DWARFISM GENES AND DWARF PLANTS

Inventors: Richard M Amasino, Madison, WI (US); Fritz M Schomburg, Madison, WI (US); Scott D Michaels, Madison, WI (US); Colleen M. Bizzell, Middleton, WI (US)

Assignee: Wisconsin Alumni Research Foundation, Madison, WI (US)

Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 263 days.

Appl. No.: 10/155,435
Filed: May 23, 2002

References Cited
U.S. PATENT DOCUMENTS
5,939,539 A 8/1999 Lange et al.
6,198,021 B1 3/2001 Lange et al.

Other publications
Lin et al. (Jan., 2001, NCBI Accession No. AC079284).*

ABSTRACT
The present invention discloses the function, the cDNA sequences, and the expressed amino acid sequences of two genes the expression of which reduced bioactive GA levels and the height of a plant. This information enables creation of dwarf transgenic plants or transgenic plants with a specific dwarf organ.

10 Claims, 2 Drawing Sheets
SEQ ID NO: 4

| L10 | MASQPPFKTNFCSIFGSSFPNSTSD5N...TNSTIQTSLKLPVIDLS | 46 |
| L10 | MDPPFNEIYNNLYQNTNKENDVSEIEFSITAVVEEVEVLTVIDVS | 48 |

SEQ ID NO: 6

| L10 | HLTSGEEVKKKRCVQMVAKAKEWGFQIVNHGIPKVDHMELLEEKLF | 96 |
| L10 | RLIDGAEEREKCKEATAARSEWGFQIVNHGISMDVLEKMRQEQIRVF | 98 |

FIG 1
DWARFISM GENES AND DWARF PLANTS

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT

This invention was made with United States government support awarded by the following agency: DOE DE-FC05-92OR22072. The United States has certain rights in this invention.

CROSS-REFERENCE TO RELATED APPLICATIONS

Not applicable.

BACKGROUND OF THE INVENTION

Gibberellins (GAs) are a group of tetracyclic diterpene carboxylic acids involved in a variety of developmental processes. They were originally identified through their effect on stem elongation (Phillips, A. L., Plant Physiol. Biochem 36: 115–124, 1998), and are now implicated in all stages of the plant life cycle including seed germination, leaf expansion, floral induction, fruit maturation, and apical dominance (Harberd, N. P. et al., BioEssays 20: 1001–1008, 1998). There are at least 126 different GAs identified in plants, fungi, and bacteria; however, most are precursors or degradation products, which are inactive forms. The bioactive GAs in higher plants include GA1, GA3, GA4, and GA7 (Hedden, P. and A. L. Phillips, Trends Plant Sci. 5: 523–530, 2000).

The GA biosynthetic pathway has three different classes of enzymes that catalyze specific reactions in the synthesis of bioactive GAs: terpene cyclases, Cyt P450 monoxygenases, and 2-oxoglutarate-dependent dioxygenases. The specific enzymatic steps for the synthesis of bioactive GAs from GA12 are species specific. The last reactions producing bioactive GAs and the first breakdown reactions involve several types of dioxygenases. The nomenclature of these dioxygenases is variable throughout the literature. Herein, the most commonly used name is listed first, followed by any other names also used. GA 20-oxidases remove the C-20, whereas 3β-hydroxylases (also called 3-oxidases) introduce the 3β-hydroxyl group; both are steps on the way to bioactive GAs. GA 2-oxidases (also called 2β-hydroxylases) introduce a 2β-hydroxyl group resulting in inactive products that cannot be converted to active forms (Thomas, S. G. et al., Proc. Natl. Acad. Sci. USA 96: 4698–4703, 1999). GA 2-oxidases generally act on GAs with 19 carbons, although there is evidence of 2β-hydroxylation of C20-GAs (Hedden, P. and Y. Kamiya, Annu. Rev. Plant Physiol. Plant Mol. Biol. 48: 431–60, 1997).

GA-modifying enzymes produce a vast number of GAs, although most are precursors or inactive forms. Many dioxygenases have been shown to be multifunctional, catalyzing consecutive reactions in the pathway, or modifying different, but structurally similar, GAs. For example, GA5, a GA 20-oxidase, converts GA12 to GA15 to GA24 to GA9, and GA53 to GA20 (Yamaguchi, S. and Y. Kamiya, Plant Cell Physiol. 41: 251–257, 2000). This multifunctional property allows many different GAs to be formed from relatively few enzymes.

Several of the dioxygenases can be grouped into small gene families. In Arabidopsis, GA 20-oxidases and GA 3β-hydroxylases are each encoded by at least four genes, and GA 2-oxidases are claimed in one review to be encoded by at least six genes (Hedden, P. and A. L. Phillips, Trends Plant Sci. 5: 523–530, 2000). Although the three groups of dioxygenases act on similar GA substrates, cluster analysis shows that they are no more closely related to each other than to any other plant dioxygenase (Hedden, P. and A. L. Phillips, Trends Plant Sci. 5: 523–530, 2000). The identity between different groups of GA dioxygenases is approximately 20–30% within one species, such as Arabidopsis (Table 1). Within a group, however, the identity is higher, even among species. Arabidopsis GA 20-oxidases are approximately 55–70% identical to each other, and 50–60% identical to 20-oxidases of other species (Prescott, A. G. and P. John, Annu. Rev. Plant Physiol. Plant Mol. Biol. 47: 245–71, 1996). The three published Arabidopsis GA 2-oxidases are 49–68% identical to each other (Thomas, S. G. et al., Proc. Natl. Acad. Sci. USA 96: 4698–4703, 1999), and 35–65% identical to GA 2-oxidases of other species (Table 2). The various members of each dioxygenase family are differentially expressed within the plant, and may be involved in different developmental processes (Hedden, P. and A. L. Phillips, Trends Plant Sci. 5: 523–530, 2000).

Chemical modification of GA levels is common in agriculture and horticulture. Seedless grapes are often treated with GA3 to increase berry size. Conversely, many crops and ornamental plants are treated with chemicals that act to inhibit enzymes in the GA-biosynthetic pathway (Hedden, P. and A. L. Phillips, Trends Plant Sci. 5: 523–530, 2000). Height reduction in ornamentals is currently achieved in many plants, such as poinsettias and petunias, via treatment with GA-inhibiting chemicals to produce compact plants that are more desired by consumers. Height reduction in a number of crop plants has resulted in increased yields and yield stability. In fact, compact crop plants have been a cornerstone of the great enhancements in agriculture yields over the past three decades. Compact plants can be grown more densely and are more resistant to storm damage (lodging) than taller wild type varieties. Compact plants are easier to harvest because they hold the seed products closer together, reducing loss during harvesting.

Many groups have manipulated GA levels by transgenically altering the expression of genes involved in GA metabolism. Overexpression of GA 20-oxidases in Arabidopsis has yielded plants with decreased GA levels which results in plants that are taller and have lighter green leaves than wild-type plants (Huang, S. et al., Plant Physiol. 118: 773–781, 1998). Suppression of GA 20-oxidases by antisense RNA has produced Arabidopsis plants that display phenotypes similar to weak GA-deficient plants; these plants have darker green cotyledons, were about 40% shorter than wild-type plants at maturity, and flowered slightly later than wild type in short-day conditions (Coles, J. P. et al., Plant J. 17: 547–556, 1999). Overexpression of a unique pumpkin 20-oxidase, which produces an inactive GA, has produced plants with a weak GA-deficient phenotype in Solanum dulcamara. These plants are semi-dwarfs, have smaller, darker green leaves, flower earlier, and produce more fruit and seed per fruit than wild type plants (Curtis, I. S. et al., Plant J. 23: 329–338, 2000). Overexpression of a bean 2-oxidase in Arabidopsis has produced plants with a variety
various host cells containing the nucleic acid molecules. The usage in plant production, as well as decreasing energy and time expenditures in chemical applications.

**BRIEF SUMMARY OF THE INVENTION**

The present invention discloses the function, the cDNA sequences, and the expressed amino acid sequences of two genes, the expression of which reduced bioactive GA levels and the height of a plant. The present invention includes various nucleic acid molecules and polypeptides that are related to the two genes and useful in various applications such as detecting the genes, generating antibodies and generating dwarf plants. The present invention also includes various host cells containing the nucleic acid molecules. The present invention also includes methods of generating dwarf plants using the nucleic acid molecules and the polypeptides described above and the resulted dwarf plants themselves.

It is an object of the present invention to provide a tool to creators of new plant varieties to alter the height of a plant or the size of a specific plant organ.

It is an advantage of the present invention that the two genes are dominant with regard to the dwarf phenotype so that a dwarf transgenic plant is easy to create.

Other objects, advantages and features of the present invention will become apparent from the following specifications and claims.

**BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWINGS**

FIG. 1 shows alignment of L10 and 69 proteins.

FIG. 2 shows alignment of other GA-modifying enzymes to the L10 and 69 proteins.

**DETAILED DESCRIPTION OF THE INVENTION**

Set forth below are the cDNA (SEQ ID NO:3 and SEQ ID NO:5) and the deduced amino acid (SEQ ID NO:4 and SEQ ID NO:6) sequences of two plant dwarfism genes, here named L10 and 69, respectively. The names L10 and 69 are for identification purpose only and may be changed to other names (for example, GA_ox_ , wherein “_” is a number) to reflect the function of these two genes when the sequences are submitted for publication. When over expressed in a plant, each of these two genes reduced the bioactive GA level and the height of the plant. Prior to the present invention, the genomic DNA sequences (SEQ ID NO:1 and SEQ ID NO:2), but not the cDNA sequences, the amino acid sequences and the function, of these two genes were known. The present invention provides plant breeders and creators a unique tool so as to sculpt the height of a plant to more closely follow the desires of the breeder.

As shown in the examples below, overexpression of either L10 or 69 cause GA-deficiency indicating they are involved in GA degradation, not biosynthesis. L10 and 69 proteins have 44% identity and 54.5% similarity to each other. Both are listed in the database as giberellin 20-oxidase-like proteins. GA 20-oxidases, however, are involved in biosynthesis, not degradation, although there is one report of a unique pumpkin 20-oxidase whose activity results in an inactive product and causes a dwarf phenotype when over-expressed in certain species (but not in Arabidopsis) (Curtis, I. S. et al., Plant J. 23: 329–338, 2000).

