

TITLE OF THE INVENTION

Novel Plant Homeodomain Protein-encoding Genes and Their Uses

CROSS-REFERENCES TO RELATED APPLICATIONS

Not applicable.

STATEMENT REGARDING FEDERALLY SPONSORED
RESEARCH OR DEVELOPMENT

Not applicable.

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention relates to polynucleotides that induce or promote regeneration of plants.

2. Description of the Related Art

A bibliography follows at the end of the Detailed Description of the Invention. The listed references are all incorporated herein by reference.

Organogenesis and embryogenesis are two pathways leading to plant regeneration in a plant tissue culture. Traditionally, these two pathways are achieved through the manipulation of the contents and ratios of hormones in a plant-cell-culture medium as well as environmental conditions of a plant cell culture. Today, it is well known that these hormones and environmental cues activate specific proteins that play pivotal roles in the initiation of organogenesis or embryogenesis. This understanding opens a new door to novel methods for manipulation of plant regeneration through a direct control of gene expression. For example, over-expression of Arabidopsis *ESR1* and *ESR2*, which

are transcriptional factors belonging to AP2 family, induces shoot regeneration in a plant-cell-culture medium that does not contain any cytokinins (Banno et al., 2001; US Pat. 6,441,276; 6,407,312). Likewise, genes involved in cytokinin production, such as *ipt*, can be over-expressed to substitute for a cytokinin in a culture medium (Ooms et al., 1983, Smigocki and Owens, 1988; Ebinuma et al., 1997). Similarly, genes involved in transduction of cytokinin signals, such as cytokinin activated histidine kinase (CKI1) and two-component transcription activators (ARR1 and ARR2), can promote shoot regeneration in Arabidopsis (Kakimoto, 1996; Sakai et al., 2001; Hwang and Sheen, 2001; Imamura et al., 2003).

Gene transformation is an important tool of molecular biology as well as crop improvement. Although currently many plants can be transformed using methods such as Agrobacterium-mediated T-DNA conjugation and particle gun bombardment, the efficiency of those methods varies, and there is a need to improve transformation efficiency of important crops, such as cotton, maize and soybean. Improved transformation efficiency is especially important to perform a large-scale analysis of gene functions in these crops. Improved transformation efficiency can be achieved by a higher rate of successful transformation, easier selection of successful transformants, shorter time for such a selection, or lesser use of antibiotics or herbicides for such a selection, and will generate economical, industrial or academic benefits.

Cotton is an economically important crop but is one of the most difficult plants to transform. It usually takes about 1.5 years to produce transgenic seeds. No cotton gene has so far been reported to improve plant regeneration or transformation although some Arabidopsis genes have been reported to improve transformation efficiency of root

explants of *Arabidopsis* (Banno et al., 2001; US Pat. 6,441,276; 6,407,312). Therefore, there is a need for such a cotton gene.

Although currently antibiotics or herbicide resistance markers are almost exclusively used for selection of plant transformants (Yoder and Goldsbrough, 1994), they generally have negative effects on proliferation and differentiation and sometimes retard differentiation of adventitious shoots during the transformation process (Ebinuma et al., 1997). Further, they pose environmental or health risks (Bryant and Leather, 1992; Gressel, 1992; Flavell et al., 1992). The availability of new selection markers also facilitates stacking of multiple transgenic traits. Consequently, there have been considerable efforts to develop alternative selection systems for plant transformants. US Pat. 6,441,276 and 6,407,312 disclose methods of selecting transformants using ESR genes.

Arabidopsis *Wuschel* was first described as a mutant defective in shoot apical meristem initiation and maintenance (Endrizzi et al., 1996; Laux et al., 1996; Mayer et al., 1998). The phenotype was attributed to mutation of a single gene (*Wus*) that encodes a homeodomain ("HD") protein, a probable transcriptional factor (Mayer et al., 1998). Interestingly, although *Arabidopsis* contains at least 14 genes that are predicted to encode a highly similar HD domain at the N-terminus and a highly divergent sequence at the C-terminus to *Wus* gene, none of them could substitute for *Wus* in the *Wuschel* mutant. This indicates that these genes have different functions. Recently it was demonstrated that the mRNAs of these genes have unique expression profiles (Haecker et al., 2004). Except for *Wus* and *Prs*, the functions of these genes have not yet been defined. Although over-expression of the *Arabidopsis* *Wus* induces shoot or somatic embryo

formation in *Arabidopsis* and rice (Zuo et al., 2002; Kamiya et al., 2003), not every HD domain protein has this property. The *Arabidopsis PRS*, for example, was required for flower development and its over-expression induces cell proliferation rather than shoot regeneration (Matsumoto and Okada, 2001). Another HD domain protein, At1g46480, was not able to induce shoot regeneration in a cytokinin-free medium (Table 1). A PCT publication WO 01/23575 A2 discloses several putative *Wuschel* homologues from maize and soybean. It was also shown that over-expression of *Wuschel* also induced shoot regeneration both in *Arabidopsis* and rice (Zuo et al., 2002; Kamiya et al., 2003).

Transcription factors generally consist of at least two modules that are often exchangeable between different members or classes. This type of chimeric transcriptional factors has been well documented in the literature. For example, VXE transcriptional factor is a fusion protein of a viral activation domain VP16, an *E.coli* LexA DNA binding domain and a human estrogen receptor regulatory domain (Zou and Chua, 2000). Surprisingly, a fusion protein containing the HD domain of GhCIR1 and the VP16 activation domain did not promote shoot regeneration (Table 5). This indicates that the HD domain alone is not sufficient for the shoot regeneration enhancing function of certain HD domain proteins. Phylogenic analysis showed that none of the three polynucleotides described in this disclosure, i.e., SEQ ID NOs: 1, 3 and 5, are closely related to *Wuschel* (Fig. 3).

Development of efficient and simple transformation techniques has made great contribution to rapid advance in molecular genetics in *Arabidopsis*. Currently, both *in planta* (flora dip) and root explant methods allow a large number of genes to be mutated or transformed (Clough and Bent, 1998; Valvekens et al., 1988). For example, Banno

and Chua (2001) have identified an Arabidopsis gene by functional screening of an Arabidopsis cDNA library using Arabidopsis root explants. On the other hand, we identified a non-Arabidopsis (*GhCIR1*) gene by direct functional screening using Arabidopsis root explants (Example 2).

SUMMARY OF THE INVENTION

One aspect of the invention relates to novel genes, such as *GhCIR1* from cotton (*Gossypium hirsutum*), *AtCIR1* and *AtCIR2* from *Arabidopsis thaliana*, and a synthetic chimeric genes such as SEQ ID NO: 15.

Another aspect of the invention relates to homologues of the novel genes.

Another aspect of the invention relates to the novel polypeptides encoded by the novel genes.

Another aspect of the invention relates to the homologues, especially, polypeptides containing a signature sequence.

Another aspect of the invention relates to the homologues of the novel genes encoding the polypeptides containing a signature sequence.

Another aspect of the invention relates to transformed cells with the novel genes or the homologues of the novel genes the polypeptides containing a signature sequence.

Another aspect of the invention relates to inducing or promoting regeneration of plants, such as Arabidopsis and cotton, using the novel genes.

Another aspect of the invention relates to a method for functional identification of a gene that induces or promotes plant regeneration.

Another aspect of the invention relates to selection or screening of transformants using the novel genes or their homologues as a selection marker.

Another aspect of the invention relates to improving regeneration or transformation efficiency of plants using the novel genes or their homologues.

BRIEF DESCRIPTION OF THE FIGURES

Fig. 1 shows the gene organization in T-DNA region of pSK36. Promoters are shown as empty boxes. Protein coding regions are shown as dark arrows. Gene transcriptional terminators are shown as gray boxes. Cotton cDNAs are directionally inserted between *AscI* and *NotI* sites so that the expression of the inserted cDNAs is under the regulation of CaMV 35S promoter (P35S). NPTI and NPTII are kanamycin resistance markers for bacteria and plants. *Pnos* and *Tnos* are the promoter and terminator of the *Agrobacterium* nopaline synthase gene. LB and RB are the left and right borders.

Fig. 2 shows the sequence alignment of Wuschel-related HD domain proteins. Sequences were aligned with the Clustal V method using DNASTar™-MegAlign™ Program set to default parameters (Gap penalty=10; gap length penalty=0.2; delayed divergent sequences=30%; DNA transit weight=0.5). The HD domain and the hexapeptide are marked with line and dotted line, respectively. AtWOX2 (AY251392) (SEQ ID NO: 17), At1g46480 (NM_103605) (SEQ ID NO: 8), At3g11260 (NM_111961) (SEQ ID NO: 6), and PRS (BAB79446) (SEQ ID NO: 18) are Arabidopsis proteins deduced from mRNA sequences. Wus (NP_565429) (SEQ ID NO: 19), LeWUS (CAD61961) (SEQ ID NO: 21), and PhWUS (AAM90847) (SEQ ID NO: 20) are the

Arabidopsis Wuschel and its homologues of tomato and *Petunia*, respectively.

P0529E05.28 (BAB84412) (SEQ ID NO: 22) is a rice protein expressed specifically in the root quiescent centre. Genebank accession numbers are shown in the parenthesis.

Seq.ID3, Seq.ID5, Seq.ID7, Seq.ID9, Seq.ID13, Seq.ID15, Seq.ID17, Seq.ID19, Seq.ID21 and Seq.ID23 (SEQ ID NOs: 23-32) are polypeptide sequences predicted from the polynucleotide sequences disclosed in PCT publication WO 01/23575 A2 as sequence identification numbers 3, 5, 7, 9, 13, 15, 17, 19, 21 and 23 there.

Fig. 3 shows the evolutionary relationship of Wuschel-related HD domain proteins. The amino acid sequences of nineteen GhCIR1-related HD domain proteins were aligned by the Clustal V method using DNASTar™-MegAlign™ Program. The sequence identity and divergence output is shown in A and phylogenetic tree is shown in B.

Fig. 4 shows the morphology of *Arabidopsis* root regenerants. *Arabidopsis* root explants were co-cultured with AGL1 carrying pSK36-GhCIR1 (A), pSK36-AtCIR1 (B), pSK36-AtCIR2 (C) and control vector pSK36 (D). The regenerants were photographed using a dissecting microscope (Nikon™ SMZ 800™) after culturing for three weeks at 22°C in SR medium. The white arrows in panel A and B point somatic embryos. The black arrow in panel C points a typical regenerant derived from pSK36-AtCIR2.

Fig. 5 shows cotton somatic embryos derived from transformation of pSK36-GhCIR1. A cotton (Coker 312) pro-embryogenic suspension culture was co-cultured with AGL1 carrying pSK36-GhCIR1 (Panel A) or empty vector pSK36 (Panel B). The transformed cells were allowed to develop somatic embryos for three months under kanamycin selection in hormone-free medium and then photographed using a dissecting microscope. The black arrow in panel A points to a non-separated somatic embryo-like

structure and the long white arrows point to normal globular stage somatic embryos. The short white arrow points to an over-sized globular somatic embryo.

Fig. 6 shows the gene organization of T-DNA region in pX6-GFP and its derivatives. Promoters are shown as empty boxes. Protein coding regions are shown as solid black arrows. Gene transcriptional terminators are shown in gray boxes. CRE recombinase is under control of CaMV 35S-*LexA* hybrid promoter, which is inducible by estrogen (e.g., estradiol) that activates the VXE transcriptional activator. P35S::*GhCIR1* and P35S::*AtCIR2* gene cassettes were inserted into the *ApaI* site to create pX6-*GhCIR1* and pX6-*AtCIR2*, respectively. Pg10-90 is a synthetic promoter that drives the constitutive expression of VXE. LB and RB are the left and right borders.

Fig. 7 shows the use of *GhCIR1* and *AtCIR2* as transformation selection markers. Arabidopsis root explants were co-cultured with AGL1 carrying pX6-GFP (A), pX6-*GhCIR1* (B and C) or pX6-*AtCIR2* (D) for 3 days. Plantlets were regenerated from SR medium supplemented with 50mg/ml kanamycin (A); RI medium supplemented with 50mg/l kanamycin (B); or kanamycin-free RI medium (C and D). The plates were incubated at 22°C for 31 days.

Fig. 8 shows the chemically-induced excision of transformation selection marker. Arabidopsis root explants were co-cultured with AGL1 carrying pX6-*GhCIR1* for 3 days and plated in kanamycin-free RI medium. The plates were incubated at 22°C for 3 weeks. Green calli were transferred to RI medium with 10µM estradiol and incubated for 2 weeks. The shoots generated were photographed using a dissecting microscope (NikonTM SMZ800TM) with a florescent attachment. Panel A shows GFP florescence in the upper part of the shoot. Panel B shows a bright-field picture of the shoot.

