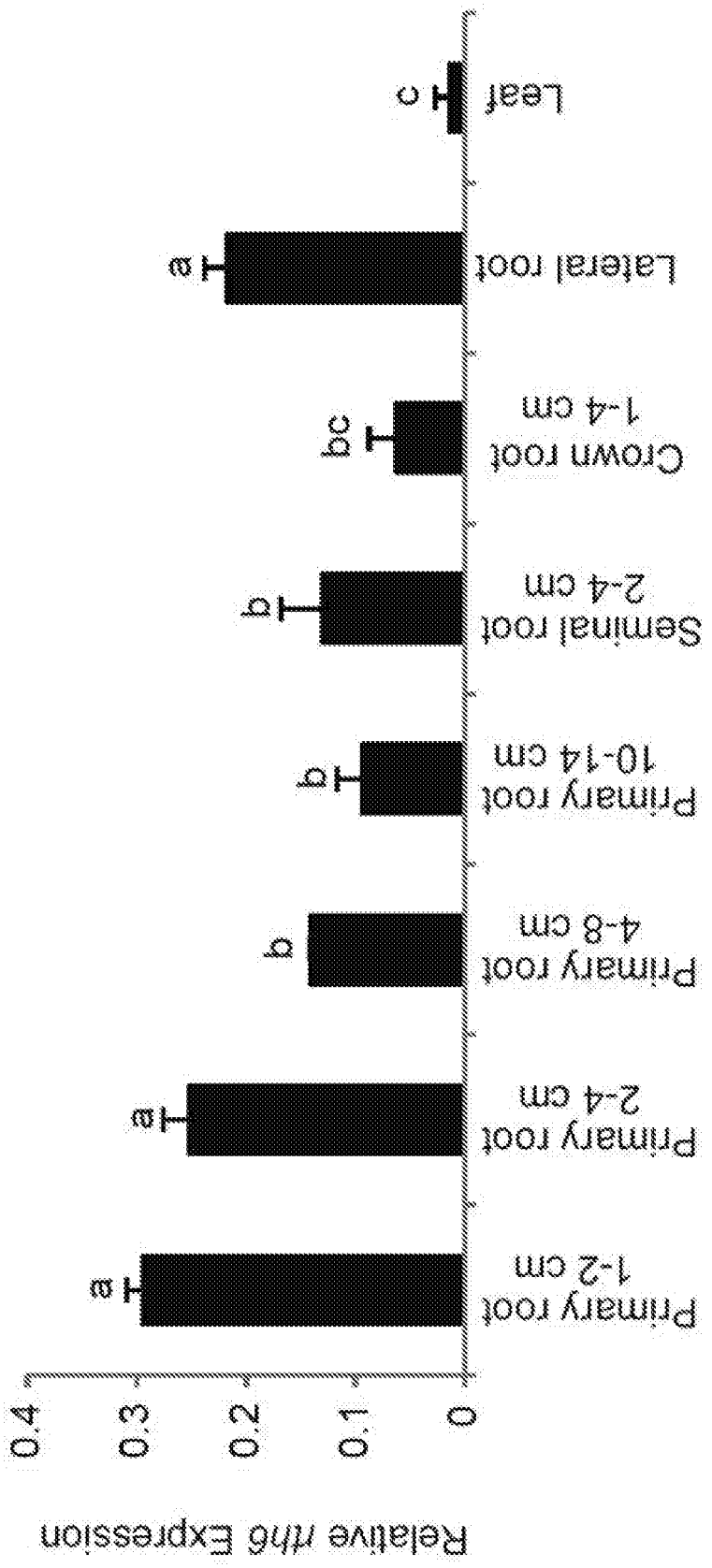

Relative rth6 Expression
0.4
0.3
0.2
0.1
0
Primary root 1-2 cm
a
Primary root 2-4 cm
a
Primary root 4-8 cm
b
Primary root 10-14 cm
b
Seminal root 2-4 cm
b
Crown root 1-4 cm
bc
Lateral root
a
Leaf
c