By sequence analysis, L10 and 69 do not fit well into any of the three groups of dioxygenases (Tables 1 and 2). In a BLAST search, the GA 20-oxidases from a variety of species show up before any 3β-hydroxylases or 2-oxidases; however, there is only a 28–33% identity between the 20-oxidases and our novel dioxygenases. There is a 24–30% identity between the novel dioxygenases and 3β-hydroxylases or 2-oxidases (Table 1 and 2) from various species. Thus, L10 and 69 dioxygenases do not seem to be significantly more similar to 20-oxidases than to the other dioxygenases (Table 1 and FIG. 2), and their overexpression phenotypes indicate that they are not 20-oxidases or 3β-hydroxylases. 20-oxidases and 3β-hydroxylases are biosynthetic enzymes and their overexpression should therefore lead to taller plants, but overexpression of either L10 or 69 leads to dwarf plants. Thus, if the L10 and 69 dioxygenases are part of a currently recognized class, based upon the overexpression of dwarf phenotype, it is more likely that they are 2-oxidases than either 3β-hydroxylases or 20-oxidases. A complete comparison of the amino acid sequences of all cloned 2-oxidases are shown in Table 2. The unique 20-oxidase from pumpkin is also included. As can be seen in Tables 1 and 2, the L10 and 69 dioxygenases are not as similar to the 2-oxidases as the rest of the 2-oxidases are to each other, even between species. The L10 and 69 dioxygenases are no more similar to the 2-oxidases than the 2-oxidases are to the 20-oxidases or 3β-hydroxylases (Table 1 and 2). Thus, the L10 and 69 dioxygenases are either a new class of dioxygenases or a unique, more distant subgroup of an existing class of dioxygenases.

In one aspect, the present invention relates to a polypeptide including an amino acid sequence that has at least 35%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, or 99% identity to and over the entire length of that of SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:4 with conservative substitutions, or SEQ ID NO:6 with conservative substitutions. The present invention also relates to a polypeptide including a novel fragment of the amino acid sequence described above, especially a fragment that is immunogenic or has a biological activity of reducing the bioactive GA level or the height of a plant. Besides the amino acid sequence described above, the polypeptide of the present invention can include a native or non-native amino acid sequence at the N- or C-terminus or both, which will not interfere with the function of the amino acid sequence described above. The flanking native or non-native amino acid sequence can but does not have to be one that assists in purification, detection, or stabilization of the amino acid sequence described above.

As used herein, “percent identity” of the two amino acid sequences or of two nucleic acids is synonymous to “percent homology,” which is determined using the algorithm of Karlin and Altschul (Proc. Natl. Acad. Sci. USA 87:2264–2268, 1990), modified as in Karlin and Altschul (Proc. Natl. Acad. Sci. USA 90:5873–5877, 1993). Such an algorithm is incorporated into the NBLAST and XBLAST programs of Altschul et al. (J. Mol. Biol. 215:403–410, 1990). BLAST nucleotide searches are performed with the NBLAST program, score=100, wordlength=12, to obtain nucleotide sequences homologous to a nucleic acid molecule.
of the invention. BLAST protein searches are performed with the XBLAST program, score=50, wordlength=3, to obtain amino acid sequences homologous to a reference polypeptide (e.g., SEQ ID NO:4 or SEQ ID NO:6). To obtain gapped alignments for comparison purposes, Gapped BLAST is utilized as described in Altschul et al. (Nucleic Acids Res. 25:3389–3402, 1997). When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) are used. The referenced programs are available on line from the National Center for Biotechnology Information, National Library of Medicine, National Institute of Health.

Also within the scope of the present invention are polypeptides that bind specifically to an antibody that binds specifically to protein L10 or 69.

In another aspect, the present invention relates isolated nucleic acid molecules as described below. An “isolated nucleic acid molecule” used herein is a nucleic acid sequence, the structure of which is not identical to that of any naturally occurring nucleic acid or to that of any fragment of a naturally occurring genomic nucleic acid spanning more than three separate genes. The term therefore covers, for example, (a) a DNA which has the sequence of part of a naturally occurring genomic DNA molecules but is not flanked by both of the coding sequences that flank that part of the molecule in the genome of the organism in which it naturally occurs; (b) a nucleic acid incorporated into a vector or into the genomic DNA of a prokaryote or eukaryote in a manner such that the resulting molecule is not identical to any naturally occurring vector or genomic DNA; (c) a separate molecule such as a cDNA, a genomic fragment, a fragment produced by polymerase chain reaction (PCR), or a restriction fragment, and (d) a recombinant nucleotide sequence that is part of a hybrid gene, i.e., a gene encoding a fusion protein. Specifically excluded from this definition are nucleic acids present in mixtures of (i) DNA molecules, (ii) transfected cells, and (iii) cell clones, e.g., as these occur in a DNA library such as a cDNA or genomic DNA library. An isolated nucleic acid molecule can be modified or unmodified DNA or RNA, whether fully or partially single-stranded or double-stranded or even triple-stranded. A modified nucleic acid molecule can be chemically or enzymatically induced and can include so-called non-standard bases such as inosine.

An isolated nucleic acid molecule of the present invention is one that includes a polynucleotide having an uninterrupted coding sequence that encodes a polypeptide the amino acid sequence of which is at least 35%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 97%, or 99% identical to SEQ ID NO:4 or SEQ ID NO:6, a complement of the foregoing, or a novel fragment of any of the foregoing. A preferred nucleic acid molecule includes a polynucleotide having a sequence that is at least 35%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 97%, or 99% identical to SEQ ID NO:4 or SEQ ID NO:6; a complement of the foregoing, or a novel fragment of any of the foregoing. A preferred nucleic acid molecule includes a polynucleotide having a sequence that is at least 35%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 97%, or 99% identical to SEQ ID NO:5.

The invention also includes nucleic acid molecules that hybridize under stringent hybridization conditions (as defined herein) to all or a portion of the nucleotide sequence represented by SEQ ID NO:3 or its complement, or SEQ ID NO:5 or its complement. The hybridizing portion of the hybridizing nucleic acid molecules is typically at least 15 (e.g., 20, 25, 30, or 50) nucleotides in length. The hybridizing portion of the hybridizing nucleic acid molecules is at least 80%, e.g., at least 95%, or at least 99%, identical to the sequence of a portion or all of a nucleic acid encoding a L10 or 69 polypeptide, or the sequence’s complement. Hybridizing nucleic acid molecules of the type described herein can be used, for example, as a cloning probe, a primer (e.g., a PCR primer), or a diagnostic probe. Hybridization of the oligonucleotide probe to a nucleic acid sample typically is performed under stringent conditions. Nucleic acid duplex or hybrid stability is expressed as the melting temperature or Tm, which is the temperature at which a probe dissociates from a target DNA. This melting temperature is used to define the required stringency conditions. If sequences are to be identified that are related and substantially identical to the probe, rather than identical, then it is useful to first establish the lowest temperature at which only homologous hybridization occurs with a particular concentration of salt (e.g., SSC or SSPE).

Then, assuming that 1% mismatching results in a 1° C. decrease in the Tm, the temperature of the final wash in the hybridization reaction is reduced accordingly (for example, if sequences having >95% identity with the probe are sought, the final wash temperature is decreased by 5° C.). In practice, the change in Tm can be between 0.5° C. and 1.5° C. per 1% mismatch. Stringent conditions involve hybridizing at 68° C. in 5xSSC/5x Denhardt’s solution/1.0% SDS, and washing in 0.2xSSC/0.1% SDS at room temperature. Moderately stringent conditions include washing in 3xSSC at 42° C. The parameters of salt concentration and temperature can be varied to achieve the optimal level of identity between the probe and the target nucleic acid. Additional guidance regarding such conditions is readily available in the art, for example, by Sambrook et al., 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, N.Y.; and Ausubel et al. (eds.), 1995, Current Protocols in Molecular Biology, (John Wiley & Sons, N.Y.) at Unit 2.10.

Isolated nucleic acid molecules of the invention can be obtained by several methods. For example, they can be isolated using procedures which are well known in the art. These include, but are not limited to: (a) hybridization of detectably labeled probes representing all or part of the L10 or 69 gene to genomic or cDNA libraries to detect similar nucleic acid sequences; (b) antibody screening of expression libraries to detect similar structural features; (c) synthesis by the polymerase chain reaction (PCR); and (d) chemical synthesis of a nucleic acid molecule. Sequences for specific coding regions of genes can also be found in GenBank, the National Institutes of Health computer database.

For the identification of isolated nucleic acid molecules using detectably labeled probes, or for the identification of polynucleotide fragments whose complements hybridize to L10 or 69, stringent hybridizing conditions described above can be used. Alternatively, higher stringency conditions can be used. Typically, lower stringency hybridization conditions permit hybridization of related but not identical L10 or 69 gene, and thereby allow identification of the L10 or 69 gene in other species.

In a related aspect, any polynucleotide sequence of the present invention, or an antiseren version thereof, can be provided in a vector or genetic construct in a manner known to those skilled in the art. A polypeptide-encoding polynucleotide so provided in a vector can, but need not, be under the transcriptional control of one or more regulatory elements which can include a promoter not natively found adjacent to the polynucleotide such that the encoded polypeptide can be produced when the vector is provided in a compatible host cell or in a cell-free transcription and translation system. Such cell-based and cell-free systems are well known to the skilled artisan. Cells comprising a vector containing a polynucleotide sequence of the invention are themselves within the scope of the invention.

In another related aspect, the present invention encompasses a polynucleotide having a nucleotide sequence that encodes
a polypeptide the amino acid sequence of which is at least 35%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 97%, or 99% identical to SEQ ID NO:4 or SEQ ID NO:6, operably linked to a non-native expression control sequence which can include a promoter. Such a polynucleotide of the present invention can be provided in a vector such that the encoded polypeptide can be produced when the vector is provided in a compatible host cell or in a cell-free transcription and translation system. Such cell-based and cell-free systems are well known to the skilled artisan. Cells comprising the vector are themselves within the scope of the invention.