Fig. 9 shows the gene organization of the T-DNA region in pER10-EGFP-HH. Panel A shows the T-DNA region of pER10-EGFP-HH. Promoters are shown as empty boxes. Protein coding regions are shown as solid black arrows. Gene transcriptional terminators are shown as gray boxes. LB and RB define the left and right borders. Panel B shows detailed information on the double-HA-hexahistidine (HH) tag. The double HA tag is dotted-underlined and the *PacI* site that can be used for gene fusion is underlined (DNA sequence: SEQ ID NO: 62, polypeptide sequence: SEQ ID NO: 63).

DETAILED DESCRIPTION OF THE INVENTION

A “gene” as used in this application means a segment of DNA, encoding a polypeptide or an RNA molecule. A gene can include regions preceding and following the coding DNA, which regions control or regulate the expression of the encoded polypeptide. A gene can include one or more introns.

“Isolated” as used in this application means that a molecule is separated from its naturally occurring environment. Isolated as used in this application does not necessarily mean a molecule is contained in isolation from all other molecules of the same or different kind, or disconnected to any other molecules; therefore, an isolated molecule may still be covalently bonded to a heterogeneous molecule to form a so-called fusion protein or an artificially created vector for cloning or expression and the fusion protein or artificially created vector still can be an isolated molecule of a different kind. In the context of an isolated nucleotide, “isolated” means that the nucleotide is separated at least from its naturally occurring macromolecular structure, for example, chromosome.

“Plant” as used in this application means only biological organisms classified in Plantae kingdom, and cells generated or derived from such organisms. Preferred plants in

this application are crops and Arabidopsis. More preferred plants in this application are cotton, rice, corn and Arabidopsis. Most preferred plants in this application are cotton and Arabidopsis.

“Non-supportive of plant regeneration” as used in this application means that there is no substantial plant regeneration from plant explants, and includes no plant regeneration from plant explants.

“Conservatively substituted homologue” as used in this application means that a protein in which one or more residues are substituted by amino acids with similar chemical or biophysical properties, e.g., hydrophobicity, ionic charges, side-chain structure.

“Transformation” or “transforming a cell” as used in this application broadly means any kind of transfer or introduction of a polynucleotide or gene to a cell, including conjugation, transformation, transfection and transduction. Preferred methods include transferring a polynucleotide or gene to a cell by *Agrobacterium*-mediated T-DNA conjugation or particle bombardment.

“Identity” or “percent identity” or “% identity” as used in this application means the identity calculated by Clustral V method using DNASTar™-MegAlign™ Program.

“Over-expression” as used in this application means that expression of a polynucleotide or gene occurs more than it would naturally occur in a system such as a cell, tissue or organism, including ectopic expression. The over-expression means, preferably at least 25%, more preferably at least 50% and the most preferably at least 100% more expression than the natural expression level.

A nuclear localizing signal (NLS) is a short amino acid sequence which acts like a 'tag' on the exposed surface of a protein. This sequence is used to confine the protein to the cell nucleus and to direct a newly synthesized protein into the nucleus via its recognition by cytosolic nuclear transport receptors. This signal often comprises positively charged amino acid, e.g., lysines or arginines. A protein can be directed to the nucleus by this NLS. An example of NLS is Pro-Pro-Lys-Lys-Lys-Arg-Lys-Val. Additional examples and sequence characteristics of nuclear localization signals can be found in numerous publications, e.g., Cokol et al., 2000, Tinland et al., 1992, which are incorporated by reference. Additional examples can also be found in NLSdb, which is a database of nuclear localization signals (NLSs) and of nuclear proteins targeted to the nucleus by NLS motifs, produced by Rajesh Nair and Phil Carter with Columbia University, New York. As of October 28, 2005, NLSdb contains 308 signals.

By functional screening of approximately 100,000 *Gossypium hirsutum* cDNAs, four candidate genes have been identified based on the ability to form either a large green callus or complete plantlet in a cytokinin-free medium. PCR amplification and sequencing of one of the candidate genes revealed that the T-DNA contained a cDNA of 1034 nucleotides (SEQ ID NO: 1) encoding an open reading frame of 244 amino acids (SEQ ID NO: 2). Re-cloning of the cDNA into pSK36 confirmed its cytokinin independent regeneration phenotype. As expected, multiple shoots were induced in a single callus and the plants were sterile or had low fertility. To overcome this problem, the gene was placed under control of the VXE inducible promoter in pER10 (Zou et al, 2002). When inducer estradiol was present in the 2ip-free medium, abundant shoots were

generated. The gene was therefore named *GhCIR1* (*Gossypium hirsutum Cytokinin Independent Regeneration*).

A BLAST search of the predicted protein showed that it shared a substantial homology to a group of proteins containing a N-terminal homeodomain (HD), a motif of approximately 60 amino acids, known to modulate DNA sequence recognition of transcription factors (Gehring et al., 1994). It is most similar to an Arabidopsis HD domain protein encoded by the putative gene At5g59340 (GeneBank accession NO. NP_200742.1), with the homology extending beyond the HD domain (Fig. 2). Using default parameters for NCBI Blast search and sequence alignment, the two polypeptide sequences shared 44% identity over the entire open reading frame. It has been reported that the gene contains a 436 nucleotide intron (underlined in SEQ ID NO: 3 in Figure 12) and the predicted protein sequence is shown in SEQ ID NO: 4. The mRNA is expressed in the apical part of zygotic embryos. Mutations in the gene cause only partial defects in embryo development (Haecker et al, 2004). However, there has been no information on its over-expression phenotype.

The HD domains of over a dozen more genes in the Arabidopsis genome share high homology with GhCIR1. Over-expression of the putative *GhCIR1* homologue (At5g59340; Genebank accession no. NP_200742) by the estrogen inducible VXE promoter or CaMV 35S promoter, as GhCIR1 did, induced an abundant regeneration of shoots in a cytokinin-free medium and significantly enhanced the production of shoots in a normal medium. A similar effect was observed with At3g11260 as well. On the other hand, At1g46480 (GeneBank No. NP_175145.2) was unable to induce shoot regeneration in a cytokinin-free medium and showed negligible effect to promote shoot regeneration in

medium containing 10 μ M 2ip (Table 1 and Table 2). We designated the two Arabidopsis genes, At5g59340 and At3g11260, as *AtCIR1* (*Arabidopsis thaliana* *Cytokinin Independent Regeneration 1*) and *AtCIR2*, respectively.

Although *GhCIR1*, *AtCIR1* and *AtCIR2* were all able to induce shoot regeneration in a cytokinin-free medium, plantlets induced by *AtCIR2* were distinctively different from those by *GhCIR1* and *AtCIR1*. Both *GhCIR1* and *AtCIR1* predominantly induced formation of spiky dark-green calli, which may develop into plantlets later. Microscopic examination of the calli regenerated from cytokinin-free medium revealed the spikes resembled somatic embryos. This was supported by the appearance of the root and cotyledon-like structures emerging from the opposite side of the spike (Fig. 4). Furthermore, no trichomes were observed in the putative cotyledons. In contrast, shoots regenerated by over-expressing *AtCIR2* in a cytokinin-free medium showed little difference to those transformed with pSK36 vector in a cytokinin-containing regeneration medium. Trichomes were seen in the first two leaves, indicating that they were true leaves and shoots were regenerated via the organogenesis pathway (Fig. 4).

To see if *GhCIR1* has a similar function in cotton as was found in Arabidopsis, we transformed a cotton suspension culture that is poorly embryogenic with pSK36-*GhCIR1*. In non-transformed cells or cells transformed with pSK36, globular somatic embryos may multiply or start to elongate after reaching certain size in a cytokinin-free medium. On the other hand, when the transformed cells with pSK36-*GhCIR1* were regenerated via somatic embryogenic pathway in a cytokinin-free medium, no normal somatic embryos were obtained, but those cells regenerated into embryo-like clusters in which individual embryos failed to separate or globular stage embryos that were much larger than usual

(Fig. 5A). These embryos usually died at a later stage. Nevertheless, over-expression of *GhCIR1* resulted in about 2-fold more kanamycin resistant somatic embryos compared to the control vector alone (Table 3). This result confirmed that over-expression of *GhCIR1* promotes cytokinin-independent plant regeneration across different plant species.

As over-expression of *GhCIR1*, *AtCIR1* or *AtCIR2* was able to induce shoot regeneration in a cytokinin-free medium, which is normally non-supportive of regeneration of Arabidopsis root explants, the three genes can be used as a selection marker, and even replace antibiotic or herbicide resistance genes. Inducible promoters can be used to overcome problems associated with plant abnormality resulting from over-expression of regeneration-enhancing genes. Alternatively, these regeneration enhancing genes can be removed using inducible excision systems, for example, estrogen-inducible gene excision system based on VXE inducible expression system and *Cre*-mediated site-specific recombination (Zuo et al., 2001). In this vector (pX6-GFP), activation of *Cre* recombinase by estrogen leads to the deletion of the DNA region between the *LoxP* sites (Fig. 6). The successful excision of the region can be confirmed by the expression of the GFP due to fusion of the GFP ORF with *G10-90* promoter.

We cloned *P35S::GhCIR1* and *P35S::AtCIR2* gene cassettes into the pX6-GFP vector (Fig. 6). Over-expression of *GhCIR1* or *AtCIR2* gene in arabidopsis root explants induced an abundant regeneration of shoots in a medium free of plant-active antibiotics such as kanamycin (Fig. 7). When the early-stage somatic embryos or shoots were transferred to a medium containing estradiol, green, florescent and normal looking shoots were produced (Fig. 8). This indicates that the gene cassettes between the GFP ORF and the *G10-90* promoter have been successfully excised. The dramatic drop in shoot

regeneration from root explants in RI+ medium also indicates that excision is a highly efficient process (Table 4).

The two putatively separate domains of GhCIR1 were independently over-expressed. Although the HD domains of Wuschel-related proteins are highly conserved, this domain alone did not show the cytokinin-independent regeneration or enhancement of regeneration efficiency phenotypes. Moreover, the fusion protein of the HD domain and an activation domain such as VP16 did not have the transcription activator function (Example 8; Tables 5 and 6). On the other hand, fusion of GhCIR1 HD domain with the putative activation domain of AtCIR2 fully re-constituted a protein with cytokinin-independent regeneration and transformation enhancing functions (Example 10, Table 7). Therefore, the entire coding sequence of GhCIR1, AtCIR1 or AtCIR2 is preferred for cytokinin-independent inducement of regeneration, or promotion of regeneration efficiency phenotypes.

Media Composition

G medium: 1x MS salts, 3% sucrose, 0.8% Agar, pH 5.7

B5 medium: 1x B5 salts, 1 X B5 vitamins, 3% sucrose, 0.5 g/L MES, pH 5.7

F1 medium: 1x B5 salts, 2% glucose, 0.5 g/L MES, 0.5 mg/L 2,4-D, 0.05 mg/L Kinetin, 0.2% Phytigel, pH 5.7.

F2 medium: F1, 20 mg/L Acetosyringone.

SR medium: 1x MS salts, 1% sucrose, 0.5 g/L MES, 2 mg/L 2ip, 0.15mg/L IAA, 200 mg/L Cefotaxim, 35 mg/L kanamycin, 0.2% Phytigel, pH 5.7

SR+ medium: SR medium, 10 μ M estradiol.

SR' medium: identical to SR except phytigel is replaced with 0.6 % lower melting agarose.

RI medium: 1X MS salts, 1% sucrose, 0.5 g/L MES, 0.15mg/L IAA, 200 mg/L

Cefotaxim, 35 mg/L kanamycin, 0.2% Phytigel, pH 5.7

RI+ medium: RI plus 10 μ M estradiol.

W solution: sterile water + 400mg/L Cefotaxim.

Example 1: Construction of a normalized cDNA library.