FIG. 4A

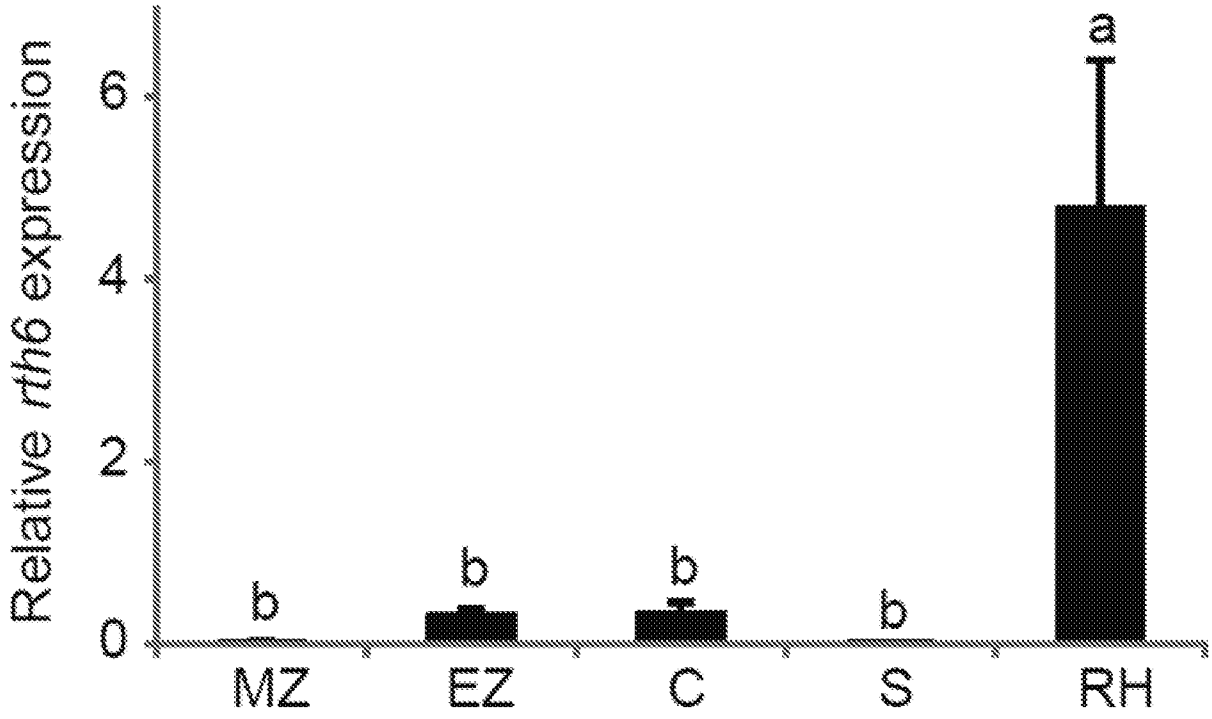

Relative rth6 expression
0
2
4
6
a
b
b
b
b
MZ
EZ
C
S
RH

FIG. 4B

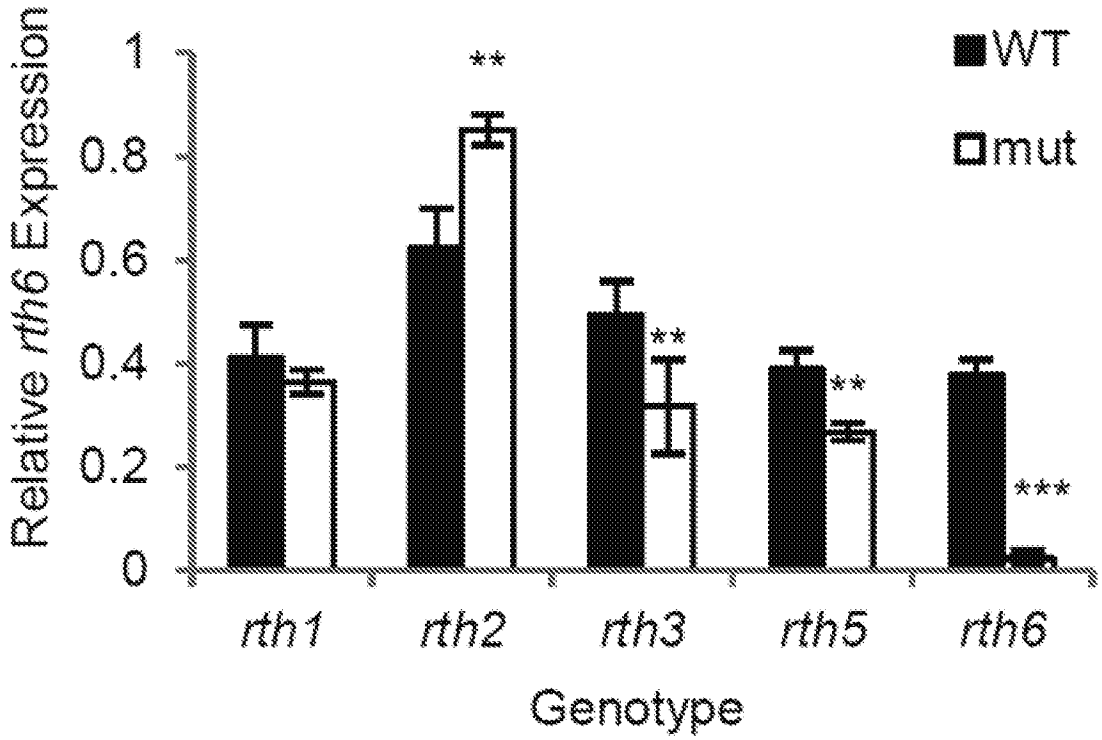

Relative rth6 Expression
0
0.2
0.4
0.6
0.8
1
WT
mut
**
**
**
***
rth1
rth2
rth3
rth5
rth6
Genotype