In yet another aspect, the present invention relates to a method of reducing the height of a plant and the resulting dwarf plant. One way to reduce the height of a plant is to increase the transcription or translation rate, or the stability of the mRNA or protein products of the endogenous L10 or 69 gene. Another way to reduce the height of a plant is to make a transgenic plant to express certain isolated nucleic acid molecules of the present invention, which include for example, the L10 or 69 gene of the same or a different species (either the genomic DNA or cDNA of the L10 or 69 gene), a portion of a L10 or 69 gene the protein product of which retains the function of reducing bioactive GA level, and other nucleic acid molecules of the present invention that are effective when expressed in the transgenic plant to cause the transgenic plant to be shorter compared to a non-transgenic plant of the same genetic background.

The examples below showed that expressing the Arabidopsis L10 or 69 gene introduced into a Arabidopsis plant or the Arabidopsis L10 gene introduced into a tobacco plant reduced the height of the Arabidopsis or tobacco plant. Identical or similar techniques can be used to express a L10 or 69 gene in other plants species to reduce the height of those species. In addition, this Arabidopsis or tobacco plant system can be used to test possible L10 or 69 genes from other plant species and those nucleic acid molecules of the present invention that are effective to cause a transgenic plant to be shorter compared to a non-transgenic plant of the same genetic background.

It should be understood that techniques of plant genetic engineering have been developed to the point where it is now practical to place any genetic construct into almost any useful plant species. The process does, however, still involve some random processes, most notably that insertions of foreign DNA into the genome of plants still occurs at random sites in the plant genome. As a result, in any group of plants emerging from a plant transformation process, the results achieved for the different gene insertion events will vary, sometimes dramatically. For example, for a simple gene insertion of another copy of an endogenous plant gene, many plants produced will have a slightly higher level of activity of the endogenous protein, others will have no measurable change or even a decrease in measurable activity, while a few will have substantial increases in activity levels. However, this variation does not mean stable results cannot be achieved, since the results tend to be consistent generation-to-generation for each specific genetic insertion. Thus the high activity plants have, in effect, a high activity allele that can be transferred by normal mendelian inheritance to their progeny.

To make a transgenic plant, as is known to those of skill in the art, one needs to make a genetic construction capable of expressing an inserted protein coding sequence, whether foreign or endogenous, in a plant. One also needs a method to insert the genetic construction into the plant.

The tools and techniques for making genetic constructions that will express proteins in plants are now widely known.

Any genetic construction intended to cause the synthesis in the cells of the plant of a polypeptide or protein must include a sequence of DNA known as a protein coding sequence (can be a genomic DNA or a cDNA), which specifies the sequence of the polypeptide or protein to be produced in the resultant plant. For a protein coding sequence to be expressed in a plant to produce a polypeptide or protein, it must be placed under the control of a plant expressible promoter and be followed by a plant transcriptional terminator sequence, also known as a polyadenylation sequence. The plant expressible promoter is a promoter which will work in plants, usually either of plant origin or from a plant pathogen like a virus (e.g. Cauliflower mosaic virus) or a bacteria (e.g. Agrobacterium promoters like the nopalin synthase promoter). Plant promoters from pathogen tend to be constitutive promoters, meaning that they actually express the protein coding sequence in all of the tissues of the plant at all times. Other plant promoters are known to be tissue specific (e.g. to fruit or to flower) or developmentally specific (e.g. to stage of plant life such as emergent specific or senescent specific), while others are intended to be inducible (e.g. heat shock or metal ion induced promoters). Any of these types of promoters may by used in the practice of this invention depending on the intended affect on the transgenic plant to be produced. For example, a plant with a specific height or stature may be obtained through adjusting the expression level of a transgene by varying promoter strength. One may also use a tissue specific promoter to limit the dwarfing effect such as changing inflorescence architecture, stem elongation, or fruit development without changing any other aspect of the plant.

Several methods have been demonstrated to insert genes into plants to make them transgenic. The most widely used methods, broadly defined, are Agrobacterium-mediated transformation or accelerated particle mediated transformation. The various techniques of Agrobacterium-mediated plant transformation make use of the natural ability of the plant pathogens of the Agrobacterium genus to transfer DNA from a plasmid in the bacteria into the genome of a plant cell. Particle-mediated plant transformation techniques utilize DNA-coated small carrier particles accelerated from a device, often referred to as a gene gun, into the cells of a plant. The full implementation of either approach requires techniques to recover a fully mature, morphologically normal plant from the transformed cells. The techniques often therefore involve either selection or screening protocols to identify which plant cells were transformed and regeneration protocols to recover whole plants from the single transformed plants cells. As mentioned above, these techniques have been worked out for many plant species and many, and perhaps all, of the economically important plant species.

Other techniques, such as electroporation have also been used to make transgenic plants. But fundamentally for the invention disclosed here, the particular technique of plant transformation does not matter. Once the plant has been genetically engineered, and a transgenic plant has been created, the method of transformation of the original plant becomes irrelevant. A transgene inserted into the genome of one plant is then fully inheritable by progeny plants of the original genetically engineered plant by normal rules of classical plant breeding. The term transgene is here used to apply to an inserted genetic construction carried in the cells of a target plant. Thus, the term transgenic plant, as used here, refers to a plant that carries such a transgene.

Plants in which a copy of a L10 or 69 gene is introduced may also contain a wild-type (i.e., endogenous) plant height coding region which acts to control the height of the plant.
Upon introduction into the genome of a plant, the L10 or 69 gene can act to augment the activity of an endogenous height coding region to make the plant shorter. For instance, a second copy of a height coding region can be introduced into a plant to increase the amount of height reduction L10 or 69 protein present in the plant.

The present invention also provides a genetically modified plant, characterized as having the phenotypic trait of general dwarfing of the whole plant or dwarfing of a specific plant organ. By this it is meant that the modified plants of the present invention, whether modified by incorporating a L10 or 69 gene expressing a new or additional L10 or 69 protein in the plant, demonstrate a reduced height or size in at least one tissue or organ relative to the same plant without the transgene inserted. Preferably, the dwarfing of the whole transgenic plant or a specific tissue or organ (on average) of the transgenic plant is at least about 20%, more preferably at least about 100%, most preferably at least about 200% in comparison to the same plant without the transgene. Preferably, the genetically modified plant and the same plant without the transgene are grown under the same conditions.

Plants included in the invention are any plants amenable to transformation techniques, including both monocotyledonous and dicotyledonous plants. Examples of monocotyledonous plants include, but are not limited to, vegetables such as asparagus, onions and garlic; cereals such as maize, barley, wheat, rice, sorghum, pearl millet, rye and oats; and grasses such as forage grasses and turfgrasses. Examples of dicotyledonous plants include, but are not limited to, vegetables, feed, and oil crops such as tomato, beans, potatoes, soybeans, peppers, lettuce, peas, alfalfa, clover, Brassica species (e.g., cabbage, broccoli, cauliflower, brussel sprouts, rapeseed, and radish), carrot, beets, eggplant, spinach, cucumber, squash, melons, cantaloupe, sunflowers; fiber crops such as cotton; and various ornamentals such as flowers and shrubs.

In another related aspect, the isolated nucleic acid molecules of the present invention can be used to analyze and determine the pattern of L10 or 69 gene activity in a transgenic or non-transgenic plant as an aid to breeding or creating plants having desired heights.

The above disclosure generally describes the present invention. A more complete understanding can be obtained by reference to the following specific examples which are provided herein for purposes of illustration only and are not intended to limit the scope of the invention.

EXAMPLES

1. Isolation of Mutants that Possessed Altered GA-Metabolism.

Arabidopsis was mutagenized by T-DNA derived from Agrobacterium tumefaciens to generate plants with altered phenotypes. The mutagenesis was designed to isolate plants that contained dominant phenotypes by use of a T-DNA vector that contained a transcriptional enhancer region (PSK115 activation vector) (Weigel, D. et al., Plant Physiol. 122: 1003–1013, 2000). Plants were screened in the initial (T1) generation to identify mutant plants that displayed altered inflorescence architecture.

Two mutant plants (Line 10 (L10) and Line 69, (69)) were identified that displayed smaller and darker green leaves, delayed floral induction, and reduced primary inflorescence length (Table 3). These phenotypes are similar to loss-of-function alleles of GA biosynthetic enzymes. When GA levels were directly measured it was found that bioactive GAs in the mutants L10 and 69 were substantially lower than wild-type plants. Bioactive GA (GA4) was measured in wild-type plants at 1.84 ng/g dry weight; whereas, both L10 and 69 plants did not contain detectable levels of GA4. Additionally, other GA forms such as GA3 were also undetectable in the L10 and 69 lines (Table 4). Both mutant phenotypes could be rescued by the external application of bioactive GAs, which is consistent with the notion that the phenotypes were a result of reduced levels of bioactive GAs.

Heterozygous plants of L10 and 69 were self-pollinated to create a segregating population. A 3:1 ratio of mutant to wild-type plants was observed in the segregating population of both L10 and 69 indicating that both phenotypes are dominant. There are no detectable differences between heterozygous and homozygous phenotypes for either mutant and thus each appears to behave in a fully dominant manner.

The phenotype of the mutants cosegregated with the T-DNA present in the mutants. More than 100 plants from the segregating population were assayed for both the altered GA phenotype and the presence of the T-DNA. All of the plants that displayed the GA-deficient phenotype contained the T-DNA while none of the wild-type plants contained the T-DNA. The probability of this cosegregation occurring randomly is less than 0.00001 and therefore indicated that the T-DNA caused the mutant phenotype seen in the L10 and 69 lines.

2. Cloning of Genes that Caused the L10 and 69 Mutant Phenotypes.

Since the T-DNA cosegregated with the mutant phenotype, it was possible to sequence the genomic DNA near the T-DNA to determine where in the genome the T-DNA was located. A piece of the genomic DNA near the T-DNA from both mutants was sequenced and used to search the Arabidopsis data base (which includes sequences from many organisms) available on line from the National Center for Biotechnology Information, National Library of Medicine, National Institute of Health. The search revealed that genomic DNA sequences corresponding to predicted genes were directly adjacent to the enhancer region of the inserted T-DNA in both L10 and 69 plants. These genes were designated L10 and 69, respectively. In the Arabidopsis data base, 69 is on BAC F7J7 (accession number AL021960: Arabidopsis thaliana DNA chromosome 4, BAC clone F7J7 (ESSA project)). It is number 140 (F7J7.140) on that BAC.