Total RNA was extracted from a pool of different cotton tissues which included somatic embryos, pro-embryogenic suspension cultures, roots, cotyledons, shoot tips and auxiliary shoot buds. Poly(A)+ RNA was isolated by PolyA-Tract™ mRNA purification system (Promoge™, UAS) and cDNA was synthesized with GeneRacer™ kit (Invitrogen™, USA). Approximately 0.75 μ g polyA+ RNA was ligated with GeneRacer™ RNA oligo (CGACUGGAGCACGAGGACACUGACAUGGACUGAAGGAGUAGAAA [SEQ ID NO: 33]) and the mRNA was converted to single stranded cDNA using GeneRacer™ oligo dT primer (GCTGTCAACGATACGCTACGTAACGGCATGACAGTG(T)18 [SEQ ID NO: 34]). After RNaseH digestion, single-stranded cDNA was subtracted once with 1 μ g biotinylated polyA+ mRNA (labelled with Ambion™ BrighStar™ psoralen-biotin labelling kit) from fully expanded leaves. Hybridized RNA/cDNA was removed with Streptavidin-magnetic beads (Promega™, USA) and the cDNA was recovered by isopropanol precipitation. A fraction of the cDNA was amplified with GR5-2 (TTTTGGCGCGCCGGACACTGACTTGGACTGAAGGAGTAGA [SEQ ID NO: 35]) and GR3-2N (TTTTTATTGCGGCCGCGCTACGTAACGGCATGACAGTG [SEQ ID

NO: 36]) primers and labeled with psoralen-biotin. The subtracted cDNA was further normalized with approximately 50ng psoralen-biotin labeled double-stranded cDNA. After hybridization, biotinylated cDNA was removed by Streptavidin-magnetic bead and the remaining cDNA was recovered by ethanol precipitation. The cDNA was amplified by *Pfu* ultra DNA polymerase (Stratagen™, USA) using primers GR5-2 and GR3-2N, digested with *AscI* and *NotI* and ligated to into similarly digested pSK36 (Fig. 1; Kojima et al., 1999). The cDNA library was amplified in XL10 and transferred to *Agrobacterium tumefaciens* strain AGL1 by electroporation.

Example 2: Functional screening of cDNA libraries

Root transformation and cDNA library screening were carried out essentially as described by Banno et al., 2001. Approximately 100,000 transformation events were screened and four candidate calli were identified. One of the calli initially had a dark green appearance but later developed into shoots. DNA was extracted from the plantlets and cDNA inserts were recovered by amplification with *Pfu* DNA polymerase using primers GR5-2 and GR3-2P (TTTTTAATTAAGCTACGTAACGGCATGACAGTG [SEQ ID NO: 37]). A PCR product of approximately 1.1 Kb was generated and cloned between the *AscI* and *PacI* sites in pSK36 and pER10 (see Zou et al, 2002), respectively creating pSK36-GhCIR1 and pER10-GhCIR1. Subsequent DNA sequencing and BLAST search proved GhCIR1 is a novel gene. SEQ ID NOs: 1 and 2 show the cDNA and predicted protein sequences of GhCIR1. SEQ ID.No. 57 shows the protein encoding region (ORF) of GhCIR1

Example 3: Cloning of Arabidopsis At5g59340, At3g11260 and At1g46480.

At5g59340 (SEQ ID NO: 3) was amplified from genomic DNA using primers 59340U (AAGGCGCGCCATGGAAAACGAAGTAAACGCAG [SEQ ID NO: 38]) and 59340L (CGTTAATTAATTACAACCCATTACCATTACTATC [SEQ ID NO: 39]). The PCR products were digested with *AscI* and *PacI* and cloned into the corresponding sites in pSK36 or pER10. Its predicted protein sequence is shown in SEQ ID NO: 4. The Arabidopsis At3g11260 (SEQ ID NO: 5) and At1g46480 (SEQ ID NO: 7) genes were amplified by RT-PCR of total RNA extracted from young seedlings (Columbia-O) using primers 11260U (AAGGCGCGCCAGTTGAGGACTTTACATCTGAACA [SEQ ID NO: 40]) and 11260L (AATTAATTAACCATGCATTGGAAAATATCT [SEQ ID NO: 41]); and 46480U (AGGCGCGCCAAAATGAAGGTTTCATGAGTTTTTCGAATG [SEQ ID NO: 42]) and 46480L (AGTTAATTAATCATCTCCCTTCAGGATGGAGAGG [SEQ ID NO: 43]). The cDNA was cloned in pSK36 and pER10 between the *AscI* and *PacI* sites (Fig. 1). The DNA sequences of the clones were verified by DNA sequencing. The predicted protein sequences are shown in SEQ ID NOs: 6 and 8.

Example 4: Effects of transient expression of GhCIR1, At5g59340, At3g11260 and At1g46480.

Root explants were transformed with pER10-GhCIR1, pER10-AtCIR1, pER10-AtCIR2 or pER10-AtESR3 by Agrobacterium mediated transformation. For comparison, pER10-ESR1 was also included in the experiment. The number of shoots regenerated in SR or RI medium that were with or without 10 μ M estradiol were counted 3-4 weeks after co-culture. The results are summarized in Table 1. Except pER10 and pER10-

At4g46480, shoots were observed in all constructs in RI medium containing inducer estradiol (RI+ medium). No shoots were seen in RI medium lacking estradiol irrespective of the constructs used. For cytokinin-independent regeneration, AtCIR2 is as efficient as ESR1, followed by GhCIR1 and AtCIR1. However, hormone and inducer contents can be individually optimized for the best regeneration result. In a normal regeneration medium supplemented with 10 μ M estradiol (SR1+), the expression of *GhCIR1*, *At5g59340* or *At3g11260* causes a significant increase in the number of shoots regenerated, compared with the control vector pER10. Each of *GhCIR1* and *At3g11260* produced approximately 15 times more shoots than pER10 vector alone. This was about 37%-62% better than *ESR1*. The effect of At1g46480 was weak under the condition used. As *At5g59340* and *At3g11260* were able to induce shoot regeneration in the absence of cytokinins, we named them AtCIR1 (*Arabidopsis thaliana* Cytokinin Independent Regeneration 1) and AtCIR2 (*Arabidopsis thaliana* Cytokinin Independent Regeneration 2), respectively.

Table 1. Effects on genes on plant regeneration in Arabidopsis root explants

Constructs	RI	SR	RI+	SR+
pER10	0	125 $\pm$ 15	0	87 $\pm$ 17
pER10-GhCIR1	0	127 $\pm$ 12	65 $\pm$ 1	1304 $\pm$ 190
pER10-At5g59340	0	116 $\pm$ 10	36 $\pm$ 4	438 $\pm$ 74
pER10-At3g11260	0	122 $\pm$ 2	112 $\pm$ 13	1537 $\pm$ 375
pER10-At1g46480	0	125 $\pm$ 14	1 $\pm$ 1	155 $\pm$ 16
pER10-ESR1	0	109 $\pm$ 12	110 $\pm$ 8	950 $\pm$ 73

Note: RI+: RI medium plus 10 μ M 17 β -estradiol (SigmaTM, USA); SR+: SR medium plus 10 μ M 17 β -estradiol. The number of shoots was counted on the 21st day (SR and SR+

media) and 30th day (RI and RI+ media). The numbers in the table are the total number of shoots per gram of root explants, averaged from three independent transformation experiments.

Example 5. Effects of constitutive expression of *GhCIR1*, *At5g59340* (*AtCIR1*), *At3g11260* (*AtCIR2*) in Arabidopsis.

AGL1 carrying constructs pSK36-GhCIR1, pSK36-AtCIR1 and pSK36-ARCIR2 were used to infect Arabidopsis root explants, which were plated on RI or SR medium. Significantly more rapid and efficient green callus regeneration was obvious from the 7-10th day in SR medium. Unlike in pER10, an abundant number of shoots was also observed in RI medium. GhCIR1 and AtCIR2 were of a similar strength but AtCIR1 was significantly weaker than GhCIR1 and AtCIR2 (Table 2). While the regenerants derived from the AtCIR2 construct showed a similar morphology to those derived from the vector-only control, the shoots derived from the GhCIR1 or AtCIR1 construct mostly appeared as big clusters. Microscopic examination of the calli revealed that the shoot clusters mainly comprised germinated somatic embryos because developing roots were observed at one end and trichome-less cotyledon-like structures were observed at the other end even when cytokinin (2ip) was added in the medium (Fig. 4).

Table 2

	RI	SR
pSK36	0	1423
pSK36-GhCIR1	541	2650
pSK36- At5g59340	224	2200
pSK36-At3g11260	762	2690

Note: The numbers in the table are the total number of shoots and dark-green calli per gram root explants on the 21st day (SR medium) or 30th day (RI medium).

Example 6: Effect of constitutive expression of *GhCIR1* in cotton suspension cultures

pSK36-GhCIR1 was transformed to a cotton suspension cell culture S4 by *Agrobacterium* mediated transformation (Ji and Cai, 2004). The transformed cells were selected and regenerated in a hormone-free MS medium with 100mg/l kanamycin and 300mg/l Cefotaxim. Although many somatic embryos died during the regeneration process, a substantial increase in the number of somatic embryos was observed (Table 3). Embryos developed abnormally when GhCIR1 was constitutively over-expressed (Fig. 5).

Table 3

Construct	Embryo No.
pSK36	14
pSK36-GhCIR1-a	24
pSK36-GhCIR1-b	30

Note: Approximately 0.2g suspension-cultured cells were divided to 100 sectors after co-culture. The numbers in the table are the total number of the sectors that produced somatic embryos. Only viable embryos at or after globular stages were counted. pSK36-GhCIR1-a and pSK36-GhCIR1-b were two experiments of the same condition.

Example 7: *GhCIR1* as a silent re-useable selection marker

To transfer the *P35S::GhCIR1* and *P35S::AtCIR2* to excision vector pX6-GFP, pSk36-GhCIR1 and pSk36-AtCIR2 were digested with *SpeI* and the restriction site was polished with T4 DNA polymerase. The gene cassettes plus portion of the *nptII* gene were released by *SacII* digestion and ligated with *SacII*-*ApaI* digested pX6-GFP vector (the *ApaI* site was polished by T4 DNA polymerase). The resultant constructs were named pX6-GhCIR1 and pX6-AtCIR2, respectively. Both were introduced to *Agrobacterium* strain AGL1.

Although no transformants were observed in explants infected with pX6-GFP as a control in the RI medium without kanamycin, a large number of shoots were seen in those infected with pX6-GhCIR1 and pX6-AtCIR2. The lack of kanamycin had no significant effect on transformation efficiency. The results are summarized in Table 4.

Table 4

	RI	RI-Kan ⁻	RI-Kan ⁺	SR
pX6-GFP	0	0	0	142±22
pX6-GhCIR1	126±44	116±12	19±4	1430±212
pX6-AtCIR2	374±54	330±44	3±3	2349±125

Note: RI-Kan⁻ is the same as RI except RI contains no kanamycin. RI-Kan⁺ is the same as RI-Kan⁻ except RI-Kan⁺ contains 10μM 17β-estradiol. The numbers in the table are the total number of shoots per gram of root explants, averaged from three independent transformation experiments.

Example 8: Effect of over-expression of truncated GhCIR1

The HD and the putative activation domains of GhCIR1 were separately cloned into pER10-EGFP-HH (Ji LH, unpublished data), which contains 2X HA-6X histidine epitope tag at the C-terminus (Fig. 9). A protein of interest can be tagged by cloning in-frame of the coding region at the unique *PacI* site. The HD domain was amplified by primers GR5-2 and GhCIR1NL (GGTTAATTAATTGAGCTCGATGATGATGGT [SEQ ID NO: 44]), and the C-terminal part was amplified by primers GhCIR1CU (AAGGCGCGCCAAAATGCCTGTTTTTCATCCTCCTCC [SEQ ID NO: 45]) and GhCIR1T (CCTTAATTAAGAAGAAATCAATGAAACGATGTTC [SEQ ID NO: 46]). The entire ORF was amplified by GR5-2 and GhCIR1T. All amplified fragments were digested with *AscI* and *PacI*, and cloned into the corresponding sites in pER10-EGFP-HH. The new constructs were named pER10-GhCIR1HH (the entire ORF), pER10-GhCIR1NHH (N-terminal) and pER10-GhCIR1CHH (C-terminal), respectively.

A *PacI* digested PCR fragment encoding the VP16 domain was amplified from pER10 using primers VP16U (CCTTAATTAACGCCCCCGACCGATGTCAGCCT [SEQ ID NO: 47]) and VP16L (TTTTAATTAACCCACCGTACTCGTCAATTCCAA [SEQ ID NO: 48]). That digested PCR fragment was then inserted at the *PacI* site of pER10-GhCIR1N, creating pER10-GHCIR1NVHH. The VP16 fragment was also inserted into the *PacI* site of pER10-GhCIR1HH, creating pER10-GhCIR1VHH.

pER10-GhCIR1HH efficiently induced shoot regeneration in RI medium as pER10-GhCIR1 does although a small drop in regeneration efficiency was observed. In contrast, no shoots were seen in root explants co-cultured with pER10-GhCIR1NHH or pER10-GhCIR1CHH. (Table 5). Moreover, addition of a VP16 activation domain to the HD domain (SEQ ID NOs: 9 and 10) did not restore the original function of GhCIR1.