FIG. 4C

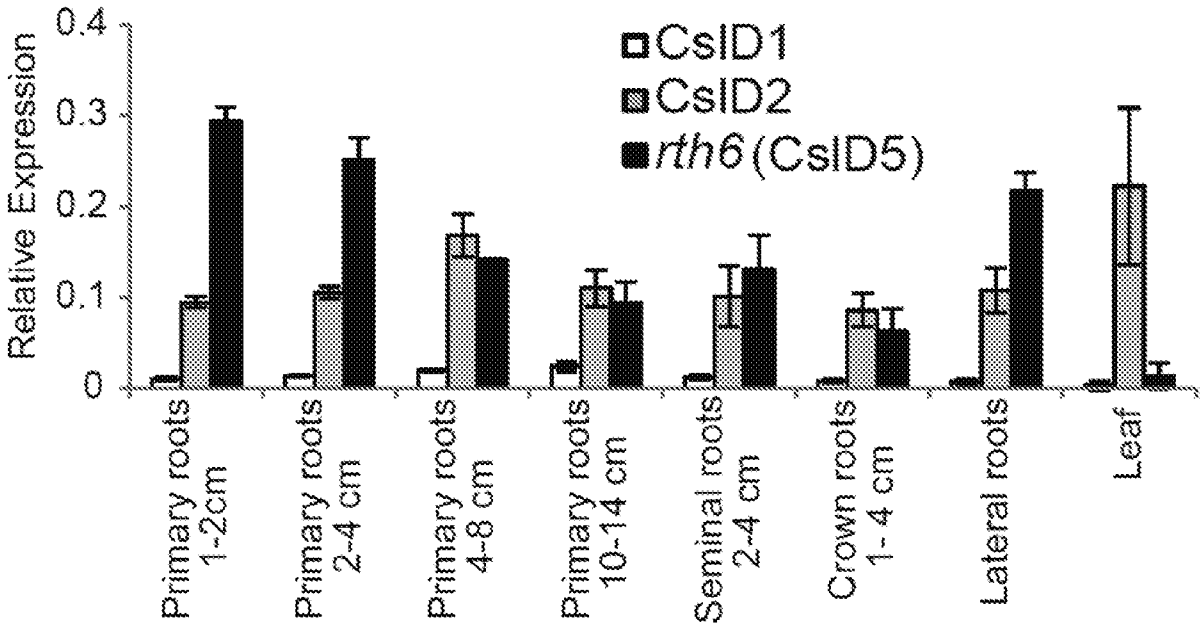

Relative Expression
0.4
0.3
0.2
0.1
0
CslD1
CslD2
rth6 (CslD5)
Primary roots 1-2cm
Primary roots 2-4 cm
Primary roots 4-8 cm
Primary roots 10-14 cm
Seminal roots 2-4 cm
Crown roots 1- 4 cm
Lateral roots
Leaf