The nucleotide sequence is annotated as “similarity to gibberellin C-20 oxidase, Oryza sativa PATCHX.G1854637.” The predicted protein (CAA17538) is annotated as “gibberellin 20-oxidase-like protein.” L10 is on BAC F8A12 (accession number AC079284: Arabidopsis thaliana chromosome 1 BAC F8A12 genomic sequence, complete sequence). It is number 18 (F8A12.18) on that BAC. The nucleotide sequence is annotated as “similar to gibberellin 20 oxidase (Triticum aestivum) GI2222796.” The predicted protein (AAG50945.1) is annotated as “gibberellin 20-oxidase, putative.”

Based on the above information, we have determined the cDNA sequences for L10 and 69 as SEQ ID NO:3 and SEQ ID NO:5, respectively.

Reverse-transcription-based PCR was used to determine the expression levels of the L10 and 69 gene in the two mutants. Both L10 and 69 plants had substantially increased mRNA levels of their respective enzymes. This observation is consistent with the hypothesis that the phenotypes of the mutants were due to activation of gene expression caused by the enhancer region of the T-DNA.

3. Ectopic Expression of L10 and 69 in Arabidopsis.

To test the hypothesis that the L10 and 69 phenotypes were due to the activation of the respective GA-modifying
genes, L10 and 69 were constitutively overexpressed. The genomic region of the respective genes was cloned into a vector that contained a cauliflower mosaic virus 35S promoter (3S5) that provides constitutive mRNA expression in most plant tissues. These new vectors that contained the 35S::L10 or 35S::69 constructs were transformed into wild-type Arabidopsis. First generation transformed plants were screened for phenotypes similar to the respective initial L10 or 69 lines. Approximately half of the transformed plants displayed phenotypes similar or identical to that of the initially isolated L10 or 69 lines, respectively. Thus, increased expression of the GA-modifying genes was sufficient to cause the alterations in plant growth and stature that were seen in the initially isolated mutant lines. This data confirmed that the activation of the GA-modifying genes near the T-DNA inserts had caused the dominant GA-deficient phenotypes. In addition to the 35S-driven genomic clones, the cDNAs for each of the L10 and 69 lines were also placed under the transcriptional control of the 35S promoter and were found to also cause a dwarf, GA-deficient-like phenotype. This indicates that the cDNAs are functional and sufficient for the purposes of altering GA metabolism to produce the aforementioned phenotypes.

4. Ectopic Expression of L10 Functions in Tobacco to Produce GA-Deficient-Like Plants.

Introduction of the 35S::L10 or 35S::69 into wild-type tobacco (Wisconsin 38) produces plants that appear to be deficient in bioactive GAs. Many phenotypic changes are similar to the phenotypic changes in Arabidopsis. For example, the leaves are smaller and darker green, plant height is reduced, and internode distance is shortened (Table 5). The 35S::L10 and 35S::69 tobacco plants and the wild-type plant had similar seed yield.

5. Sequence Alignments of L10 and 69 to GA-Modifying Enzymes.

L10 and 69 are more similar to each other than to any other protein in the Arabidopsis database. When L10 is used to BLAST search the Arabidopsis database of proteins, the closest match to L10 is 69. Likewise, when 69 is used to search the database L10 is the closest match to 69. This implies that L10 and 69 may define a group of GA-modifying enzymes that may be functionally distinct from other GA-modifying enzymes (Tables 1 and 2). An alignment of L10 to 69, depicted in FIG. 1, reveals that there is 44% identity and 54.5% similarity shared by the two proteins. In FIG. 1, lines denote identity and colons and periods denote degree of similarity.

An alignment of L10 and 69 to AtGA2ox1, AtGA2ox2, AtGA2ox3, AtGA20ox1, AtGA20ox2, AtGA20ox3, and AtGA4 (a 20-oxidase), as depicted in FIG. 2, reveals that these enzymes contain regions of similarity (see also Table 2). In FIG. 2, amino acid residues that are identical with 69 protein are designated by a black box surrounding the amino acid residue and similarities in amino acid residues to 69 proteins are denoted by gray shading around residues. However, L10 and 69 contain unique regions that are similar to each other but show little or no relatedness to the other GA-modification enzymes. The two most prominent examples of this are the sequence from L10 at amino acid 115 through 137 and the carboxy terminus of L10 and 69 defined by the L10 protein sequence at amino acid 304 through 335 (FIGS. 1 and 2).

### TABLE 1

Percent identity between novel dioxygenases and other known Arabidopsis dioxygenases (Numbers in the table are percent identity. Thick lines separate the groups of dioxygenases and the values under these lines illustrate the high percent of identity within each group. All of the dioxygenases are from Arabidopsis. At 2ox1-3 are the three cloned 2-oxidases (accession nos. AJ32435, AJ132436, and AJ132437). At20ox1-3 are three 20-oxidases (accession nos. X83379, X83380, and X83381). At 3ox1-2 are two 3β-hydroxylase (accession nos. L71729 and TS1691)).

<table>
<thead>
<tr>
<th>L10</th>
<th>69</th>
<th>At 2ox1</th>
<th>At 2ox2</th>
<th>At 2ox3</th>
<th>At 20ox1</th>
<th>At 20ox2</th>
<th>At 20ox3</th>
<th>At 3ox1</th>
<th>At 3ox2</th>
</tr>
</thead>
<tbody>
<tr>
<td>L10</td>
<td>69</td>
<td>44</td>
<td>27</td>
<td>25</td>
<td>25</td>
<td>32</td>
<td>31</td>
<td>31</td>
<td>28</td>
</tr>
<tr>
<td>At 2ox1</td>
<td>29</td>
<td>26</td>
<td>27</td>
<td>32</td>
<td>30</td>
<td>31</td>
<td>29</td>
<td>29</td>
<td>25</td>
</tr>
<tr>
<td>At 2ox2</td>
<td>55</td>
<td>53</td>
<td>31</td>
<td>29</td>
<td>29</td>
<td>31</td>
<td>31</td>
<td>32</td>
<td>28</td>
</tr>
<tr>
<td>At 2ox3</td>
<td>55</td>
<td>53</td>
<td>31</td>
<td>29</td>
<td>29</td>
<td>31</td>
<td>31</td>
<td>32</td>
<td>28</td>
</tr>
<tr>
<td>At 20ox1</td>
<td>69</td>
<td>28</td>
<td>28</td>
<td>28</td>
<td>34</td>
<td>33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At 20ox2</td>
<td>27</td>
<td>30</td>
<td>28</td>
<td>33</td>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At 20ox3</td>
<td>69</td>
<td>28</td>
<td>28</td>
<td>28</td>
<td>34</td>
<td>33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At 3ox1</td>
<td>73</td>
<td>61</td>
<td>31</td>
<td>29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At 3ox2</td>
<td>64</td>
<td>32</td>
<td>32</td>
<td>32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At 3ox1</td>
<td>74</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

These data indicate that L10 and 69 define a group of GA-modifying enzymes that are similar to each other and to AtGA2ox1, AtGA2ox2, AtGA2ox3, AtGA20ox1, AtGA20ox2, AtGA20ox3, and AtGA4, but they are not related to the other GA-modification enzymes.
Percent identity between the L10 and 69 dioxygenases and other known GA-degrading enzymes (Numbers in the table are percent identity. A thick line separates the known 2-oxidases. L10 and 69 are our two novel dioxygenases. At 2ox1-3 are the three 2-oxidases in Arabidopsis (accession nos. AJ132435, AJ132436, and AJ132437). Rice 2ox is the 2-oxidase from Oryza sativa (Sakamoto, 2001). Bean 2ox is the 2-oxidase from Phaseolus coccineus (accession no. AJ132438). Rice 2ox is the 2-oxidase from Oryza sativa (Sakamoto, 2001). Bean 2ox is the 2-oxidase from Phaseolus coccineus (accession no. AJ132438). Rice 2ox is the 2-oxidase from Oryza sativa (Sakamoto, 2001). Bean 2ox is the 2-oxidase from Phaseolus coccineus (accession no. AJ132438). Rice 2ox is the 2-oxidase from Oryza sativa (Sakamoto, 2001). Bean 2ox is the 2-oxidase from Phaseolus coccineus (accession no. AJ132438). Rice 2ox is the 2-oxidase from Oryza sativa (Sakamoto, 2001). Bean 2ox is the 2-oxidase from Phaseolus coccineus (accession no. AJ132438). Rice 2ox is the 2-oxidase from Oryza sativa (Sakamoto, 2001). Bean 2ox is the 2-oxidase from Phaseolus coccineus (accession no. AJ132438). Rice 2ox is the 2-oxidase from Oryza sativa (Sakamoto, 2001). Bean 2ox is the 2-oxidase from Phaseolus coccineus (accession no. AJ132438). Rice 2ox is the 2-oxidase from Oryza sativa (Sakamoto, 2001). Bean 2ox is the 2-oxidase from Phaseolus coccineus (accession no. AJ132438). Rice 2ox is the 2-oxidase from Oryza sativa (Sakamoto, 2001). Bean 2ox is the 2-oxidase from Phaseolus coccineus (accession no. AJ132438). Rice 2ox is the 2-oxidase from Oryza sativa (Sakamoto, 2001). Bean 2ox is the 2-oxidase from Phaseolus coccineus (accession no. AJ132438). Rice 2ox is the 2-oxidase from Oryza sativa (Sakamoto, 2001). Bean 2ox is the 2-oxidase from Phaseolus coccineus (accession no. AJ132438). Rice 2ox is the 2-oxidase from Oryza sativa (Sakamoto, 2001). Bean 2ox is the 2-oxidase from Phaseolus coccineus (accession no. AJ132438). Rice 2ox is the 2-oxidase from Oryza sativa (Sakamoto, 2001). Bean 2ox is the 2-oxidase from Phaseolus coccineus (accession no. AJ132438). Rice 2ox is the 2-oxidase from Oryza sativa (Sakamoto, 2001). Bean 2ox is the 2-oxidase from Phaseolus coccineus (accession no. AJ132438). Rice 2ox is the 2-oxidase from Oryza sativa (Sakamoto, 2001). Bean 2ox is the 2-oxidase from Phaseolus coccineus (accession no. AJ132438). Rice 2ox is the 2-oxidase from Oryza sativa (Sakamoto, 2001). Bean 2ox is the 2-oxidase from Phaseolus coccineus (accession no. AJ132438). Rice 2ox is the 2-oxidase from Oryza sativa (Sakamoto, 2001). Bean 2ox is the 2-oxidase from Phaseolus coccineus (accession no. AJ132438). Rice 2ox is the 2-oxidase from Oryza sativa (Sakamoto, 2001). Bean 2ox is the 2-oxidase from Phaseolus coccineus (accession no. AJ132438).