This suggests that the HD domain alone is insufficient for the observed functions of GhCIR1, AtCIR1 and AtCIR2.

Table 5

	RI	SR	RI+	SR+
pER10	0	769±64	0	685±114
pER10-GhCIR1	0	800±70	372±51	2884±148
pER10-GhCIR1HH	0	749±81	187±17	2289±100
pER10-GhCIR1VHH	0	777±52	97±16	1439±97
pER10-GhCIR1NHH	0	710±58	0	684±72
pER10-GhCIR1NVHH	0	728±54	0	635±96
pER10-GhCIR1CHH	0	820±56	0	583±58

Note: RI+: RI medium plus 10µM estradiol; SR+: SR medium plus 10 µM estradiol. The number of shoots was counted on the 21st day (SR and SR+ media) and 30th day (RI and RI+ media). The numbers in the table are the total number of shoots per gram of root explants, averaged from three independent transformation experiments.

As GhCIR1 contains a basic peptide KRRGRP (SEQ ID NO: 59) (aa 138-143) that is similar to the N-terminal NLS (RK GK) (SEQ ID NO: 64) of *Agrobacterium virD2* (Tinland et al., 1992), it is likely that the removal of the C-terminal part of GhCIR1 affects the nuclear targeting of GhCIR1N and GhCIR1N-VP16. To confirm this possibility, we created a second construct expressing a fusion protein comprising the GhCIR1N (amino acids 1-96), the VP16 activation domain and the estrogen regulatory domain that contains a nuclear localization signal (GhCIR1N-VE: SEQ ID NO: 12). This construct, pSK36-GhCIR1N-VE, was made in pSK36 and expected to simulate pER10-GhCIR1 if the fusion protein was functional. Results in Table 6 support our previous data in Table 5 as no shoot regeneration was observed in RI+ medium with this construct.

Table 6

	RI	SR	RI+	SR+
pSK36	0	1208	0	1008
pSK36-GhCIR1N-VE	0	1349	0	1221
pSK36-GhCIR1N	0	1235	0	1028

Note: The numbers in the table are the total number of shoots regenerated from 1 gram of root explants.

Example 9. Root transformation method

Root transformation was performed according to Banno et al (2001). *Arabidopsis thaliana* Wassilewskija (WS) seeds were sterilized and sown on G medium and grown for 7 days at 22 °C under 16 hours day/night lighting cycles. Root cultures were prepared by transferring 7-day-old seedlings to B5 liquid medium and culturing for 15 days with shaking at 100 rpm. White portions of roots were excised as short explants and incubated in a callus-inducing medium (F1 medium) for 2-3 days and co-cultured in F2 medium with *Agrobacterium* strains for 2 days at 24°C under constant lighting. After thorough washing in W solution, the root explants were suspended in SR' medium and plated on selection medium (RI, RI+, SR or SR+). The plates were sealed with Micropore™ tape (3M™) and incubated at 22°C with 16 hours lighting cycles.

Example 10. Vital roles of the Non-HD domains of GhCIR1 and AtCIR2

To demonstrate the importance of the non-HD domain of the encoded polypeptides disclosed herein, we created chimeric constructs of *GhCIR1* and *AtCIR2*. pER10-GhCIR1N2HH (encoding aa 1-147 of GhCIR1) was made inserting *AscI-PacI* digested PCR product of primers GR5-2 and GhCIR1N2L (GATTAATTAAAGTTTTACCAATAGGCCTGC) (SEQ ID NO: 49) into the corresponding sites of pER10-GhCIR1HH. GhCIR1N1C2 (Fig. 22 and 23) is a fusion of HD-domain of GhCIR1 (aa 1-93) and the C-terminal domain (C2 fragment) of AtCIR2 (aa 107-192) and was made by inserting a *PacI*-digested PCR product of primers At3g11260U2 (TTTTAATTAATCCATCAACTAGAGATGTTTTTG) (SEQ ID NO: 50) and At3g11260L (AATTAATTAACCATGCATTGGAAAATATCT) (SEQ ID NO: 41) into the *PacI* site of pER10-GhCIR1N1HH. Similarly, pER10-GhCIR1N2C2 (Fig. 24 and 25) was made by inserting the C2 fragment into the *PacI* site of pER10-GhCIR1N2HH. The ORF of GhCIR1N2C2 (SEQ ID. 58) is shown in Fig. 28. pER10-GhCIR1N1C2 and pER10-GhCIR1N2HH were similar to the empty vector pER10 in inducing cytokinin-independent shoot regeneration or improving transformation efficiency in the presence of cytokinin, i.e., pER10-GhCIR1N1C2 or pER10-GhCIR1N2HH did not show any activity in inducing cytokinin-independent shoot regeneration or improving transformation efficiency in the presence of cytokinin. pER10-GhCIR1N2C2, however, not only improved transformation efficiency in the presence of 2ip (cytokinin) but also induced shoot regeneration in the absence of 2ip (cytokinin) (Table 7). An interesting difference between pER10-GhCIR1N1C2 and pER10-GhCIR1N2C2 (compare SEQ ID NOs:14 and 16) is a short motif KRRGRP (SEQ ID NO: 59) resembling the monopartite nuclear localization signal (Tinland et al., 1992).

Therefore, this motif may serve as a nuclear targeting signal in GhCIR1 and GhCIR1N2C2.

The C-termini of GhCIR1, AtCIR1 and AtCIR2 all contain a conserved hexapeptide with a consensus sequence of TLxLFP (SEQ ID NO: 65) near the C-terminal ends. This peptide is not found in other proteins except Wuschel-related HD-domain proteins. We noticed the presence of a threonine residue in the peptide. To see if phosphorylation of this residue has any effect on the function of Wuschel-related proteins, threonine 188 (T¹⁸⁸) of GhCIR1 was mutated by site-directed mutagenesis to encode either aspartic acid (D) that mimics phosphorylation, or alanine (A) that is resistant to phosphorylation. Primer pairs GR5-2 (SEQ ID NO: 35)/GhCIR1T188AU (SEQ ID NO: 51) and GhCIR1T188AL (SEQ ID NO: 52)/GR3-2P (SEQ ID NO: 37) were used for amplification of 5' and 3' parts of GhCIR1T188A, respectively, which were joined by PCR using primer pair GR-5-2/GR3-2P and cloned in pER10-EGFP-HH. pER10-GhCIR1T188D was similarly made using primers GhCIR1T188DU (SEQ ID NO: 53) and GhCIR1T188DL (SEQ ID NO: 54).

Strikingly, mutant GhCIR1 T¹⁸⁸D totally lost its function to enhance transformation in the presence of 2ip or induce shoot regeneration in the absence of 2ip while GhCIR1 T¹⁸⁸A was only marginally weakened compared to the wild-type GhCIR1 (Table 7). These findings suggest that a non-phosphorylated TLxLFP (SEQ ID NO: 65) motif is vital for functions of the Wuschel-related proteins. Therefore, the threonine residue in this hexapeptide may be substituted by non-negatively charged amino acids,

and substitution by a negatively charged amino acid, i.e., aspartic acid (D) or glutamic acid (E)) is not favored. Preferably the non-negatively charged amino acids are neutral amino acids, e.g., alanine, cysteine, methionine, isoleucine, leucine, valine, proline, glycine, phenylalanine and tryptophan. More preferably the non-negatively charged amino acids are phosphorylation-resistant amino acids with simple side chains, e.g. alanine, glycine, methionine, valine, isoleucine, leucine. The most preferred non-negatively charged amino acid is alanine.

As GhCIR1 loses its cytokinin-independent shoot regeneration activity by the phosphorylation of T¹⁸⁸, it can be inferred that over-expression of theonine/serine phosphatase will have a similar effect as over-expression of GhCIR1. An example of such phosphatases is one encoded by the arabidopsis *Poltergeist (Pol)* gene (Yu et al, 2003). Preferably, phosphatase mutants with deleted negative regulation domain are to be over-expressed.

A detailed examination of sequence alignment of Wuschel-related proteins (Fig. 2) revealed that the TLxLFP (SEQ ID NO: 65) motif is followed by only 3 amino acids in At1g46480 compared to 48 amino acids in GhCIR1, 55 amino acids in AtCIR1 and 54 amino acids in AtCIR2. To see if the lack of significant sequence extension accounted for the virtual lack of the cytokinin-independent shoot regeneration function of At1g46480, the C-terminal 45 amino acids of GhCIR1 was deleted by PCR using primers GR5-2 and GhCIR1 Δ 45L (TTAATTAAAACACCAGTCGGGTGCAATG) (SEQ ID NO: 55). Indeed, this construct, pER10-GhCIR1 Δ 43HH, contained only negligible activity in promoting shoot regeneration or cytokinin-independent shoot regeneration (Table 7). In addition, we noticed that a glutamic acid is conserved in GhCIR1, AtCIR1

and AtCIR2 in deleted region. However, converting this residue to alanine, which was achieved by one-step PCR using primers GR5-2 and GhCIR1E236AL (TTTTAATTAAGAAGAAATCAATGAAACGATGTGCCCCAGAA) (SEQ ID NO: 56), did not change the activity of GhCIR1 (Table 7, pER10-GhCIR1E236AHH). This result indicates that the size of this highly degenerate region is vital for both cytokinin-independent shoot regeneration and transformation enhancing activities.

Table 7

Construct	RI	SR	RI+	SR+
pER10	0	417±83	0	293±51
pER10-GhCIR1HH	0	507±74	354±45	1905±615
pER10-GhCIR1N1HH	0	463±85	0	327±15
pER10-GhCIR1N1VHH	0	417±201	0	220±50
pER10-GhCIR1N1C2	0	447±124	0	410±40
pER10-GhCIR1N2HH	0	370±46	0	297±107
pER10-GhCIR1N2VHH	0	395±65	0	273±85
pER10-GhCIR1N2C2	0	570±50	130±20	1727±196
pER10-GhCIR1CHH	0	493±183	0	245±55
pER10-GhCIR1T ¹⁸⁸ AHH	0	407±125	217±76	2040±478
pER10-GhCIR1T ¹⁸⁸ DHH	0	523±182	0	317±100
pER10-GhCIR1D43HH	0	537±86	7±6	563±76
pER10-GhCIR1E236AHH	0	490±95	400±41	2075±400

Note: RI+: RI medium plus 10μM estradiol; SR+: SR medium plus 10 μM estradiol. The number of shoots was counted on the 21st day (SR and SR+ media) and 30th day (RI and RI+ media). The numbers in the table are the total number of shoots per gram of root explants, averaged from three independent transformation experiments.

Example 11: Signature Sequence of Cytokinin-Independent Regeneration Proteins

From analysis of Fig.2 and data presented in Table 1, 2, 5 and 7, we concluded that functional proteins contained a C-terminal extension of about 50-60 residues (50 aa

in GhCIR1, 57 aa in AtCIR1 and 52 aa in AtCIR2) whereas non-functional proteins had less than 12 aa in the C-terminal extension (12 aa in PRS, 5 aa in At1g46480 and GhCIR1Δ45C). We also have found that GhCIR1, AtCIR1 and AtCIR2, which have cytokinin-independent shoot regeneration activity, share common elements including 1) a homeodomain near the N-terminus, 2) a nuclear localization signal (NLS), 3) a hexapeptide TLxLFP (SEQ ID NO: 65) that may be subjected to negative regulation by phosphorylation near the C-terminus and 4) a C-terminal extension to the TLxLFP (SEQ ID NO: 65) hexapeptide. These findings can be extrapolated into a signature sequence. A NLS sequence is not shown at any specific position in the signature sequence since it can be located anywhere in a protein as long as it does not disrupt the homeodomain near the N-terminus and the hexapeptide TLxLFP (SEQ ID NO: 65) near the C-terminus and is preferably exposed on the surface of the folded protein so as to be recognized by the cellular nuclear importing machinery.