FIG. 5A

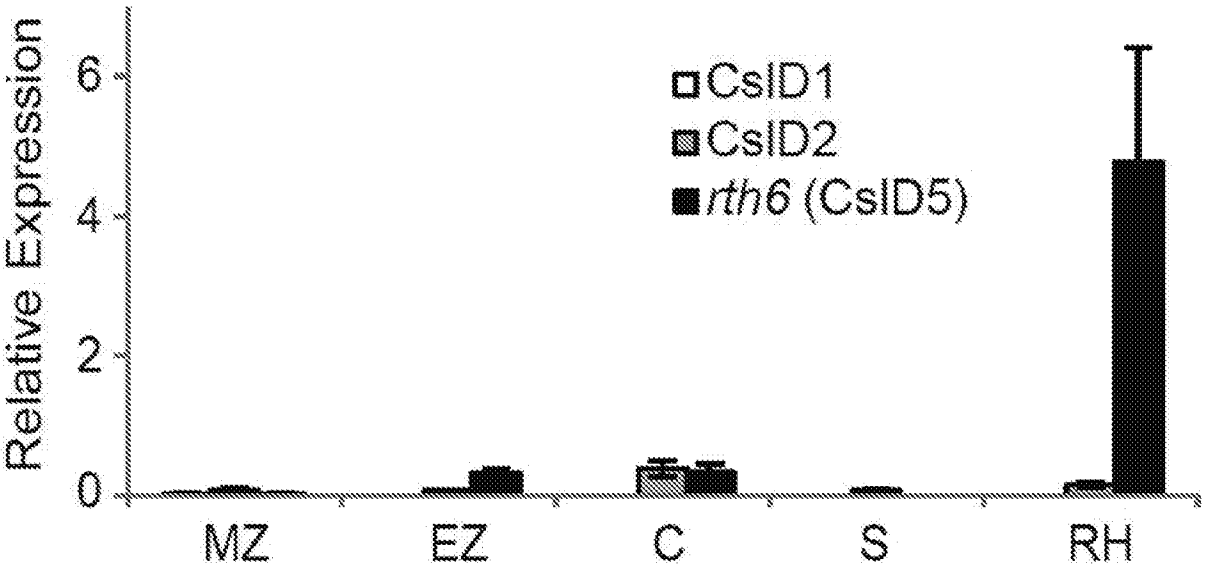

Relative Expression
6
4
2
0
CslD1
CslD2
rth6 (CslD5)
MZ
EZ
C
S
RH

FIG. 5B

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US16/55749

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8) - C07H 21/04; C12N 5/10, 15/82; C12 Q1/68 (2017.01)
CPC - C07H 21/04; C07K 14/415; C12N 5/10, 15/82, 15/8201, 15/8241, 15/8262; C12Q 1/68
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)
IPC(8): A01H 5/00; C07H 21/04; C12N 5/10, 15/82; C12 Q1/68 (2017.01)
CPC: A01H 5/00; C07H 21/04; C07K 14/415; C12N 5/10, 15/82, 15/8201, 15/8241, 15/8262; C12Q 1/68; C12Y 207/11001

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

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PatSeer (US, EP, WO, JP, DE, GB, CN, FR, KR, ES, AU, IN, CA, INPADOC Data); EBSCO Discovery; PubMed; Google Scholar; Patseer; The Lens; ENA; NCBI Blast; KEYWORDS: improve, agronomic, characteristic, plant, promoter, operably linked , polynucleotide