<table>
<thead>
<tr>
<th></th>
<th>L10</th>
<th>69</th>
<th>At</th>
<th>Rice</th>
<th>Bean</th>
<th>Pea</th>
<th>Pumpkin</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2ox1</td>
<td>2ox2</td>
<td>2ox3</td>
<td>2ox1</td>
<td>2ox2</td>
<td>2ox3</td>
<td>2ox2.2</td>
</tr>
<tr>
<td>L10</td>
<td>-</td>
<td>44</td>
<td>27</td>
<td>25</td>
<td>24</td>
<td>26</td>
<td>29</td>
</tr>
<tr>
<td>69</td>
<td>-</td>
<td>29</td>
<td>26</td>
<td>27</td>
<td>23</td>
<td>25</td>
<td>28</td>
</tr>
<tr>
<td>At</td>
<td></td>
<td>55</td>
<td>53</td>
<td>38</td>
<td>57</td>
<td>47</td>
<td>55</td>
</tr>
<tr>
<td>2ox1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At</td>
<td></td>
<td>69</td>
<td>37</td>
<td>63</td>
<td>45</td>
<td>56</td>
<td>32</td>
</tr>
<tr>
<td>2ox2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>At</td>
<td></td>
<td>38</td>
<td>56</td>
<td>47</td>
<td>55</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>2ox3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rice</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2ox</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2ox1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pea</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2ox1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2ox2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TABLE 3

Characterization of the Mutant Phenotypes of L10 and 69 Lines.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Wild-type</th>
<th>L10</th>
<th>69</th>
<th>35S::L10</th>
<th>35S::69</th>
<th>GA24</th>
<th>GA3</th>
<th>GA3</th>
<th>GA4</th>
<th>GA19</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flowering Time (Number of Leaves)</td>
<td>8</td>
<td>17</td>
<td>15</td>
<td>16</td>
<td>15</td>
<td>8</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Height cm</td>
<td>47</td>
<td>7.2</td>
<td>9.4</td>
<td>8.2</td>
<td>9.6</td>
<td>30</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Flowering Branches</td>
<td>41</td>
<td>77</td>
<td>84</td>
<td>&gt;100</td>
<td>&gt;100</td>
<td>65</td>
<td>65</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internode Length mm</td>
<td>8.9</td>
<td>1.8</td>
<td>2.3</td>
<td>2.4</td>
<td>2.6</td>
<td>7</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TABLE 4

Abundance of GAs Present in Wild Type and Mutant Lines (All values are in ng/g dry weight. ND = not detectable).

<table>
<thead>
<tr>
<th>GAs</th>
<th>Ws (wild type)</th>
<th>69</th>
<th>L10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-13-Hydroxylated:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GA24</td>
<td>51.8</td>
<td>0.23</td>
<td>0.06</td>
</tr>
<tr>
<td>GA30</td>
<td>1.01</td>
<td>0.05</td>
<td>0.02</td>
</tr>
<tr>
<td>GA20</td>
<td>1.84</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>13-Hydroxylated:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GA29</td>
<td>6.43</td>
<td>0.30</td>
<td>0.39</td>
</tr>
<tr>
<td>GA33</td>
<td>0.79</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>GA40</td>
<td>9.29</td>
<td>0.09</td>
<td>0.02</td>
</tr>
</tbody>
</table>

The complete disclosures of all publications that are cited herein are hereby incorporated by reference as if individually incorporated. It is also understood that, given the limitations of the state of the art, occasional sequence errors or deletions may occur without affecting the usefulness of the data presented. Various modifications and alterations of this invention will become apparent to those skilled in the art without departing from the scope and spirit of this invention, and it should be understood that this invention is not to be unduly limited to the illustrative embodiments set forth herein, but rather is to be construed to be of spirit and scope defined by the appended claims.
SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 12

<210> SEQ ID NO 1
<211> LENGTH: 2171
<212> TYPE: DNA
<213> ORGANISM: Arabidopsis

<400> SEQUENCE:

atggcttctc aacctccctt taagacaaat ttctgctcga ttttcggaag ctcatttcca 60
aattcaacta gtgagagcaa tacaaacaca tcaactatcc aaacctcagg cataaagctt 120
cctgtgatcg atctcagcca tctaactagt ggtgaggagg tcaaacgcaa 480
aagatgtgtg aaacaaatgg ttgcagctgc gaaagagtgg ggattttttc aaattgtgaa ccatggaatt 240
cccaaagacg tctttgagat gatgctcctc gaagagaaga aactctttga ccaacctttt 300
tctgtgaaag tcagagaacg tttttcggac ttatcgaaga atagttaccg ttggggaaac 360
cctagcgcca cttctcccgc tcagtactcc gttttcggaag cgtttcacat cattctttca 420
gaggttttca ggatttctga tgcattcagtt gctgaggaa gatttttttgt tcagcataa 480
acatctattg attacatcct ggtttttccttttgcagccttattacttttctttttttggacc 540
tttaataa taatgtactc aacgtgatat gcgttttttttaaatggattttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
aaactgagat cgattgcctt gggtatccaa agaagtatag aagattcagt ttcagagagt 2100
acaaagagca
gagtgaacat gatgttaaag aaactggtga taaggtaggc ttgtccaggt 2160
ttctcatttg a 2171
<210> SEQ ID NO 2
<211> LENGTH: 3861
<212> TYPE: DNA
<213> ORGANISM: Arabidopsis
<400> SEQUENCE: 2
agatcctctc atggatccac cattcaacga aatatacaat aaccttttgt 60
agatcctctc atggatccac cattcaacga aatatacaat aaccttttgt 120
cagggagctg gagtttctg tctggagct cagcgttctt attgtgagg cagggagctg 180
cagggagctg gagtttctg tctggagct cagcgttctt attgtgagg cagggagctg 240
gataaaacc gtgaatactt gcgagagttt ggtgcttctt gcgagagttt ggtgcttctt 300
gtttaggggc cagcggctgc ctctctactc gcgctttct tcggtctagct 360
tgccagttgct gtttggagag cttctctct gcgctttct tcggtctagct 420
ttcagttgct gtttggagag cttctctct gcgctttct tcggtctagct 480
attatataac atatttttag attataataa attataataa attataataa 540
gtttaggggc atcccaataa caagatgtaa attaaataac atctctactt attactttaa 600
taatttttag attataataa attataataa attataataa attataataa 660
attatataac atatttttag attataataa attataataa attataataa 720
gtttaggggc atcccaataa caagatgtaa attaaataac atctctactt attactttaa 780
taatttttag attataataa attataataa attataataa attataataa 840
attatataac atatttttag attataataa attataataa attataataa 900
gtttaggggc atcccaataa caagatgtaa attaaataac atctctactt attactttaa 960
taatttttag attataataa attataataa attataataa attataataa 1020
attatataac atatttttag attataataa attataataa attataataa 1080
attatataac atatttttag attataataa attataataa attataataa 1140
attatataac atatttttag attataataa attataataa attataataa 1200
attatataac atatttttag attataataa attataataa attataataa 1260
attatataac atatttttag attataataa attataataa attataataa 1320
attatataac atatttttag attataataa attataataa attataataa 1380
attatataac atatttttag attataataa attataataa attataataa 1440
atccataac atatttttag attataataa attataataa attataataa 1500
attatataac atatttttag attataataa attataataa attataataa 1560
attatataac atatttttag attataataa attataataa attataataa 1620
attatataac atatttttag attataataa attataataa attataataa 1680
attatataac atatttttag attataataa attataataa attataataa 1740
attatataac atatttttag attataataa attataataa attataataa 1800
attatataac atatttttag attataataa attataataa attataataa 1860
attatataac atatttttag attataataa attataataa attataataa 1920
<210> SEQ ID NO 3
<211> LENGTH: 1011
<212> TYPE: DNA
<213> ORGANISM: Arabidopsis
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1) .. (1008)
<400> SEQUENCE: 3

```
1980
atg get tct caa cct ccc ttt aag aca aat ttc tgc tcg att ttc gga  48
2040
Met Ala Ser Gln Pro Pro Phe Lys Thr Asn Phe Cys Ser Ile Phe Gly
```