A signature sequence for a protein that is able to enhance plant transformation and cytokinin-independent regeneration is defined as follows:

(X¹³⁻²³)RWNPT(KV)(DE)Q(IL)(STK)(MIL)L(ET)(SND)L(FY)(KR)(QEA)G(IL)
 RTP(ST)(AT)DQIQ(QK)I(ST)(GST)(RE)L(KRS)(AF)YG(KHT)IE(GS)KNVIFYWFQN
 HKAR(QE)RQK(QR)(KR)(X²⁹⁻¹⁰¹)(EK)(STACIGVLMNQP)L(PQ)LFP(LV)
 (X¹³⁻⁵⁷) (SEQ ID NO: 60)

wherein

X represents any amino acids and the numbers following the X represent the range of the size of the domain represented by X,

amino acids in a bracket represent alternatives for that position, for example, RWNPT(KV) (SEQ ID NO: 66) means that RWNPTK (SEQ ID NO: 67) or RWNPTV (SEQ ID NO: 68), and the protein comprises an NLS with the proviso that the NLS does not change the specified amino acids, excluding prior art sequences such as AtCIR2 and AtCIR1.

It is well known that a protein sequence could have been evolved considerably without changing its function. Various mathematical matrices have been made to define the weight of amino acid similarity. Today the most commonly used matrix for sequence alignment and homology search, e.g., BLAST search, is the BLOSUM series of matrices (Henikoff and Henikoff, 1992). Among them, BLOSUM 62 is often the default matrix for BLAST search. This mathematical model may be used for *in silico* evolution of the signature sequence that was described above. By substituting the less conserved positions with conserved residues that weigh at least one in the BLOSUM 62 substitution table, the signature sequence for a protein that is able to enhance plant transformation and cytokinin-independent regeneration may be expanded as follows.

(X¹³⁻²³)RWNPT(KEQRIMLV)(EDNKQ)Q(ILMV)(EKNSTQR)(MILV)L
 (DEKQST)(ADEHNST)L(FYW)(EQKR)(ADEKRQS)G(ILMV)
 RTP(AST)(AST)DQIQ(EKQR)I(AST)(AGST)
 (DEKQR)L(AEQKRST)(ASFWY)YG
 (EQHKTY)IE(AGST)KNV FYWFQNHKAR(DEQKR)RQK(EKQR)(EQKR) (X²⁹⁻¹⁰¹)(DEKQR)(STACIGVLMNQP)L(EPQKR)LFP(MILV)(X¹³⁻⁵⁷) (SEQ ID NO: 61),
 wherein

X represents any amino acids and the numbers following the X represent the range of the size of the domain represented by X,

amino acids in a bracket represent alternatives for that position, for example, RWNPT(KV) (SEQ ID NO: 66) means that RWNPTK (SEQ ID NO: 67) or RWNPTV (SEQ ID NO: 68), and

the protein comprises an NLS with the proviso that the NLS does not change the specified amino acids, excluding prior art sequences such as AtCIR2. It is also well known that a polypeptide, e.g., β -galactosidase (LacZ), green fluorescent protein (GFP), β -glucuronidase (GUS), etc, may be inserted into a protein to form a fusion protein without substantial changes in the protein's activity. Therefore, insertion of one or more of heterogeneous sequences in the signature sequence and/or addition to the signature sequence at the X domains is also contemplated with the proviso that such insertion does not change the specified amino acids.

Example 12: Effect of constitutive expression of *GhCIR1* and *AtCIR2* in cotton root explants

To further demonstrate that GhCIR1 enhances transformation efficiency in cotton, pX6-GhCIR1 was used to transform cotton root explants. Surface sterilized cotton seeds (Coker 312) were germinated in half-strength, sugar-free MS medium. Root explants of about 1 cm in length were pre-cultured for 3 days at 28°C in MS medium supplemented with 0.1mg/l kenetin and 0.1 mg/l 2,4-D. Agrobacterium culture (AGL1) harboring pX6-GhCIR1 or pEX6-GFP was washed and adjusted to about 0.3 OD₆₀₀ units in MS medium supplemented 0.1mg/l kenetin, 0.1mg/l 2,4-D and 100 μ M acetosyringon. Pre-cultured

root explants were soaked in the AGL1 suspensions for about 10 minutes. After briefly blotted dry in filter paper, the root explants were cultured at 28°C on solid MS medium supplemented with 1 mg/l zeatin riboside, 0.01mg/l NAA, 100mg/l kanamycin and 300 mg/l Cefotaxime. The number of kanamycin resistant calli at the end of the second month are scored and summarized in Table 8. In two independent experiments, over-expression of GhCIR1 lead to 2.3-fold and 4-fold more transformed calli, respectively. Furthermore, pX6-GhCIR1 calli were 2-3 times larger than the empty vector.

Table 8.

Experiment 1			
construction	Kan ^r Calli	Total Explants	Transformation efficiency
pX6	13	465	3
pX6-CIR1	44	642	7
Experiment 2			
construction			
pX6	9	416	2
pX6-CIR1	54	695	8

Note: Transformation efficiency is calculated as the percentage of kanamycin resistant calli over total number of explants used for co-cultured with Agrobacterium strains.

Thus, one aspect of the present invention relates to an isolated polynucleotide comprising SEQ ID NO: 1. Another aspect of the present invention relates to an isolated polynucleotide comprising SEQ ID NO: 15. Another aspect of the present invention relates to an isolated polynucleotide comprising a polynucleotide of at least 65% identity to SEQ ID NO: 1 wherein the polypeptide encoded by the polynucleotide promotes plant regeneration, induces plant regeneration in the absence of exogenous hormones, or increases plant transformation efficiency. Another aspect of the present invention relates

to an isolated polynucleotide comprising a polynucleotide of at least 65% identity to SEQ ID NO: 15 wherein the polypeptide encoded by the polynucleotide promotes plant regeneration, induces plant regeneration in the absence of one or more exogenous hormones, or increases plant transformation efficiency. Another aspect of the present invention relates to an isolated polynucleotide comprising a polynucleotide encoding SEQ ID NO: 2. Another aspect of the present invention relates to an isolated polynucleotide comprising a polynucleotide encoding SEQ ID NO: 16. Another aspect of the present invention relates to an isolated polynucleotide comprising a polynucleotide encoding a polypeptide that promotes plant regeneration, induces plant regeneration in the absence of one or more exogenous hormones, or increases plant transformation efficiency. The polypeptide preferably contains a signatures sequence:

(X¹³⁻²³)RWNPT(KEQRIMLV)(EDNKQ)Q(ILMV)(EKNSTQR)(MILV)L
 (DEKQST)(ADEHNST)L(FYW)(EQKR)(ADEKRQS)G(ILMV)
 RTP(AST)(AST)DQIQ(EKQR)I(AST)(AGST)
 (DEKQR)L(AEQKRST)(ASFWY)YG
 (EQHKTY)IE(AGST)KNVFYWFQNHKAR(DEQKR)RQK(EKQR)(EQKR) (X²⁹⁻¹⁰¹)(DEKQR)(STACIGVLMNQP)L(EPQKR)LFP(MILV)(X¹³⁻⁵⁷) (SEQ ID NO: 61),
 wherein

X represents any amino acids and the numbers following the X represent the range of the size of the domain represented by X,

amino acids in a bracket represent alternatives for that position, for example, RWNPT(KV) (SEQ ID NO: 66) means that RWNPTK (SEQ ID NO: 67) or RWNPTV (SEQ ID NO: 68), and

the protein comprises an NLS with the proviso that the NLS does not change the specified amino acids and the protein excludes prior art sequences such as AtCIR2. A preferred signature sequence is

(X¹³⁻²³)RWNPT(KV)(DE)Q(IL)(STK)(MIL)L(ET)(SND)L(FY)(KR)(QEA)G(IL)
RTP(ST)(AT)DQIQ(QK)I(ST)(GST)(RE)L(KRS)(AF)YG(KHT)IE(GS)KNVIFYWFQN
HKAR(QE)RQK(QR)(KR)(X²⁹⁻¹⁰¹)(EK)(STACIGVLMNQP)L(PQ)LFP(LV)
(X¹³⁻⁵⁷) (SEQ ID NO: 60),

wherein

X represents any amino acids and the numbers following the X represent the range of the size of the domain represented by X,

amino acids in a bracket represent alternatives for that position, for example, RWNPT(KV) (SEQ ID NO: 66) means that RWNPTK (SEQ ID NO: 67) or RWNPTV (SEQ ID NO: 68), and the protein comprises an NLS with the proviso that the NLS does not change the specified amino acids and the protein excludes prior art sequences such as AtCIR2. Further, insertion of one or more of heterogeneous sequences such as reporter enzymes, e.g., beta-galactosidase (LacZ), green florescent protein (GFP), beta-glucuronidase (GUS), etc, in the signature sequence and/or addition to the signature sequence at the X domains is also contemplated with the proviso that such insertion does not change the specified amino acids.

The isolated polynucleotides may comprise an operably linked promoter active in a plant. The promoter may be an inducible or constitutive promoter. The isolated polynucleotides may comprise other regulatory regions such as an enhancer, or a repressor binding site. The isolated polynucleotides may comprise fusion partners for

various purposes such as over expression, increased stability, targeting to a specific type of cells or intracellular compartments, or tagging for easier identification or purification. The isolated polynucleotide may be a gene. The isolated polynucleotide may be a vector, or when the isolated polynucleotide is a gene, the polynucleotide may be incorporated into a vector. The vector can be a plasmid. The isolated polynucleotide may be an expression vector, or when it is a gene, it can be incorporated into an expression vector so that the polypeptide encoded by the isolated polynucleotide can be produced in a cell, preferably bacterial cell, plant cell, or both. The present invention also relates to a cell that has been transformed (transformed in a broad sense, including transformation, transduction, transfection and conjugation) with the isolated polynucleotide or when the isolated polynucleotide is a gene, the gene or a vector incorporating the gene. The transformed cells may further comprise a second vector. The first and second vector may comprise a binary vector system.

Another aspect of the present invention relates to an isolated polypeptide comprising SEQ ID NO: 2. Another aspect of the present invention relates to an isolated polypeptide comprising SEQ ID NO: 16. Another aspect of the present invention relates to an isolated polypeptide of at least 65% identity to SEQ ID NO: 2, wherein said polypeptide promotes plant regeneration, induces plant regeneration in the absence of exogenous hormones, or increases plant transformation efficiency. Another aspect of the present invention relates to an isolated polypeptide of at least 65% identity to SEQ ID NO: 16, wherein said polypeptide promotes plant regeneration, induces plant regeneration in the absence of exogenous hormones, or increases plant transformation efficiency. Another aspect of the present invention relates to an isolated polypeptide that

promotes plant regeneration, induces plant regeneration in the absence of exogenous hormones, or increases plant transformation efficiency. The polypeptide preferably contains a signatures sequence:

(X¹³⁻²³)RWNPT(KEQRIMLV)(EDNKQ)Q(ILMV)(EKNSTQR)(MILV)L
 (DEKQST)(ADEHNST)L(FYW)(EQKR)(ADEKRQS)G(ILMV)
 RTP(AST)(AST)DQIQ(EKQR)I(AST)(AGST)
 (DEKQR)L(AEQKRST)(ASFWY)YG
 (EQHKTY)IE(AGST)KNVIFYWFQNHKAR(DEQKR)RQK(EKQR)(EQKR) (X²⁹⁻¹⁰¹)(DEKQR)(STACIGVLMNQP)L(EPQKR)LFP(MILV)(X¹³⁻⁵⁷) (SEQ ID NO: 61),

wherein

X represents any amino acids and the numbers following the X represent the range of the size of the domain represented by X,

amino acids in a bracket represent alternatives for that position, for example, RWNPT(KV) (SEQ ID NO: 66) means that RWNPTK (SEQ ID NO: 67) or RWNPTV (SEQ ID NO: 68), and

the protein comprises an NLS with the proviso that the NLS does not change the specified amino acids and the protein excludes prior art sequences such as AtCIR2. A more preferred signature sequence is

(X¹³⁻²³)RWNPT(KV)(DE)Q(IL)(STK)(MIL)L(ET)(SND)L(FY)(KR)(QEA)G(IL)
 RTP(ST)(AT)DQIQ(QK)I(ST)(GST)(RE)L(KRS)(AF)YG(KHT)IE(GS)KNVIFYWFQNHKAR(QE)RQK(QR)(KR)(X²⁹⁻¹⁰¹)(EK)(STACIGVLMNQP)L(PQ)LFP(LV)
 (X¹³⁻⁵⁷) (SEQ ID NO: 60),

wherein

X represents any amino acids and the numbers following the X represent the range of the size of the domain represented by X, amino acids in a bracket represent alternatives for that position, for example, RWNPT(KV) (SEQ ID NO: 66) means that RWNPTK (SEQ ID NO: 67) or RWNPTV (SEQ ID NO: 68), and the protein comprises an NLS with the proviso that the NLS does not change the specified amino acids and the protein excludes prior art sequences such as AtCIR2. Further, insertion of one or more of heterogeneous sequences such as reporter enzymes, e.g., beta-galactosidase (LacZ), green florescent protein (GFP), beta-glucuronidase (GUS), etc, in the signature sequence and/or addition to the signature sequence at the X-domains is also contemplated with the proviso that such insertion does not change the specified amino acids.