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2012/0110700 A1 (TARAMINO, G et al.) May 3, 2012; paragraphs [0020], [0021], [0029], [0134], [0148], [0188], [0239], [0250], [0255], [0261], [0297], [0302], [0318], [0330], [0353], [0357], [0390], [0448] ; claims 48, 49, 51, 62	1-2, 3/1-2, 13-16, 21-30, 31/21-30, 35-44, 47-49, 50/48-49, 51/48-49, 52/51/48-49, 72-75
Y	"NCBI Reference Sequence: XM_008660569.1" Publication [online]. 02 August 2014 [retrieved 01 February 2017]. Retrieved from the Internet: <URL: https://www.ncbi.nlm.nih.gov/nucleotide/670364419?report=genbank&log$=nucltop&blast_rank=1&RID=93UJU2SJ014>; pp 1-2	1-2, 3/1-2, 13-16, 21-30, 31/21-30, 35-44, 47-49, 50/48-49, 51/48-49, 52/51/48-49, 72-75 27, 31/27, 40
Y	US 2015/0159166 A1 (PIONEER HI BRED INTERNATIONAL INC.) June 11, 2015; paragraphs [0203], [0404]	47
Y	US 5,606,823 A (SOUZA, E et al.) March 4, 1997; abstract; column 4, lines 59-61; claim 1	48-49, 50/48-49, 51/48-49, 52/51/48-49, 72
Y	CN 102925404 A (UNIVERSITY NANJING) February 13, 2013; abstract; paragraph [0043]	75
Y	US 2015/0067922 A1 (THE PENN STATE RESEARCH FOUNDATION) March 5, 2015; abstract	

☒ 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
06 February 2017 (06.02.2017)	24 FEB 2017

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	Shane Thomas PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (January 2015)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US16/55749

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WEN, T. Characterization Of Three Root Hair Mutants In Maize. Iowa State University. 1998. [retrieved 12 December 2016]. Retrieved from the Internet: <URL: https://www.google.com/webhp?sourceid=chrome-instant&ion=1&espv=2&ie=UTF-8#q=wen+Characterization+of+three+root+hair+mutants+in+maize>; abstract	45-46
A	US 2011/0004955 A1 (ABAD, M et al.) January 6, 2011; abstract; paragraph [0010]	56-69, 70/56-69

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US16/55749

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.: 4-12, 17-20, 32-34
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:

-***-Please See Supplemental Page-***-

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, 3/1-2, 13-16, 21-30, 31/21-30, 35-49, 50/48-49, 51/48-49, 52/51/48-49, 56-69, 70/56-69, 71-75; SEQ ID NOs: 1-3

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
Information on patent family members

International application No.
PCT/US16/55749

-***-Continued from Box III -***-

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 examined, the appropriate additional examination fees must be paid.

Group I, Claims 1-3, 13-16, 21-31, 35-52 and 56-75 are directed toward a method for improving at least one agronomic characteristic in a plant, the method comprising introducing into the plant an expression construct, the construct comprising a promoter operably linked to a polynucleotide, wherein the polynucleotide comprises a nucleotide sequence selected from the group consisting of: (a) a nucleotide sequence comprising SEQ ID NO: 2; (b) a nucleotide sequence encoding the amino acid sequence set forth in SEQ ID NO: 3; an expression cassette comprising the polynucleotide and a plant comprising the expression construct, as well as methods for selecting al allelic variant of the gene in a maize plant, reducing soil erosion in a crop field, and increasing infection of a plant with a beneficial microorganism therewith.

Group II, Claims 53-55 are directed toward a method of increasing resistance of a plant to a root pathogen, the method comprising the step of reducing the expression of a rth6 gene family member or inhibiting the activity a rth6 polypeptide.

The inventions listed as Groups I and II 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: the special technical features of Group I include SEQ ID NO: 2, not present in Group II; the special technical features of Group II include increasing resistance of a plant to a root pathogen, not present in Group I.

Groups I and II share the technical features including: rth6.

However, these shared technical features are previously disclosed by the document 'Characterization of three root hair mutants in maize' by Wen (hereinafter 'Wen').

Wen discloses rth6 (rth6; abstract).

Since none of the special technical features of the Groups I and II inventions is found in more than one of the inventions, and since all of the shared technical features are previously disclosed by the Wen reference, unity of invention is lacking.