Title: GENE REGULATORY REGION THAT PROMOTES ROOT-SPECIFIC TRANSCRIPTION AND ITS USES

Abstract: Nucleic acid sequence which directs root-specific transcription of contiguous genes in plants is provided. This sequence may be used in transgenic plants to promote expression of endogenous and foreign genes in roots to enhance plant resistance to diseases and pests, stress (salt) tolerance, and improve nutritive value of edible root plants. In addition, this sequence may be useful for the production of recombinant proteins in roots aimed at molecular farming and phytoremediation.
GENE REGULATORY REGION THAT PROMOTES ROOT-SPECIFIC
TRANSCRIPTION AND ITS USES

This application derives priority from U.S. Provisional Application No. 60/206,788,
which was filed May 24, 2000.

Field of the Invention

This invention relates to a nucleic acid sequence which directs root-specific
transcription of contiguous genes in plants. The nucleic acid sequence of the present
invention can be used in transgenic plants to promote expression of endogenous and foreign
genes in roots to enhance plant resistance to diseases and pests, stress (salt) tolerance, and
improve nutritive value of edible root plants. In addition, the nucleic acid sequence of the
present invention can be useful for the production of recombinant proteins in roots, with
applications in molecular farming and phytoremediation.

BACKGROUND

Plant improvement using genetic engineering techniques requires the availability of
highly tissue-specific promoters, so that the expression of genes of interest may be targeted to
appropriate tissues. The two principal reasons for that are as follows. First, high levels of
expression in the desired tissue, rather than constitutive expression, is likely to be less
metabolically demanding. Second, it is desirable to direct expression of foreign genes to
tissues which are not used for human or animal consumption, whenever the introduced gene
does not encode a nutrient protein.
Despite their important role in plant development, relatively little work has been done on the regulation of gene expression in roots in comparison to shoots, and as a result regulatory sequences that promote root-specific expression are not abundant. However, the emerging biotechnologies including phytoremediation, the use of plants to remove pollutants from the environment, and molecular farming, the use of plants for the production of valuable natural products and recombinant proteins are rapidly changing this situation (Gleba et al., 1999). These technologies rely on the natural ability of plant roots to absorb, bind and concentrate toxic metals from soil (phytoextraction), and continuously exude a diverse array of plant products, including proteins, into their immediate environment (rhizosecretion).

Strategies for increasing the efficiency of phytoextraction involve generating superior plant varieties by genetic engineering with increased capacity for metal uptake and binding, as well as enhanced root resistance to high levels of heavy metals. Similarly plant varieties optimized for the rhizosecretion of recombinant proteins and valuable natural products into simple hydroponic media are being developed. For the development of transgenic plants with improved phytoextraction and rhizosecretion properties the availability of strong promoters which control expression of genes in roots over the lifetime of a plant is essential.

**SUMMARY OF THE INVENTION**

The present invention involves the isolation and use of a nucleic acid sequence which directs root-specific transcription of contiguous genes in plants. The inventors have isolated a nucleic acid sequence which acts as a promoter from gene encoding a root-specific condensing enzyme of very long chain fatty acid biosynthesis in *Lesquerella fendleri*, *LfKCS45*. As a non-limiting example, the present invention demonstrates that the *LfKCS45* 5' regulatory sequence is useful in promoting transcription of heterologous genes in *Arabidopsis* roots.
The nucleic acid sequence of the present invention can be used in transgenic plants to promote expression of endogenous and foreign genes in roots to enhance plant resistance to diseases and pests, stress (salt) tolerance, and improve nutritive value of edible root plants. In addition, the nucleic acid sequence of the present invention can be useful for the production of recombinant proteins in roots, with applications in molecular farming and phytoremediation.

**BRIEF DESCRIPTION OF THE DRAWING**

Figure 1 shows the sequence of the insert in the plasmid pLfKCS3-GUS.

**DETAILED DESCRIPTION**

The present invention provides a nucleic acid sequence which directs root-specific transcription of contiguous genes in plants. By way of non-limiting example, the present invention shows that an isolated transcription regulatory region from the LfKCS45 gene is capable of directing expression of desired genes in a root-specific manner.

Because this regulatory sequence can also promote transcription in roots of a different plant species, it is useful for a number of different applications in a variety of dicotyledonous plants.

Examples of applications wherein the nucleic acid sequence of the present invention can be useful includes for example:

1. the improvement of nutritive value of edible roots, by affecting content and composition of nutrients such as various proteins,

2. the protection of roots against root-feeding insects (expression of genes encoding insecticidal toxin of Bacillus *thuringiensis*, chitinase, proteinase inhibitor, an antibody, etc.),
nematodes (expression of genes encoding a collagenase, a lectin, a proteinase inhibitor, an antibody, etc.) and disease (expression of Pto gene, viral coat protein, an antibody, etc.),

(3) production of novel recombinant proteins in the root for pharmaceutical, industrial or nutritional purposes,

(4) production of novel compounds/products in the root, ie. secondary metabolites, biodegradable plastic polymers, etc., and

(5) increasing the efficiency of phytoextraction (phytoremediation) through expression of high affinity, high selectivity metal-binding peptides and transporters in roots and by enhancing metal tolerance of roots to prolong plant survival in metal-contaminated sites (6) enhancing stress (salt) tolerance of plants.