Another aspect of the present invention relates to a method of selecting a polynucleotide that enhances plant regeneration from non-Arabidopsis mRNAs comprising the steps of constructing a cDNA library from cells, wherein the cDNAs in said cDNA library can be optionally normalized and are regulated by a plant-active promoter; transforming Arabidopsis explants with vectors comprising said normalized cDNA library; culturing said Arabidopsis explants in a plant-cell-culture medium that is non-supportive of plant regeneration with conventional vectors, wherein said plant-cell-culture medium can be free of one or more plant hormones or with only trace amount of them; selecting Arabidopsis explants regenerating in said plant-cell-culture medium; identifying the nucleotide sequence of the polynucleotide that was transferred from said cDNA library to said Arabidopsis explants regenerating in said plant-cell-culture medium

and makes said Arabidopsis explants regenerate in said plant-cell-culture medium by sequencing a PCR product amplified from said Arabidopsis explants regenerating in said plant-cell-culture medium using a primer from a group consisting of SEQ ID NOs. 35, 37, 38, 39, 40, 41, 42 and 43; and isolating a polynucleotide of said nucleotide sequence from a cell, or synthesizing a polynucleotide of said nucleotide sequence.

Another aspect of the present invention relates to a method of selecting a plant cell transformant comprising the steps of transforming plant cells with vectors, wherein said vector comprises the isolated polynucleotide described above, and a promoter that is active in a plant and operably linked to said isolated polynucleotide; culturing said plant cells in a plant-cell-culture medium that is non-supportive of plant regeneration with conventional vectors; and selecting regenerating plant cells. The vector may further comprise a polynucleotide of interest.

Another aspect of the present invention relates to a method of improving plant transformation efficiency comprising the steps of transforming plant cells with vectors, wherein said vector comprises the isolated polynucleotide described above, and a promoter that is active in a plant and operably linked to said isolated polynucleotide; culturing said plant cells in a plant-cell-culture medium that is optionally cytokinin-free; and selecting regenerating plant cells.

Another aspect of the present invention relates a method for improving plant regeneration or transformation efficiency by over-expression of a theonine/serine phosphatase in vectors, wherein said vector comprises the isolated theonine/serine phosphatase gene or unregulated mutants of the genes, and a promoter that is active in a plant and operably linked to said isolated genes; culturing said plant cells in a plant-cell-

culture medium that is optionally free of cytokinins; and selecting regenerating plant cells. The serine/threonine phosphatase is preferably the one encoded by arabidopsis Poltergeist (Pol) gene (Yu et al, 2003). The serine/threonine phosphatase polypeptide can be one sharing at least 65% identity with the arabidopsis Poltergeist. Preferably, phosphatase mutants with deleted negative regulation domain are to be over-expressed.

It will be appreciated that the polynucleotides of the present invention may be used to induce plant regeneration in the absence of one or more exogenous hormones as well as promote plant regeneration, or increase plant transformation efficiency using a normal regeneration or transformation methods.

The transformation of a cell, including a cell comprising explants, with the isolated polynucleotide of the present invention (in other words, the incorporation or transfer of the isolated polynucleotide of the present into a cell) can be achieved by a method well known one skilled in the art. Especially for a plant cell, such a transformation can be achieved by *Agrobacterium*-mediated T-DNA conjugation or particle gun bombardment.

A wide variety of plant cells can be used to practice the present invention. Particularly preferred are *Arabidopsis* and cotton; however, rice, corn (e.g. maize), beans (e.g. soybeans), wheat, barley, sugarbeet, oil palm, sunflower and others can also be used. A wide variety of plant cells can be used to practice the present invention. Preferred are explants from various origins, and especially, root explants are preferred in selecting a polynucleotide that enhances plant regeneration.

A wide variety of promoters can be used to practice the present invention. For a vector to transfer the isolated polynucleotide between two different hosts, two separate

promoters, respectively active only in one of the two hosts, can be used. The promoter can be a chemically inducible promoter, or a constitutive promoter. The isolated polynucleotide and promoter may be chemically excisable.

The plant-cell-culture medium used to practice the present invention may be free of any antibiotics or herbicides, or any plant hormones.

The invention also contemplates a complement sequence of any of the nucleic acid sequences disclosed above; the use of the complement sequences in producing the nucleic acid sequences disclosed above in a biological system; and the use of the complement sequences in probing nucleic acid sequences related to the sequences disclosed above.

Therefore, the invention can be embodied in the following ways.

1. An isolated polynucleotide comprising
 - a. SEQ ID NOs: 57 or 15; or a complement thereof;
 - b. a polynucleotide of at least 65% identity to SEQ ID NOs: 57 or 15; or a complement thereof, wherein the polypeptide encoded by the polynucleotide of at least 65% identity to SEQ ID NOs: 57 or 15; or a complement thereof promotes plant regeneration, induces plant regeneration in the absence of one or more exogenous hormones, or increases plant transformation efficiency;
 - c. a polynucleotide encoding SEQ ID NOs: 2 or 16; or a complement thereof;
 - d. a polynucleotide encoding a polypeptide comprising a signature sequence:

(X¹³⁻²³)RWNPT(KEQRIMLV)(EDNKQ)Q(ILMV)(EKNSTQR)(MILV)L
 (DEKQST)(ADEHNST)L(FYW)(EQKR)(ADEKRQS)G(ILMV)
 RTP(AST)(AST)DQIQ(EKQR)I(AST)(AGST)
 (DEKQR)L(AEQKRST)(ASFWY)YG
 (EQHKTY)IE(AGST)KNVIFYWFQNHKAR(DEQKR)RQK(EKQR)(EQKR) (X<sup>29-
 101</sup>)(DEKQR)(STACIGVLMNQP)L(EPQKR)LFP(MILV)(X¹³⁻⁵⁷) (SEQ ID NO: 61),

wherein

X represents any amino acids,

the numbers following the X represent the range of the size of the domain
 represented by X,

amino acids in each bracket represent alternatives for the respective single amino
 acid position represented by each bracket,

the polypeptide comprises an NLS with the proviso that the NLS does not change
 the specified amino acids, and

AtCIR1 and AtCIR2 are excluded from the polypeptide comprising the signature
 sequence.

2. The isolated polynucleotide of item 1, further comprising an operably linked promoter, wherein said promoter is active in a plant.
3. The isolated polynucleotide of item 1, wherein said isolated polynucleotide is a gene.
4. The isolated polynucleotide of item 2, wherein said promoter is a constitutive promoter.

5. The isolated polynucleotide of item 2, wherein said promoter is an inducible promoter.
6. A vector which comprises the polynucleotide of item 1.
7. The vector of item 6, wherein said vector is a plasmid.
8. The vector of item 6, further comprising a promoter, wherein said promoter is operably linked to said polynucleotide.
9. The vector of item 8, wherein said promoter is active in a plant cell.
10. A binary vector system comprising the vector of item 9.
11. A bacterial cell comprising the vector of 7.
12. A plant cell comprising the vector of 9.
13. The isolated polynucleotide of item 1, wherein the signature sequence is
 (X^{13-23}) RWNPT(KV)(DE)Q(IL)(STK)(MIL)L(ET)(SND)L(FY)(KR)(QEA)G(IL)
RTP(ST)(AT)DQIQ(QK)I(ST)(GST)(RE)L(KRS)(AF)YG(KHT)IE(GS)KNVIFYWFQN
HKAR(QE)RQK(QR)(KR)(X^{29-101})(EK)(STACIGVLMNQP)L(PQ)LFP(LV)
 (X^{13-57}) (SEQ ID NO: 60).
14. The nucleotide of item 1, wherein the polypeptide comprises an additional sequence in the signature sequence at the X domains.
15. An isolated polypeptide comprising
 - a. SEQ ID NOs: 2 or 16;
 - b. a polypeptide of at least 65% identity to SEQ ID NOs: 2 or 16,
wherein the polypeptide of at least 65% identity to SEQ ID NOs: 2 or 16 promotes plant regeneration, induces plant regeneration in the

absence of one or more exogenous hormones or in a trace amount of them, or increases plant transformation efficiency;

c. a signature polypeptide:

(X¹³⁻²³)RWNPT(KEQRIMLV)(EDNKQ)Q(ILMV)(EKNSTQR)(MILV)L
(DEKQST)(ADEHNST)L(FYW)(EQKR)(ADEKRQS)G(ILMV)
RTP(AST)(AST)DQIQ(EKQR)I(AST)(AGST)
(DEKQR)L(AEQKRST)(ASFWY)YG
(EQHKTY)IE(AGST)KNVIFYWFQNHKAR(DEQKR)RQK(EKQR)(EQKR) (X²⁹⁻¹⁰¹)(DEKQR)(STACIGVLMNQP)L(EPQKR)LFP(MILV)(X¹³⁻⁵⁷) (SEQ ID NO: 61),

X represents any amino acids,

the numbers following the X represent the range of the size of the domain represented by X,

amino acids in each bracket represent alternatives for the respective single amino acid position represented by each bracket,

the polypeptide comprises an NLS with the proviso that the NLS does not change the specified amino acids, and

AtCIR1 and AtCIR2 are excluded from the polypeptide comprising the signature sequence.

16. The isolated polypeptide of item 15, wherein the signature sequence is preferably

(X¹³⁻²³)RWNPT(KV)(DE)Q(IL)(STK)(MIL)L(ET)(SND)L(FY)(KR)(QEA)G(IL)
RTP(ST)(AT)DQIQ(QK)I(ST)(GST)(RE)L(KRS)(AF)YG(KHT)IE(GS)KNVIFYWFQNHKAR(QE)RQK(QR)(KR)(X²⁹⁻¹⁰¹)(EK)(STACIGVLMNQP)L(PQ)LFP(LV)

(X¹³⁻⁵⁷) (SEQ ID NO: 60).

17. The isolated polypeptide of item 16, wherein the isolated polypeptide comprises an additional sequence in the signature sequence at the X domains.

18. A method of selecting a polynucleotide from species other than *Arabidopsis*, comprising the steps of

- a. constructing a cDNA library from cells, wherein the cDNAs in said cDNA library can be expressed in a plant cell,
- b. transforming *Arabidopsis* explants with vectors in said cDNA library,
- c. culturing said *Arabidopsis* explants in a plant-cell-culture medium that is non-supportive of regeneration with conventional vectors,
- d. selecting *Arabidopsis* explants regenerating in said plant-cell-culture medium,
- f. identifying the nucleotide sequence of the polynucleotide that was transferred from said cDNA library to said *Arabidopsis* explants regenerating in said plant-cell-culture medium and enables said *Arabidopsis* explants to regenerate in said plant-cell-culture medium, and
- g. isolating a polynucleotide of said nucleotide sequence from a cell, or synthesizing a polynucleotide of said nucleotide sequence.

19. The method of item 18, wherein *Arabidopsis* explants are transformed with said vectors in said cDNA library by *Agrobacterium*-mediated T-DNA conjugation.

20. The method of item 18, wherein *Arabidopsis* explants are transformed with said vectors in said cDNA library by particle gun bombardment

21. The method of item 18, wherein said plant-cell-culture medium does not contain no cytokinins or trace amount of cytokinins.
22. The method of item 18, wherein said plant-cell-culture medium does not contain 2ip.
23. The method of item 18, wherein said Arabidopsis explants are Arabidopsis root explants.
24. A method of selecting a plant cell transformant comprising the steps of
- a. transforming plant cells with vectors, wherein said vector comprises
 - i. a polynucleotide encoding a polypeptide that enables a transformant to regenerate in a medium deficient of one or more hormones, and
 - ii. a promoter that is active in a plant and operably linked to said isolated polynucleotide,
 - b. culturing said plant cells in a plant-cell-culture medium that is deficient of one or more hormones, and
 - c. selecting regenerating plant cells.
25. The method of item 24, wherein the polypeptide that enables a transformant to regenerate in said medium is
- a. a polypeptide of at least 65% identity to SEQ ID NOs: 2 or 16; or
 - b. a signature polypeptide:
- (X¹³⁻²³)RWNPT(KEQRIMLV)(EDNKQ)Q(ILMV)(EKNSTQR)(MILV)L
 (DEKQST)(ADEHNST)L(FYW)(EQKR)(ADEKRQS)G(ILMV)
 RTP(AST)(AST)DQIQ(EKQR)I(AST)(AGST)

(DEKQR)L(AEQKRST)(ASFWY)YG

(EQHKTY)IE(AGST)KNVIFYWFQNHKAR(DEQKR)RQK(EKQR)(EQKR) (X²⁹⁻

¹⁰¹)(DEKQR)(STACIGVLMNQP)L(EPQKR)LFP(MILV)(X¹³⁻⁵⁷) (SEQ ID NO: 61),

X represents any amino acids,

the numbers following the X represent the range of the size of the domain

represented by X,

amino acids in each bracket represent alternatives for the respective single amino acid position represented by each bracket, and

the polypeptide comprises an NLS with the proviso that the NLS does not change the specified amino acids.