Met Ala Ser Gln Pro Pro Phe Lys Thr Asn Phe Cys Ser Ile Phe Gly
---continued

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>5</th>
<th>10</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>aGC tCA tTT cCO aAT tCA aCT aGT gAT aGC aAT aCA aCA tCA tCT</td>
<td>96</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tCO aCA aCC tCO gGC tTA aAG cTT cCT gTG aTC aGT aGC aAT cTA</td>
<td>144</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aCT aGT gGT gAG gGC aAA cCG aAS aAG tGT gTG aAA cAT aGT</td>
<td>192</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gOA gCT gCG aAA gAG tGG gGA tTT cCA aAT tGT aGC aAT cGA tGA</td>
<td>240</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cCC aAA gAC gTC tTT gAG aTG aTC cTC cTC gAA gAG aAA cTC tTT</td>
<td>288</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gAC aCA cCT tTT tCT gTG aAA gTC aGA gSA cGT tTT tCG gAC tTA tCG</td>
<td>336</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aSP gLN Pro Phe Ser Val Lys Ala Arg Phe Ser Are Lys Val Ser</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aAG aAT aGT tAC cGT tYG gGA aAC cCT tCT cCC gCT cAG</td>
<td>384</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aLy Ser Tyr Arg Try Gly Arg Ser Ala Thr Ser Pro Ala Gln</td>
<td>115</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tAC tCC gTT tCG gAA gCG tTT cAC aAT tCT tCA gAG tCT tCA gGG</td>
<td>432</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aSP Ser Val Ser Glu Ala His Ile Leu Ser Glu Val Ser Arg</td>
<td>130</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aTT tCT gAT gCT cGC aAC cTC aGA aCA aTC aGT gAA gCG tAT gTG</td>
<td>480</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aIle Ser Asp Arg Asp Ala Leu Arg Thr Ile Val glu Ala Tyr Val</td>
<td>145</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cCA gAG aTA gCT cGA gOA cAA aTG aTA gTG gAA aTA cTG gGG gAA</td>
<td>528</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gLN Glu Ile Ala Arg Val Ala Gln Met Ile Cys Glu Ile Leu Gly Lys</td>
<td>165</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cCA gTG aAC gTG aGT tCG gAG tAT tTC gAA aAC aAT tTT gAG cTT gAA</td>
<td>576</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gIN Val Asp Val Ser Glu Tyr Phe Arg Am Ile Phe Glu Leu Glu</td>
<td>180</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aAC aGT tTT cTA aGG cTC aAT aAG tAC cAT cCT aGT tTT gGT tCT</td>
<td>624</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aAm Ser Phe Pro Leu Arg Leu Arg Thr Ile Val glu Ala Tyr Val</td>
<td>195</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gAA gTG tTT gGT tGG cCT cAT aCC gAT cAC aAA tGG tCT cCC aCT cAT aTA</td>
<td>672</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gIN Val Phe Glu Leu Val Pro His Thr Asp Thr Ser Phe Leu Thr Ile</td>
<td>210</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cTC cTT cCO aGA aCA tCO gGA ggG tAA gAA gAT gAA cAA</td>
<td>720</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu Ser Glu Ala Gln Ile Gly Gly Leu Leu Leu Arg Gly Gly Glu</td>
<td>225</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tGG aTC aGC tTA aAA cCT gTC tGG gAA gCC cTT aCA aGC cTC aGG aTT gGG</td>
<td>768</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trp Ile Ser Val Lys Pro Cys Leu Glu Ala Leu Thr Val Arg Ile Gly</td>
<td>245</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gAT tGG tTT cAG gCO aTG aAT gAA gTG tAC cAA aGC gTG aGA cAT</td>
<td>816</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aSP Met Phe Glu Ala Leu Ser Arg Lys Val Tyr Gln Ser Val Arg His</td>
<td>260</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aGA gTG aTT tCT cCA gCO aAT aGC aAG aTG cTA cTA aGT tTC tTC aArg Val Ile Ser Pro Ala Asp Ile Glu Arg Met Ser Ile Ala Phe Phe</td>
<td>275</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gTA gTG cCT cAT cTC aGA aCT aGC aGG aCG aTC aAT cTA gCT tTC tTC</td>
<td>912</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val Cys Pro Tyr Leu Glu Thr Glu Ile Asp Cys Phe Gly Tyr Pro Lys</td>
<td>290</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aAG tAT aGA aAG cTG tAC aAT cTC aCT cGA aAG aAG cAG aGG cAT</td>
<td>861</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tyr Thr Tyr Tyr Leu Glu Thr Glu Ile Asp Cys Phe Gly Tyr Pro Lys</td>
<td>295</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gAT gTT aAA gaa act ggt gat aag gta ggc tgg tcc agg ttt ctc aTT</td>
<td>1008</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Asp Val Lys Glu Thr Gly Asp Lys Val Gly Leu Ser Arg Phe Leu Ile
325 330 335
tga

<210> SEQ ID NO 4
<211> LENGTH: 336
<212> TYPE: PRT
<213> ORGANISM: Arabidopsis
<400> SEQUENCE: 4
Met Ala Ser Gln Pro Pro Phe Thr Asn Phe Cys Ser Ile Phe Gly
1  5 10 15
Ser Ser Phe Pro Asn Ser Thr Ser Asp Ser Asn Thr Ser Thr
20 25 30
Ile Gln Thr Ser Gly Leu Lys Leu Pro Val Ile Asp Leu Ser His Leu
35  40 45
Thr Ser Gly Glu Val Lys Arg Arg Cys Val Lys Gln Met Val
50  55 60
Ala Ala Ala Lys Glu Trp Gly Phe Phe Glu Ile Val Asn His Gly Ile
65  70 75 80
Pro Lys Asp Val Phe Glu Met Met Leu Leu Glu Lys Lys Leu Phe
85  90 95
Asp Gln Pro Phe Ser Val Lys Val Arg Glu Arg Phe Ser Asp Leu Ser
100 105 110
Lys Asn Ser Tyr Arg Trp Gly Asn Pro Ser Ala Thr Ser Pro Ala Gln
115 120 125
Tyr Ser Val Ser Glu Ala Phe His Ile Ile Leu Ser Glu Val Ser Arg
130 135 140
Ile Ser Asp Ser Asp Arg Asn Ser Leu Arg Thr Ile Val Glu Ala Tyr Val
145 150 155 160
Gln Glu Ile Ala Arg Ala Val Glu Leu Cys Glu Ile Leu Gly Lys
165 170 175
Gln Val Asn Val Ser Ser Gly Tyr Phe Glu Asn Ile Phe Glu Leu Glu
180 185 190
Asn Ser Phe Leu Arg Leu Asn Tyr His Pro Ser Val Phe Gly Ser
195 200 205
Glu Val Phe Gly Leu Val Pro His Thr Asp Thr Ser Phe Leu Thr Ile
210 215 220
Leu Ser Gln Asp Gln Ile Gly Gly Leu Glu Asn Arg Gly Gln
225 230 235 240
Trp Ile Ser Val Lys Ser Cys Leu Glu Ala Leu Thr Val Asn Ile Gly
245 250 255
Asp Met Phe Gln Ala Leu Ser Asn Gly Val Tyr Gln Ser Val Arg His
260 265 270
Arg Val Ile Ser Pro Ala Asn Ile Glu Arg Met Ser Ile Ala Phe Phe
275 280 285
Val Cys Pro Tyr Leu Glu Thr Glu Ile Asp Cys Phe Gly Tyr Pro Lys
290 295 300
Lys Tyr Arg Arg Phe Ser Phe Arg Glu Tyr Lys Glu Gln Ser Glu His
305 310 315 320
Asp Val Lys Glu Thr Gly Asp Lys Val Gly Leu Ser Arg Phe Leu Ile
325 330 335

<210> SEQ ID NO 5
<211> LENGTH: 1017
US 6,921,849 B2

-continued

<212> TYPE: DNA
<213> ORGANISM: Arabidopsis
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (1)..(1014)