The methods employed in the isolation of the nucleic acid of the present invention and the uses thereof are discussed in the following non-limiting examples:

Examples:

**Isolation of a root-specific promoter region form *Lesquerella fendleri***

A *Lesquerella fendleri* genomic DNA library was obtained from Dr. Chris Somerville of the Carnegie Institution of Washington, Stanford, CA. The genomic library was plated on *E. coli* LE392 (Promega) and about 150,000 clones were screened using *Arabidopsis FAE1* gene (James et al., 1995) as a probe. The probe was prepared by PCR using pGEM-7Zf(+)–FAE1 (Millar and Kunst, 1997) as a template with *FAE1* upstream primer, 5’CCGAGCTCAAAGAGGATACATAC-3’ and *FAE1* downstream primer, 5’GATACTCGAGAACGTTGCACTCAGATAC-3’. PCR was performed in a 10μl reaction containing 10 ng of the template, 2mM MgCl2, 1.1 [μM of each primer, 100 [μM of (dCTP + dGTP + dTTP) mix, 50 μCi of [α-32P]dATP, 1X PCR buffer and 2.5 units of *Taq* DNA polymerase (Life Technologies). Amplification conditions were: 2 min of initial
denaturation at 94°C, 30 cycles of 94°C for 15 sec, 55°C for 30 sec, 72°C for 1 min and 40 sec, followed by a final extension at 72°C for 7 min. The amplified radiolabeled probe was purified by QIAquick PCR Purification Kit (Qiagen) and denatured by boiling before adding to the hybridization solution. Hybridization took place overnight at 65°C in a solution containing 6X SSC, 20 mM NaH2PO4, 0.4% SDS, SX Denhardt's solution, and 50 μg/ml sonicated, denatured salmon sperm DNA (Sigma) and washing was performed three times for 20 min each in 2X SSC, 0.5% (w/v) SDS at 65°C.

Nine clones with sequences corresponding to the *Arabidopsis FAE1* gene were isolated from the *Lesquerella fendleri* genomic library. The phage DNA from those nine clones was extracted and purified using QIAGEN Lambda Mini Kit (Qiagen) according to the manufacturer's protocol. One of the clones was digested with *KpnI* and a 6.1 kb fragment was subcloned into the pGEM-7Zf(+) vector (Promega) also cut with *KpnI*. The whole 6.1 kb insert was then sequenced with ABI automatic 373 DNA sequencer using fluorescent dye terminators. It consists of a 1025 by 5' flanking region, a 1464 bp open reading frame coding for a condensing enzyme involved in very long chain fatty acid biosynthesis in roots, and an 3609 by 3' flanking sequence. 1025 by region directly 5' of the ORF was amplified using the high fidelity *Pfu* polymerise (Life Technologies) with a forward primer 5'-GGCAAGCTTCGGTACCAGCGTAAAATTTCTTTCC-3' and a reverse primer 5'-GCCGTCGACTGACGGGTAAATGAGAAAAAG-3' and inserted upstream of the (β-glucuronidase (GUS) gene in pBI101 (Clontech) cut with *HindIII* and *SalI*, resulting in the vector pLfKCS45-GUS. The sequence of the insert in the plasmid pLfKCS3-GUS is shown in Figure 1.

Functional analysis of the *LfKCS45* promoter region

To evaluate the ability of the promoter fragment of the *LfKCS45* gene to confer root
specific regulation of gene expression in plants, the pLfKCS45-GUS construct was introduced into \textit{Agrobacterium tumefaciens} strain GV3101 (pMP90) (Koncz and Schell, 1986) by heat shock and selected for resistance to kanamycin (50 \([\mu g/mL])\. \textit{A. thaliana} ecotype Columbia was transformed with \textit{A. tumefaciens} harbouring the pLfKCS45-GUS construct using floral dip method (Clough and Bent, 1998). Screening for transformed seed was done on 50 [\mu g/mL] kanamycin as described previously (Katavic et al., 1994). 32 transgenic lines were generated and examined for GUS activity.

Histochemical localization of GUS activity in transgenic plants was done in different tissues as follows. Tissue sections were incubated in 50 mM sodium phosphate, pH 7.0, 0.5 mM potassium ferricyanide, 0.5 mM potassium ferrocyanide, 10 mM EDTA, 0.05% (w/v) triton X-100, and 0.35 mg/ml 5-bromo-4-chloro-3-indolyl-\(\beta\)-D-glucuronide (X-Gluc) for 4 to 7 hours at 37°C (Jefferson, 1987). Following staining the blue-stained samples were fixed in 70% ethanol.