26. The method of item 24, wherein said vector further comprises a polynucleotide of interest.

27. The method of item 24, wherein said promoter is a chemically inducible promoter.

28. The method of item 24, wherein said promoter is a constitutive promoter.

29. The method of item 24, wherein said isolated polynucleotide and promoter are chemically excisable.

30. The method of item 24, wherein said plant-cell-culture medium does not contain any antibiotics or herbicides.

31. The method of item 24, wherein said plant-cell-culture medium does not contain one or more hormones.

32. A method of improving plant transformation efficiency comprising the steps of

a. transforming plant cells with vectors, wherein said vector comprises:

- i. a polynucleotide encoding a polypeptide that enables a transformant to regenerate in a medium that is deficient of one or more hormones or increases regeneration rate in a regular plant-cell-culture medium, and
- ii. a promoter that is active in a plant and operably linked to said isolated polynucleotide,

- b. culturing said plant cells in a plant-cell-culture medium, which is optionally supplemented with cytokinins, and optionally supplemented with herbicides, antibiotics or both, and,
- c. selecting regenerating plant cells.

33. The method of item 32, wherein said polypeptide is

- a. a polypeptide of at least 65% identity to SEQ ID NOs: 2 or 16; or
- b. a signature polypeptide:

(X¹³⁻²³)RWNPT(KEQRIMLV)(EDNKQ)Q(ILMV)(EKNSTQR)(MILV)L
 (DEKQST)(ADEHNST)L(FYW)(EQKR)(ADEKRQS)G(ILMV)
 RTP(AST)(AST)DQIQ(EKQR)I(AST)(AGST)
 (DEKQR)L(AEQKRST)(ASFWY)YG
 (EQHKTY)IE(AGST)KNVIFYWFQNHKAR(DEQKR)RQK(EKQR)(EQKR) (X²⁹⁻¹⁰¹)(DEKQR)(STACIGVLMNQP)L(EPQKR)LFP(MILV)(X¹³⁻⁵⁷) (SEQ ID NO: 61),

X represents any amino acids,

the numbers following the X represent the range of the size of the domain represented by X,

amino acids in each bracket represent alternatives for the respective single amino acid position represented by each bracket, and

the polypeptide comprises an NLS with the proviso that the NLS does not change the specified amino acids.

34. The method of item 32, wherein the polypeptide that enables a transformant to regenerate in a medium that is deficient of one or more hormones or increases regeneration rate in a regular plant-cell-culture medium is a serine or threonine phosphatase.
35. The method of item 32, wherein said plant is cotton, maize, soybeans, rice, wheat, barley, sugarbeet, and oil palm.
36. The method of item 32, wherein said promoter is a chemically inducible promoter.
37. The method of item 32, wherein said promoter is a constitutive promoter.
38. The method of item 32, wherein said isolated polynucleotide and promoter are chemically excisable.
39. The method of item 32, wherein said plant-cell-culture medium contains antibiotics or herbicides.
40. The method of item 32, wherein said regular plant-cell-culture medium contains cytokinin.
41. The method of item 32, wherein said transformed cells are regenerated via somatic embryogenesis.
42. The method of item 32, wherein said transformed cells are regenerated via organogenesis.

43. The isolated polynucleotide of items 1, wherein the polypeptide encoded by said polynucleotides recited by d promotes plant regeneration, induces plant regeneration in the absence of one or more exogenous hormones or in the trace amount of them, or increases plant transformation efficiency.
44. The isolated polypeptide of item 15, wherein the polypeptide recited by c above promotes plant regeneration, induces plant regeneration in the absence of one or more exogenous hormones or in the trace amount of them, or increases plant transformation efficiency.
45. A method of improving plant transformation efficiency comprising the steps of
- a. transforming plant cells with vectors, wherein said vector comprises:
 - i. the isolated polynucleotide encoding a serine/threonine phosphatase, and
 - ii. a promoter that is active in a plant and operably linked to said isolated polynucleotide,
 - b. culturing said plant cells in a plant-cell-culture medium, which is optionally supplemented with cytokinins, and optionally supplemented with herbicides, antibiotics or both, and,
 - c. selecting regenerating plant cells.
46. The method of item 45, wherein said theonine/serine phosphatase is the arabidopsis Poltergeist.
47. The method of item 45, wherein said theonine/serine phosphatase shares at least 65% identity to the arabidopsis Poltergeist.

48. The methods of item 46 or 47, wherein said phosphatase is preferably a mutant deleted in negative regulation domain.

The above items are not meant to be an exhaustive list of possible claims. For example, a polypeptide of at least 65% identity to SEQ ID NOs: 2 and 16 having the signature sequence is also contemplated. Further, the Markush group style representation of possible alternatives in a bracket is meant to represent individual species covered by the formulation, and should not be narrowly construed as if the alternatives in a bracket are meant to be a range.

References

- Ainley, W. M. and Key, J. L. (1990). *Plant Mol. Biol.* 14:949-966.
- Aoyama, T. and Chua, N-H (1997). *The Plant J.* 11:605-612.
- Aoyama, T., Dong, C-H, Wu, Y., Carabelli, M., Sessa, G., Ruberti, I, Morelli, G and Chua N-H. (1995). *Plant Cell* 7:1773-1785.
- Imamura, A., Kiba, T., Tajima, Y., Yamashino, T. and Mizuno T. (2003). *Plant Cell Physiol.* 44(2): 122–131.
- Banno, H, Ikeda, Y, Niu, Q. W. and Chua, N-H. (2001). *Plant Cell.* 13:2609–2618.
- Barry, G. F., Rogers, S. G., Fraley, R. T. and Brand, L. (1984). *Proc. Natl. Acad. Sci. USA* 81:4776-4780.
- Beato M (1989). *Cell* 56:335-344.
- Bryant, J. and Leather, S. (1992). *Trends Biotechnol.* 10:274-275.
- Chang, C. and Shockey, J. A. (1999). *Curr. Opin. Plant Biol.* 2:352-358.
- Chuck, G., Lincoln, C. and Hake, S. (1996). *The Plant Cell* 8:1277-1289.
- Clough, S.J. and Bent, A. (1998). *Plant J.* 16:735-743.
- Cokol, M., Nair R., and Rost, B. (2000). *EMBO Reports* 1:411-415.
- Ebinuma, H., Sugita, K., Matsunaga, E. and Yamakado, M. (1997). *Proc. Natl. Acad. Sci. USA* 94:2117-2121.
- Endrizzi, et al., 1996, *Plant Journal* 10: 967-979;
- Faiss, M., Zalubilova, J., Strnad, M and Schmulling, T. (1997). *The Plant Journal* 12:401-415.
- Flavell, R. B., Dart, E., Fuchs, R. L. and Fraley, R. B. (1992). *Bio/Technology* 10:141-144.

- Gatz, C. (1996). *Curr. Opin. Biotechnol.* 7:168-172.
- Gatz, C., Froberg, C. and Wendenburg, R. (1992). *Plant J.* 2:397-404.
- Geer, L.Y., Domrachev, M., Lipman, D.J., Bryant, S.H. (2002). *Genome Res.* 2002 12:1619-23.
- Gehring, W.J., Affolter, M., Burglin, T. (1994). *Trends Biotechnol.* 10:382.
- Haecker, A., Gross-Hardt R, Geiges, B., Sarkar, A, Breuninger, H., Herrmann, M., Laux, T. (2004). *Development.* 131:657-68.
- Henikoff, S., and Henikoff, J.G. (1992). Amino acid substitution matrices for protein blocks. *Proc. Natl. Acad. Sci. USA* 89:10915-10919.
- Hwang, I. and Sheen, J. (2001). *Nature* 413:383-9.
- Ji, L. and Cai, L. (2004). A method for high efficiency transformation and regeneration of plant suspension cell cultures. (submitted for PCT patent application)
- Kakimoto, T. (1996). *Science* 274:982-985.
- Kamiya, N., Nagasaki, H., Morikami, A., Sato, Y., Matsuoka, M. (2003). *Plant J.* 35:429-41.
- Kojima, S., Banno, H., Yoshioka, Y., Oka, A., Machida, C. and Machida, Y. (1999). *DNA Res.* 6:407-410.
- Koncz, C., Martini, N., Mayerhofer, R., Koncz-Kalman, Z., Korber, H., Redei, G. P. and Schell, J. (1989). *Proc. Natl. Acad. Sci. USA* 86:8467-8471.
- Laux, et al. (1996). *Development* 122:87-96;
- Lincoln, C., Long, J., Yamaguchi, J., Serikawa, K. and Hake, S. (1994). *The Plant Cell* 6:1859-1876.
- Lloyd, A. M., Schena, M., Walbot, V. and Davis, R. W. (1994). *Science* 266:436-439.

- Matsumoto, N., Okada, K. (2001). *Genes Dev.* 15:3355-64.
- Mayer, et al. (1998). *Cell* 95: 805-815
- Mett, V. L., Lockhead, L. P. and Reynolds, P. H. S. (1993). *Proc. Natl. Acad. Sci. USA* 90:4567-4571.
- Okamuro, J. K., Caster, B., Villarroel, R., Van Montagu, M. and Jofuku, K. D. (1997). *Proc. Natl. Acad. Sci. USA* 94:7076-7081.
- Ooms, G., Kaup, A. and Roberts, J. (1983). *Theor. Appl. Genet.* 66:169-172.
- Picard, D. (1993). *Trends Cell Biol.* 3:278-280.
- Riechmann, J. L. and Meyerowitz, E. M. (1998). *Biol. Chem.* 379:633-646.
- Riou-Khamlichi, C., Huntley, R., Jacquemard, A. and Murray, J. A. (1999). *Science* 283:1541-1544.
- Schena, M., Lloyd, A. M. and Davis, R. W. (1991). *Proc. Natl. Acad. Sci. USA* 88:10421-10425.
- Sato, S., Kato, T., Tabata, S., Oka, A. (2001). *Science* 294:16.
- Smigocki, A. C. and Owens, L. D. (1988). *Proc. Natl. Acad. Sci. USA* 85:5131-5135.
- Smigocki, A. C. and Owens, L. D. (1989). *Plant Physiol.* 91:808-811.
- Tinland, B., Koukolíková-Nicola, Z., Hall, M. N., and Hohn, B. (1992). *Proc. Natl. Acad. Sci. USA* 89: 7442-7446.
- Valvekens, D., Montague, M.V. and Lijsbettens, M.V. (1988). *Proc. Natl., Acad. Sci, USA.* 85:5536-5540.
- Weinmann, P., Gossen, M., Hillen, W., Bujard, H. and Gatz, C. (1994). *Plant J.* 5:559-569.
- Yasutani I, Ozawa S, Niishida T, Sugiyama M and Komamine A (1994). *Plant Physiol.* 105:815-822.

Yoder, J. I. and Goldsbrough, A. P. (1994). *Bio/Technology* 12:263-267.

Yu, L.P., Miller, A.K. and Clark S.E. (2003). *Current Biol.* 3:179-188.

Zuo, J. and Chua, N-H. (2000). *Current Opinion in Biotechnology* 11:146-151.

Zuo J, Niu, Q. W. and Chua, N. H. (2000). *Plant J.* 24:265-273.

Zuo J, Niu, Q.W., Frugis, G., Chua, N.H. (2002). *Plant J.* 30:349-59.

Zuo, J., Niu, Q.W., Moller, S.G. and Chua, N.H. (2001). *Nat. Biotechnol.* 19:157-161.