<400> SEQUENCE: 5

```
atg gac cca cca ttc aac gaa ata tac aat aac ctt ttg tat aat cag
Met Asp Pro Pro Phe Asn Glu Ile Tyr Aen Aen Leu Leu Tyr Aen Gln
1 5 10 15
ato cca aac aag gaa aac gat gtt tct gaa ata ccc ttt aac gtt ttc
Ile Thr Asn Lys Glu Aen Asp Val Ser Glu Ile Pro Phe Ser Ser
20 25 30
gtc aca gtc gtc gag gag gtt gag ctt cct gtt att gag gtc aac
Val Thr Ala Val Val Glu Val Glu Val Leu Ile Asp Val Ser
35 40 45
cgt tgg att gag gca gcc gag gag gag aag ttt gaa gag gcg
Arg Leu Ile Asp Gly Ala Glu Glu Arg Glu Lys Lys Glu Ala
50 55 60
att gcc gca gca gca gca gca gaa ctc ccg ttt cga cca cca cct
Ile Ala Arg Ala Ser Arg Trp Gly Phe Glu Val Asn Pro Pro
65 70 75 80
```

```
tgt cca tca tac gac gcc gtt ata gag tgt tca aat gat cgt cct gct

Cys Pro Ser Tyr Asp Ala Val Ile Glu Cys Ser Ser Asp Arg Pro Ala

290 295 300

tat aga aat ttc age ttc aga gaa ttc aga caa caa gtt caa gaa gat

Tyr Arg Asn Phe Ser Phe Arg Glu Phe Arg Glu Glu Val Glu Glu Aasp

305 310 315 320

gtt aag aag ttt ggt ttt aaa gtt ggc ctt cct agg tct ctt aat cac

Val Lys Lys Phe Gly Phe Lys Val Gly Leu Pro Arg Phe Leu Aas His

325 330 335

Val Tyr

<210> SEQ ID NO 6
<211> LENGTH: 338
<212> TYPE: PRT
<213> ORGANISM: Arabidopsis

<400> SEQUENCE: 6
Met Asp Pro Pro Phe Asn Glu Ile Tyr Aen Aen Leu Leu Tyr Aen Gln
Ile Thr Aen Lys Glu Asn Asp Val Ser Glu Ile Pro Phe Ser Phe Ser
Val Thr Ala Val Val Glu Glu Val Glu Leu Pro Val Ile Asp Val Ser
Arg Leu Ile Asp Gly Ala Glu Glu Arg Glu Lys Cys Lys Glu Ala
Ile Ala Arg Ala Ser Arg Glu Trp Gly Phe Phe Glu Val Ile Aas His
Gly Ile Ser Met Asp Val Leu Glu Lys Met Arg Glu Glu Glu Ile Arg
Val Phe Arg Glu Pro Phe Asp Lys Lys Ser Glu Lys Phe Ser
Ala Gly Ser Tyr Arg Trp Gly Thr Pro Ser Ala Thr Ser Ile Arg Gln
Leu Ser Trp Ser Glu Ala Phe His Val Pro Met Thr Asp Ile Ser Asp
Aen Lys Asp Phe Thr Leu Ser Ser Thr Met Glu Lys Phe Ala Ser
Glu Ser Glu Ala Leu Ala Tyr Met Leu Ala Glu Val Leu Ala Glu Lys
Ala Gly Glu Lys Ser Ser Phe Phe Glu Aen Cys Val Arg Aen Thr
Cys Tyr Leu Arg Met Aas Arg Tyr Pro Pro Cys Pro Lys Pro Ser Glu
Val Tyr Gly Leu Met Pro His Thr Asp Ser Asp Phe Leu Thr Ile Leu
Tyr Glu Asp Glu Val Gly Gly Leu Glu Leu Ile Lys Asp Aen Arg Trp
Ile Ala Val Lys Pro Asn Pro Lys Ala Leu Ile Ile Aas Ile Gly Asp
Leu Phe Glu Ala Trp Ser Aen Gly Met Tyr Lys Ser Val Glu His Arg
Val Met Thr Aen Pro Lys Val Glu Arg Phe Ser Thr Ala Tyr Phe Met
Cys Pro Ser Tyr Asp Ala Val Ile Glu Cys Ser Ser Asp Arg Pro Ala
Tyr Arg Asn Phe Ser Phe Arg Glu Phe Arg Gin Gin Val Gin Glu Asp
Val Lys Lys Phe Gly Phe Lys Val Gin Leu Pro Arg Phe Lys Asn His
Val Tyr

<210> SEQ ID NO 7
<211> LENGTH: 378
<212> TYPE: PRT
<213> ORGANISM: Arabidopsis

<400> SEQUENCE: 7

Met Ala Ile Leu Cys Thr Thr Ser Pro Ala Glu Lys Glu His Glu
Pro Lys Gin Asp Leu Glu Lys Asp Gin Thr Ser Pro Leu Ile Phe Asn
Pro Ser Leu Leu Asn Leu Gin Ser Gin Ile Pro Asn Gin Phe Ile Trp
Pro Asp Glu Glu Pro Ser Ile Asp Ile Pro Glu Leu Asn Val Pro
Phe Ile Asp Leu Ser Ser Gin Asp Ser Thr Leu Gin Ala Pro Arg Val
Ile Ala Gin Ala Cys Thr Lys Gin Phe Phe Leu Val Gin His His
Gly Val Ser Gin Ser Leu Ile Gin Ala Gin Arg Leu Met Gin Met
Phe Phe Asp Met Pro Leu Ala Gly Lys Gin Lys Gin Arg Lys Pro
Gly Glu Ser Cys Gin Tyr Ala Ser Ser Phe Thr Gin Arg Phe Ser Thr
Lys Leu Pro Trp Lys Gin Thr Leu Ser Phe Gin Phe Ser Gin Asp Asn
Ser Gin Ser Arg Thr Val Gin Asp Tyr Phe Ser Gin Thr Leu Gin Gin
Glu Phe Gin Gin Gin Gin Tyr Gin Gin Gin Gin Cys Gin Ala Gin
Ser Ser Leu Ser Leu Lys Ile Gin Leu Leu Gin Leu Gin Gin Leu Gin
Val Gin Arg Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
Gly Thr Gin Gin Thr Gin Thr Gin Gin Gin Gin Gin Gin Gin Gin Gin
Glu Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
Asp Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
Lys Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
Val Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
Val Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
Lys Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin
<210> SEQ ID NO 8
<211> LENGTH: 377
<212> TYPE: PRT
<213> ORGANISM: Arabidopsis

<400> SEQUENCE: 8

Met Ala Val Ser Phe Val Thr Thr Ser Pro Glu Glu Glu Asp Lys Pro
1  5 10 15
Lys Leu Gly Leu Gly Asn Ile Gln Thr Pro Leu Ile Phe Asn Pro Ser
20 25 30
Met Leu Asn Leu Gln Ala Asn Ile Pro Asn Gln Phe Ile Trp Pro Asp
35 40 45
Asp Glu Lys Pro Ser Ile Asn Val Leu Glu Leu Asp Val Pro Leu Ile
50 55 60
Asp Leu Gln Asn Leu Ser Asp Pro Ser Ser Thr Leu Asp Ala Ser
65 70 75 80
Arg Leu Ile Ser Glu Ala Cys Lys Lys His Gly Phe Phe Leu Val Val
85 90 95
Asn His Gly Ile Ser Glu Leu Ile Ser Asp Ala His Glu Tyr Thr
100 105 110
Ser Arg Phe Phe Asp Met Pro Leu Ser Glu Gln Arg Val Leu Arg
115 120 125
Lys Ser Gly Glu Ser Val Gly Tyr Ala Ser Phe Thr Gly Arg Phe
130 135 140
Ser Thr Lys Leu Pro Trp Lys Glu Thr Leu Ser Phe Arg Phe Cys Asp
145 150 155 160
Asp Met Ser Arg Ser Lys Ser Val Gin Asp Tyr Phe Cys Asp Ala Leu
165 170 175
Gly His Gly Phe Glu Pro Phe Gly Lys Val Tyr Gin Glu Tyr Cys Glu
180 185 190
Ala Met Ser Ser Leu Ser Leu Lys Ile Met Glu Leu Leu Gly Leu Ser
195 200 205
Leu Gly Val Lys Arg Asp Tyr Phe Arg Glu Phe Glu Asn Asp
210 215 220
Ser Ile Met Arg Leu Asn Tyr Phe Phe Cys Ile Lys Pro Asp Leu
225 230 235 240
Thr Leu Gly Thr Gly Pro His Cys Asp Pro Thr Ser Leu Thr Ile Leu
245 250 255
His Gln Asp His Val Asn Gin Leu Gin Val Phe Val Glu Gin Gin Trp
260 265 270
Arg Ser Ile Arg Pro Asn Pro Lys Ala Phe Val Val Asn Ile Gly Asp
275 280 285
Thr Phe Met Ala Leu Ser Asn Arg Tyr Lys Ser Cys Leu His Arg
290 295 300
Ala Val Val Asn Ser Glu Arg Met Arg Lys Ser Leu Ala Phe Phe Leu
305 310 315 320
Cys Pro Lys Lys Asp Arg Val Val Thr Pro Pro Arg Glu Leu Leu Asp 325 330 335
Ser Ile Thr Ser Arg Arg Tyr Pro Asp Phe Thr Trp Ser Met Phe Leu 340 345 350
Glu Phe Thr Gln Lys His Tyr Arg Ala Asp Met Asn Thr Leu Gln Ala 355 360 365
Phe Ser Asp Trp Leu Thr Lys Pro Ile 370 375

<210> SEQ ID NO 9
<211> LENGTH: 358
<212> TYPE: PRT
<213> ORGANISM: Arabidopsis
<400> SEQUENCE: 9

Met Pro Ala Met Leu Thr Asp Val Phe Arg Gly His Pro Ile His Leu 1 5 10 15
Pro His Ser His Ile Pro Asp Phe Thr Ser Leu Arg Glu Leu Pro Asp 20 25 30
Ser Tyr Lys Trp Thr Pro Lys Asp Leu Leu Phe Ser Ala Ala Pro 35 40 45
Ser Pro Pro Ala Thr Gly Glu Asn Ile Pro Leu Ile Asp Leu Asp His 50 55 60
Pro Asp Ala Thr Asn Gln Ile Gly His Ala Cys Arg Thr Trp Gly Ala 65 70 75 80
Phe Glu Thr Gly Ser Leu Phe Gly Leu Pro Val Gln Arg Lys Leu 85 90 95
Glu Phe Leu Thr Gly Ser Leu Phe Gly Leu Pro Val Gln Arg Lys Leu 100 105 110
Lys Ser Ala Arg Ser Glu Thr Gly Val Ser Gly Tyr Gly Val Ala Arg 115 120 125
Ile Ala Ser Phe Phe Asn Lys Gln Met Trp Ser Glu Gly Phe Thr Ile 130 135 140
Thr Gly Ser Pro Leu Asn Asp Phe Arg Lys Leu Trp Pro Glu His Ala 145 150 155 160
Leu Asn Tyr Cys Asp Ile Val Glu Tyr Glu His Met Lys Lys 165 170 175
Leu Ala Ser Lys Leu Met Trp Leu Ala Leu Asn Ser Leu Gly Val Ser 180 185 190
Glu Glu Asp Ile Glu Trp Ala Ser Leu Ser Asp Leu Asn Trp Ala 195 200 205
Gln Ala Ala Leu Glu Asn His Tyr Pro Val Cys Pro Glu Pro Asp 210 215 220
Arg Ala Met Gly Leu Ala Ala His Thr Asp Ser Thr Leu Thr Ile 225 230 235 240
Leu Tyr Glu Asn Asn Ala Gly Leu Glu Val Phe Arg Asp Asp Leu 245 250 255 260 265
Gly Trp Val Thr Val Pro Pro Phe Pro Gly Ser Leu Val Val Val Val 270
Gly Asp Leu Phe His Ile Leu Ser Asn Gly Leu Phe Lys Ser Val Leu 275 280 285
His Arg Ala Arg Val Asn Thr Arg Ala Arg Leu Ser Val Ala Phe 290 295 300
Leu Trp Gly Pro Gln Ser Asp Ile Lys Ile Ser Pro Val Pro Lys Leu 305 310 315 320
<table>
<thead>
<tr>