Leave stems, inflorescences, roots, and siliques at different stages of development of more than 30 independent transgenic \textit{Arabidopsis} lines were examined for \(\beta\)-glucuronidase activity. Strong GUS activity was detected exclusively in root tips starting soon after germination, whereas there was no GUS staining in leaves, stems, flowers, and embryos. GUS activity in the root in all the examined transgenic lines persisted throughout subsequent plant development into maturity. Thus the \textit{LjKCS45} promoter is useful for root-specific expression of foreign genes in transgenic plants and is capable of promoting transcription in plants other than \textit{Lesquerella fendleri}.

In addition, root-specific expression conferred by the \textit{LJKCS45} promoter is independent of the native terminator at the \textit{LJKCS45} gene 3' end. In all our constructs, a terminator derived from the \textit{Agrobacterium} nopaline synthase gene was used. Thus, the
sequence in the 1025 bp promoter construct is sufficient for the desired expression profile independent of ancillary sequences.
References:


What we claim is:

1. An isolated nucleic acid fragment comprising a nucleic acid sequence encoding a promoter for directing root-specific transcription of contiguous genes in plants.

2. An isolated nucleic acid fragment according to claim 1, wherein said sequence is defined by the nucleic acid sequence of SEQ. ID. NO. 1.

3. An isolated nucleic acid fragment according to claim 1, wherein said sequence has a sequence identity of 95% or greater to the nucleic acid sequence of SEQ. ID. NO. 1.

4. An isolated nucleic acid fragment according to claim 1, wherein said sequence has a sequence identity of 85% or greater to the nucleic acid sequence of SEQ. ID. NO. 1.

5. An isolated nucleic acid fragment according to claim 1, wherein said sequence has a sequence identity of 65% or greater to the nucleic acid sequence of SEQ. ID. NO. 1.

6. An isolated nucleic acid fragment comprising a nucleic acid sequence encoding a promoter for enhancing expression of endogenous and foreign genes in roots of plants.

7. An isolated nucleic acid fragment according to claim 3, wherein said wherein said sequence is defined by the nucleic acid sequence of SEQ. ID. NO. 1.

8. An isolated nucleic acid fragment according to claim 1, wherein said sequence promotes expression of genes, which enhance plant resistance to a condition selected from the group consisting of infection of disease, attack of pests, and stress-salt tolerance.

9. An isolated nucleic acid fragment according to claim 3, wherein said genes enhance the nutritive value of edible roots of plants.

10. An isolated nucleic acid fragment comprising a nucleic acid sequence encoding a promoter for enhancing production of recombinant proteins in roots of plants.
Figure 1. *Lesquerella fendleri* LJKCS45 promoter:

(Length: 1025 bp)

```
GGTAC CAGCGTAAAA TTCTTTTCTCT -1025
CTAAGCTTTGG AGGCATTTGG AGGCCAGGGT TCTACTCAAT ATATAATCAA -1000
TAGGGAATAT TGCTTTTCTCA GTCAATTGTTAT CACCGGTTTCC TTATGATACT -950
GATATTTTTCTT GAAAGATTTC CGATAGACGGA TTGGATCCAG -900
CCTGCCAGCTG TCTGCGAAAA AACGCAAAGTG TGGCGCTGAA ATTATAGATA -850
TTTACGTATT AACCTGAGTA AAACGCGTCCG TTTTGATCTG TATTTGAAT -800
TATATCGTAAT ATATCTGTAAT ACGGCTGGAAC ACTAGCAGTCT TTTTTCAAGG -750
ACAGTGCGAC GCTGTGTATAT AACCGCGCTGT GGCATTACCT TTCAGGTGCC -700
AAAGAAATGT CAGCATGGAG AAGAATGATG AGCGCTCTGC ATAAAAATCT -650
CTATAAAATA ACGCTAATAAT ATAGTGGTAA AAAATAGTGA CAGACACAAAC -600
TTGGCCGGAAG TAGTTACGCA AATGTTGGGA GTCAAAACC AAATTCCTCA -550
GAAATAAAGC AAATAGGGTT ATGACGCTGAA AGTAGATTAG GAGGAACCTTC -500
CATGCAAGAA AAAAAAGAGT AGGAAAAACAG TGCATGAGGG TGTAAAATCTC -450
TCAACATAGT TTTCTTGTCA TTAATTTTGTT TTCTTTCTCT TTTTTGTTT -400
TCTTTTTTT TATTTCTTAT CGTTGCGTTG TATAAGAACCA AAACACGCGT -350
ATAATTATTT TTTTCATGAA TATGTGAAAG TATGTAATGA AAAGAAAAAT -300
AAAAATTTAT GCGTCGTCATG TCTATACGTT TCAGAAAGAG AAATTGGGAT -250
TACAAGATGG CAATATTTCC TAACACTGCA TACAAACAGC GATTATATAT -200
AAATATATTT TTTTATTTTAT TTTACGTTAA CAGGTGTGTTG ATGAAGAAAGG -150
TATCATCTC TATAAAATGT TGCTTGCTGT ATAGACGCAA GACAGCCAGC -100
CAGAATAAGA GAGTATCTAT ATATCTAATCT TTTTTCTCAT TACCCGTCAA -50
GCTAACG
```