Claims

We claim:

1. An isolated polynucleotide comprising
 - e. SEQ ID NOs: 57 or 15; or a complement thereof;
 - f. a polynucleotide of at least 65% identity to SEQ ID NOs: 57 or 15; or a complement thereof, wherein the polypeptide encoded by the polynucleotide of at least 65% identity to SEQ ID NOs: 57 or 15; or a complement thereof promotes plant regeneration, induces plant regeneration in the absence of one or more exogenous hormones, or increases plant transformation efficiency;
 - g. a polynucleotide encoding SEQ ID NOs: 2 or 16; or a complement thereof;
 - h. a polynucleotide encoding a polypeptide comprising a signature sequence:

(X¹³⁻²³)RWNPT(KEQRIMLV)(EDNKQ)Q(ILMV)(EKNSTQR)(MILV)L
 (DEKQST)(ADEHNST)L(FYW)(EQKR)(ADEKRQS)G(ILMV)
 RTP(AST)(AST)DQIQ(EKQR)I(AST)(AGST)
 (DEKQR)L(AEQKRST)(ASFWY)YG
 (EQHKTY)IE(AGST)KNVIFYWFQNHKAR(DEQKR)RQK(EKQR)(EQKR) (X²⁹⁻¹⁰¹)(DEKQR)(STACIGVLMNQP)L(EPQKR)LFP(MILV)(X¹³⁻⁵⁷),

wherein

X represents any amino acids,

the numbers following the X represent the range of the size of the domain represented by X,

amino acids in each bracket represent alternatives for the respective single amino acid position represented by each bracket,

the polypeptide comprises an NLS with the proviso that the NLS does not change the specified amino acids, and

AtCIR1 and AtCIR2 is excluded from the polypeptide comprising the signature sequence.

2. The isolated polynucleotide of item 1, further comprising an operably linked promoter, wherein said promoter is active in a plant.
3. The isolated polynucleotide of item 1, wherein said isolated polynucleotide is a gene.
4. The isolated polynucleotide of item 2, wherein said promoter is a constitutive promoter.
5. The isolated polynucleotide of item 2, wherein said promoter is an inducible promoter.
6. A vector which comprises the polynucleotide of item 1.
7. The vector of item 6, wherein said vector is a plasmid.
8. The vector of item 6, further comprising a promoter, wherein said promoter is operably linked to said polynucleotide.
9. The vector of item 8, wherein said promoter is active in a plant cell.
10. A binary vector system comprising the vector of item 9.
11. A bacterial cell comprising the vector of 7.

12. A plant cell comprising the vector of 9.

13. The isolated polynucleotide of item 1, wherein the signature sequence is
 (X^{13-23}) RWNPT(KV)(DE)Q(IL)(STK)(MIL)L(ET)(SND)L(FY)(KR)(QEA)G(IL)
 RTP(ST)(AT)DQIQ(QK)I(ST)(GST)(RE)L(KRS)(AF)YG(KHT)IE(GS)KNVIFYWFQN
 HKAR(QE)RQK(QR)(KR)(X^{29-101})(EK)(STACIGVLMNQP)L(PQ)LFP(LV)
 (X^{13-57}) .

14. The nucleotide of item 1, wherein the polypeptide comprises an additional sequence in the signature sequence at the X domains.

15. An isolated polypeptide comprising

- d. SEQ ID NOs: 2 or 16;
- e. a polypeptide of at least 65% identity to SEQ ID NOs: 2 or 16,
 wherein the polypeptide of at least 65% identity to SEQ ID NOs: 2 or 16 promotes plant regeneration, induces plant regeneration in the absence of one or more exogenous hormones or in a trace amount of them, or increases plant transformation efficiency;
- f. a signature polypeptide:

(X^{13-23}) RWNPT(KEQRIMLV)(EDNKQ)Q(ILMV)(EKNSTQR)(MILV)L
 (DEKQST)(ADEHNST)L(FYW)(EQKR)(ADEKRQS)G(ILMV)
 RTP(AST)(AST)DQIQ(EKQR)I(AST)(AGST)
 (DEKQR)L(AEQKRST)(ASFWY)YG
 (EQHKTY)IE(AGST)KNVIFYWFQNHKAR(DEQKR)RQK(EKQR)(EQKR) (X^{29-101})(DEKQR)(STACIGVLMNQP)L(EPQKR)LFP(MILV)(X^{13-57}),

X represents any amino acids,

the numbers following the X represent the range of the size of the domain represented by X,

amino acids in each bracket represent alternatives for the respective single amino acid position represented by each bracket,

the polypeptide comprises an NLS with the proviso that the NLS does not change the specified amino acids, and

AtCIR1 and AtCIR2 is excluded from the polypeptide comprising the signature sequence.

16. The isolated polypeptide of item 15, wherein the signature sequence is preferably

(X¹³⁻²³)RWNPT(KV)(DE)Q(IL)(STK)(MIL)L(ET)(SND)L(FY)(KR)(QEA)G(IL) RTP(ST)(AT)DQIQ(QK)I(ST)(GST)(RE)L(KRS)(AF)YG(KHT)IE(GS)KNVFWFQN HKAR(QE)RQK(QR)(KR)(X²⁹⁻¹⁰¹)(EK)(STACIGVLMNQP)L(PQ)LFP(LV) (X¹³⁻⁵⁷).

17. The isolated polypeptide of item 16, wherein the isolated polypeptide comprises an additional sequence in the signature sequence at the X domains.

18. A method of selecting a polynucleotide from species other than *Arabidopsis*, comprising the steps of

- a. constructing a cDNA library from cells, wherein the cDNAs in said cDNA library can be expressed in a plant cell,
- b. transforming *Arabidopsis* explants with vectors in said cDNA library,
- c. culturing said *Arabidopsis* explants in a plant-cell-culture medium that is non-supportive of regeneration with conventional vectors,

- d. selecting Arabidopsis explants regenerating in said plant-cell-culture medium,
 - f. identifying the nucleotide sequence of the polynucleotide that was transferred from said cDNA library to said Arabidopsis explants regenerating in said plant-cell-culture medium and enables said Arabidopsis explants to regenerate in said plant-cell-culture medium, and
 - g. isolating a polynucleotide of said nucleotide sequence from a cell, or synthesizing a polynucleotide of said nucleotide sequence.
19. The method of item 18, wherein Arabidopsis explants are transformed with said vectors in said cDNA library by *Agrobacterium*-mediated T-DNA conjugation.
20. The method of item 18, wherein Arabidopsis explants are transformed with said vectors in said cDNA library by particle gun bombardment
21. The method of item 18, wherein said plant-cell-culture medium does not contain no cytokinins or trace amount of cytokinins.
22. The method of item 18, wherein said plant-cell-culture medium does not contain 2ip.
23. The method of item 18, wherein said Arabidopsis explants are Arabidopsis root explants.
24. A method of selecting a plant cell transformant comprising the steps of
- d. transforming plant cells with vectors, wherein said vector comprises

- iii. a polynucleotide encoding a polypeptide that enables a transformant to regenerate in a medium deficient of one or more hormones, and
- iv. a promoter that is active in a plant and operably linked to said isolated polynucleotide,
- e. culturing said plant cells in a plant-cell-culture medium that is deficient of one or more hormones, and
- f. selecting regenerating plant cells.

25. The method of item 24, wherein the polypeptide that enables a transformant to regenerate in said medium is

- a. a polypeptide of at least 65% identity to SEQ ID NOs: 2 or 16; or
- b. a signature polypeptide:

(X¹³⁻²³)RWNPT(KEQRIMLV)(EDNKQ)Q(ILMV)(EKNSTQR)(MILV)L
 (DEKQST)(ADEHNST)L(FYW)(EQKR)(ADEKRQS)G(ILMV)
 RTP(AST)(AST)DQIQ(EKQR)I(AST)(AGST)
 (DEKQR)L(AEQKRST)(ASFWY)YG
 (EQHKTY)IE(AGST)KNVFYWFQNHKAR(DEQKR)RQK(EKQR)(EQKR) (X²⁹⁻¹⁰¹)(DEKQR)(STACIGVLMNQP)L(EPQKR)LFP(MILV)(X¹³⁻⁵⁷),

X represents any amino acids,
 the numbers following the X represent the range of the size of the domain represented by X,
 amino acids in each bracket represent alternatives for the respective single amino acid position represented by each bracket, and

the polypeptide comprises an NLS with the proviso that the NLS does not change the specified amino acids.

26. The method of item 24, wherein said vector further comprises a polynucleotide of interest.

27. The method of item 24, wherein said promoter is a chemically inducible promoter.

28. The method of item 24, wherein said promoter is a constitutive promoter.

29. The method of item 24, wherein said isolated polynucleotide and promoter are chemically excisable.

30. The method of item 24, wherein said plant-cell-culture medium does not contain any antibiotics or herbicides.

31. The method of item 24, wherein said plant-cell-culture medium does not contain one or more hormones.

32. A method of improving plant transformation efficiency comprising the steps of

a. transforming plant cells with vectors, wherein said vector comprises:

- i. a polynucleotide encoding a polypeptide that enables a transformant to regenerate in a medium that is deficient of one or more hormones or increases regeneration rate in a regular plant-cell-culture medium, and
- ii. a promoter that is active in a plant and operably linked to said isolated polynucleotide,

- b. culturing said plant cells in a plant-cell-culture medium, which is optionally supplemented with cytokinins, and optionally supplemented with herbicides, antibiotics or both, and,
- c. selecting regenerating plant cells.

33. The method of item 32, wherein said polypeptide is

- a. a polypeptide of at least 65% identity to SEQ ID NOs: 2 or 16; or
- b. a signature polypeptide:

(X¹³⁻²³)RWNPT(KEQRIMLV)(EDNKQ)Q(ILMV)(EKNSTQR)(MILV)L
 (DEKQST)(ADEHNST)L(FYW)(EQKR)(ADEKRQS)G(ILMV)
 RTP(AST)(AST)DQIQ(EKQR)I(AST)(AGST)
 (DEKQR)L(AEQKRST)(ASFWY)YG
 (EQHKTY)IE(AGST)KNVIFYWFQNHKAR(DEQKR)RQK(EKQR)(EQKR) (X²⁹⁻¹⁰¹)(DEKQR)(STACIGVLMNQP)L(EPQKR)LFP(MILV)(X¹³⁻⁵⁷),

X represents any amino acids,

the numbers following the X represent the range of the size of the domain represented by X,

amino acids in each bracket represent alternatives for the respective single amino acid position represented by each bracket, and

the polypeptide comprises an NLS with the proviso that the NLS does not change the specified amino acids.

34. The method of item 32, wherein the polypeptide that enables a transformant to regenerate in a medium that is deficient of one or more hormones or increases

regeneration rate in a regular plant-cell-culture medium is a serine or threonine phosphatase.

35. The method of item 32, wherein said plant is cotton, maize, soybeans, rice, wheat, barley, sugarbeet, and oil palm.
36. The method of item 32, wherein said promoter is a chemically inducible promoter.
37. The method of item 32, wherein said promoter is a constitutive promoter.
38. The method of item 32, wherein said isolated polynucleotide and promoter are chemically excisable.
39. The method of item 32, wherein said plant-cell-culture medium contains antibiotics or herbicides.
40. The method of item 32, wherein said regular plant-cell-culture medium contains cytokinin.
41. The method of item 32, wherein said transformed cells are regenerated via somatic embryogenesis.
42. The method of item 32, wherein said transformed cells are regenerated via organogenesis.
43. The isolated polynucleotide of items 1, wherein the polypeptide encoded by said polynucleotides recited by d promotes plant regeneration, induces plant regeneration in the absence of one or more exogenous hormones or in the trace amount of them, or increases plant transformation efficiency.
44. The isolated polypeptide of item 15, wherein the polypeptide recited by c above promotes plant regeneration, induces plant regeneration in the absence

of one or more exogenous hormones or in the trace amount of them, or increases plant transformation efficiency.

45. A method of improving plant transformation efficiency comprising the steps of
- a. transforming plant cells with vectors, wherein said vector comprises:
 - i. the isolated polynucleotide encoding a serine/threonine phosphatase, and
 - ii. a promoter that is active in a plant and operably linked to said isolated polynucleotide,
 - b. culturing said plant cells in a plant-cell-culture medium, which is optionally supplemented with cytokinins, and optionally supplemented with herbicides, antibiotics or both, and,
 - c. selecting regenerating plant cells.
46. The method of item 45, wherein said theonine/serine phosphatase is the arabidopsis Poltergeist.
47. The method of item 45, wherein said theonine/serine phosphatase shares at least 65% identity to the arabidopsis Poltergeist.
48. The methods of item 46 or 47, wherein said phosphatase is preferably a mutant deleted in negative regulation domain.