<th>Val Ser Pro Val Glu Ser Pro Leu Tyr Gln Ser Val Thr Trp Lys Glu</th>
<th>325 330 335</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tyr Leu Arg Thr Lys Ala Thr His Phe Asn Lys Ala Leu Ser Met Ile</td>
<td>340 345 350</td>
</tr>
<tr>
<td>Arg Asn His Arg Glu Glu</td>
<td>355</td>
</tr>
</tbody>
</table>

<210> SEQ ID NO 10
<211> LENGTH: 330
<212> TYPE: PRT
<213> ORGANISM: Arabidopsis

<400> SEQUENCE: 10

Met Ala Val Leu Ser Lys Pro Val Ala Ile Pro Lys Ser Gly Phe Ser
  1   5   10   15
Leu Ile Pro Val Ile Asp Met Ser Asp Pro Glu Ser Lys His Ala Leu
  20  25  30
Val Lys Ala Cys Glu Asp Phe Gly Phe Phe Lys Val Ile Asn His Gly
  35  40  45
Val Ser Ala Glu Leu Val Ser Val Leu Glu His Glu Thr Val Asp Phe
  50  55  60
Phe Ser Leu Pro Lys Ser Glu Lys Thr Gin Val Ala Gly Tyr Pro Phe
  65  70  75  80
Gly Tyr Gly Asn Ser Lys Ile Gly Arg Asn Gly Asp Val Gly Trp Val
  85  90  95
Glut Tyr Leu Leu Met Asn Ala Asn His Asp Ser Gly Ser Gly Gly Pro
 100 105 110
Leu Phe Pro Ser Leu Leu Lys Pro Gly Thr Phe Arg Asn Ala Leu
 115 120 125
Glu Glu Tyr Thr Ser Val Arg Lys Met Thr Phe Asp Val Leu Glu
 130 135 140
Lys Ile Thr Asp Gly Leu Gly Ile Lys Pro Arg Asn Thr Leu Ser Lys
 145 150 155 160
Leu Val Ser Asp Gln Asn Thr Asp Ser Ile Leu Arg Leu Asn His Tyr
 165 170 175
Pro Pro Cys Pro Leu Ser Asn Lys Lys Thr Asn Gly Gly Lys Asn Val
 180 185 190
Ile Gly Phe Gly Glu His Thr Asp Pro Gln Ile Ile Ser Val Leu Arg
 195 200 205
Ser Asn Asn Thr Ser Gly Leu Gln Ile Asn Leu Asn Gly Ser Trp
 210 215 220
Ile Ser Val Pro Asp His Thr Ser Phe Phe Phe Asn Val Gly Asp
 225 230 235 240
Ser Leu Gln Val Met Thr Asn Gly Arg Phe Lys Ser Val Arg His Arg
 245 250 255
Val Leu Ala Asn Cys Lys Lys Ser Arg Val Ser Met Ile Tyr Phe Ala
 260 265 270
Gly Pro Ser Leu Thr Gln Arg Ile Ala Pro Leu Thr Cys Leu Ile Asp
 275 280 285
Asn Glu Asp Glu Arg Leu Tyr Glu Gly Phe Thr Ser Glu Tyr Lys
 290 295 300
Asn Ser Thr Tyr Asn Ser Arg Leu Ser Asp Asn Leu Gln Gln Phe
 305 310 315 320
Glu Arg Lys Thr Ile Lys Asn Leu Leu Asn
<210> SEQ ID NO 11
<211> LENGTH: 342
<212> TYPE: PRT
<213> ORGANISM: Arabidopsis

<400> SEQUENCE: 11

Met Val Val Leu Pro Gln Pro Val Pro Val Leu Asp Asn His Ile Ser Leu
1  5  10  15
Ile Pro Thr Tyr Lys Pro Val Pro Val Leu Thr Ser His Ser Ile Pro
20  25  30
Val Val Asn Leu Ala Asp Pro Glu Ala Lys Thr Arg Ile Val Lys Ala
35  40  45
Cys Glu Glu Phe Gly Phe Lys Val Val Asn His Gly Val Arg Pro
50  55  60
Glu Leu Met Thr Arg Leu Glu Glu Ala Ile Gly Phe Gly Leu
65  70  75  80
Pro Gln Ser Leu Lys Asn Arg Ala Gly Pro Pro Glu Pro Tyr Gly Tyr
85  90  95
Gly Asn Lys Arg Ile Gly Pro Amly Gly Asp Val Gly Trp Ile Gly Tyr
100 105 110
Leu Leu Leu Asn Ala Asp Pro Glu Leu Ser Ser Pro Lys Thr Ser Ala
115 120 125
Val Phe Arg Glu Thr Pro Gln Ile Phe Arg Asn Ala Leu Glu Glu Tyr
130 135 140
Thr Thr Ser Val Arg Lys Met Thr Phe Asp Val Leu Glu Lys Ile Thr
145 150 155 160
Asp Gly Leu Gly Ile Lys Pro Arg Asn Thr Leu Ser Lys Leu Val Ser
165 170 175
Asp Gln Asn Thr Asp Ser Ile Leu Arg Leu Asn His Tyr Pro Pro Cys
180 185 190
Pro Leu Ser Asn Lys Lys Thr Arg Gly Gly Asn Val Ile Gly Phe
195 200 205
Gly Glu His Thr Asp Pro Gln Ile Ile Ser Val Leu Arg Ser Asn Asn
210 215 220
Thr Ser Gly Leu Glu Ile Asn Leu Asp Gly Ser Trp Ile Ser Val
225 230 235 240
Pro Pro Asp His Thr Ser Phe Phe Pro Val Gly Asp Ser Leu Glu
245 250 255
Val Met Thr Asn Gly Arg Phe Lys Ser Val Arg His Arg Val Leu Ala
260 265 270
Asn Cys Lys Lys Ser Arg Val Ser Met Ile Tyr Phe Ala Gly Pro Ser
275 280 285
Leu Thr Glu Arg Ile Ala Pro Leu Cys Leu Ile Asp Asn Glu Asp Glu
290 295 300
Arg Leu Tyr Glu Glu Phe Thr Trp Ser Glu Tyr Lys Asn Ser Thr Tyr
305 310 315 320
Asn Ser Arg Leu Ser Asp Asn Arg Leu Glu Glu Phe Glu Arg Lys Thr
325 330 335
Ile Lys Asn Leu Leu Leu
340

<210> SEQ ID NO 12
<211> LENGTH: 335
<212> TYPE: PRT
<213> ORGANISM: Arabidopsis

<400> SEQUENCE: 12

Met Val Ile Val Leu Gln Pro Ala Ser Phe Asp Ser Asn Leu Tyr Val
1   5   10   15

Asn Pro Lys Cys Lys Pro Arg Pro Val Leu Ile Pro Val Ile Asp Leu
20  25  30

Thr Asp Ser Asp Ala Lys Thr Gln Ile Val Lys Ala Cys Glu Glu Phe
35  40  45

Gly Phe Phe Lys Val Ile Asn His Gly Val Arg Pro Asp Leu Leu Thr
50  55   60

Gln Leu Glu Glu Ala Ile Asn Phe Phe Ala Leu His His Ser Leu
65  70  75  80

Lys Asp Lys Ala Gly Pro Pro Asp Pro Phe Gly Tyr Gly Thr Lys Arg
85  90  95

Ile Gly Pro Asn Gly Asp Leu Gly Trp Leu Glu Tyr Ile Leu Leu Asn
100 105 110

Ala Asn Leu Cys Leu Glu Ser His Tyr Thr Ala Ile Phe Arg His
115 120 125

Thr Pro Ala Ile Phe Arg Glu Ala Val Glu Gly Tyr Ile Lys Glu Met
130 135 140

Lys Arg Met Ser Ser Lys Phe Leu Glu Met Val Glu Glu Leu Lys
145 150 155 160

Ile Glu Pro Lys Glu Lys Leu Ser Arg Leu Val Lys Val Lys Glu Ser
165 170 175

Asp Ser Cys Leu Arg Met Asn His Tyr Pro Glu Lys Glu Glu Thr Pro
180 185 190

Val Lys Glu Glu Ile Gly Phe Gly Glu His Thr Asp Pro Gln Leu Ile
195 200 205

Ser Leu Leu Arg Ser Asn Thr Glu Leu Gln Ile Cys Val Lys
210 215 220

Asp Gly Thr Trp Val Asp Val Thr Pro Asp His Ser Ser Phe Phe Val
225 230 235 240

Leu Val Gly Asp Thr Leu Gln Val Met Thr Asn Gly Arg Phe Lys Ser
245 250 255

Val Lys His Arg Val Val Thr Tyr Arg Ser Arg Ile Ser Met
260 265 270

Ile Tyr Phe Ala Gly Pro Pro Leu Ser Glu Lys Ile Ala Pro Leu Ser
275 280 285

Cys Leu Val Pro Lys Gln Asp Cys Leu Tyr Asn Glu Phe Thr Trp
290 295 300

Ser Gln Tyr Lys Leu Ser Ala Tyr Lys Thr Lys Leu Gly Asp Tyr Arg
305 310 315 320

Leu Gly Leu Phe Glu Lys Arg Pro Pro Phe Ser Leu Ser Asn Val
325 330 335
We claim:

1. An isolated nucleic acid molecule comprising a polynucleotide having an uninterrupted coding sequence that encodes the amino acid sequence of SEQ ID NO:4.

2. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide has the nucleotide sequence of SEQ ID NO:3.

3. A nucleic acid construct comprising:
   a polynucleotide having a nucleotide sequence that encodes a polypeptide, wherein the polypeptide has the amino acid sequence of SEQ ID NO:4 or an amino acid sequence exhibiting 95% sequence identity to SEQ ID NO:4, operably linked to a heterologous plant expressible promoter, wherein expression of the nucleic acid in a transgenic plant causes the plant to be shorter compared to a non-transgenic plant of the same genetic background.

4. A nucleic acid construct comprising: a polynucleotide having an uninterrupted coding sequence that encodes a polypeptide, wherein the polypeptide has the amino acid sequence of SEQ ID NO:4 or an amino acid sequence exhibiting 95% sequence identity to SEQ ID NO:4, operably linked to a heterologous plant expressible promoter wherein expression of the nucleic acid in a transgenic plant causes the plant to be shorter compared to a non-transgenic plant of the same genetic background.

5. A transgenic plant comprising in its genome the nucleic acid construct of claim 4.

6. The transgenic plant of claim 5, wherein the transgenic plant is at least about 20% shorter than a non-transgenic plant of the same genetic background while being grown under the same conditions.

7. A seed of the transgenic plant of claim 5 wherein the seed contains the same transgene as the transgenic plant.

8. A purified polypeptide comprising the amino acid sequence of SEQ ID NO:4.

9. The nucleic acid construct of claim 3, wherein the polynucleotide has a nucleotide sequence that encodes SEQ ID NO:4.

10. The nucleic acid construct of claim 4, wherein the polynucleotide has an uninterrupted coding sequence that encodes SEQ ID NO:4.

* * * * *