US 20100287660A1

(19) United States

(12) Patent Application Publication

Spangenberg et al.

(10) Pub. No.: US 2010/0287660 A1

(43) Pub. Date: Nov. 11, 2010

(54) **MODIFICATION OF LIGNIN BIOSYNTHESIS VIA SENSE SUPPRESSION**

(75) Inventors: **German Spangenberg**, Bundoora (AU); **Angela Jane Lidgett**, Kew (AU); **Robyn Louise Heath**, Clifton Hill (AU); **Russell Leigh McInnes**, Bundoora (AU); **Damian Paul Lynch**, Northcote (AU); **Ulrik Peter John**, Westgarth (AU); **Aidyn Mouradov**, Mill Park (AU); **Megan Elizabeth Griffith**, Templestowe (AU)

Correspondence Address:
Larson & Anderson, LLC
P.O. BOX 4928
DILLON, CO 80435 (US)

(73) Assignees: **DAIRY AUSTRALIA LIMITED**, Southbank, Victoria (AU); **MOLECULAR PLANT BREEDING NOMINEES LTD.**, Bundoora, Victoria (AU)

(21) Appl. No.: **12/669,659**

(22) PCT Filed: **Jul. 17, 2008**

(86) PCT No.: **PCT/AU08/01034**

§ 371 (c)(1),
(2), (4) Date: **Jul. 23, 2010**

(30) **Foreign Application Priority Data**

Jul. 19, 2007 (AU) 2007203378

Publication Classification

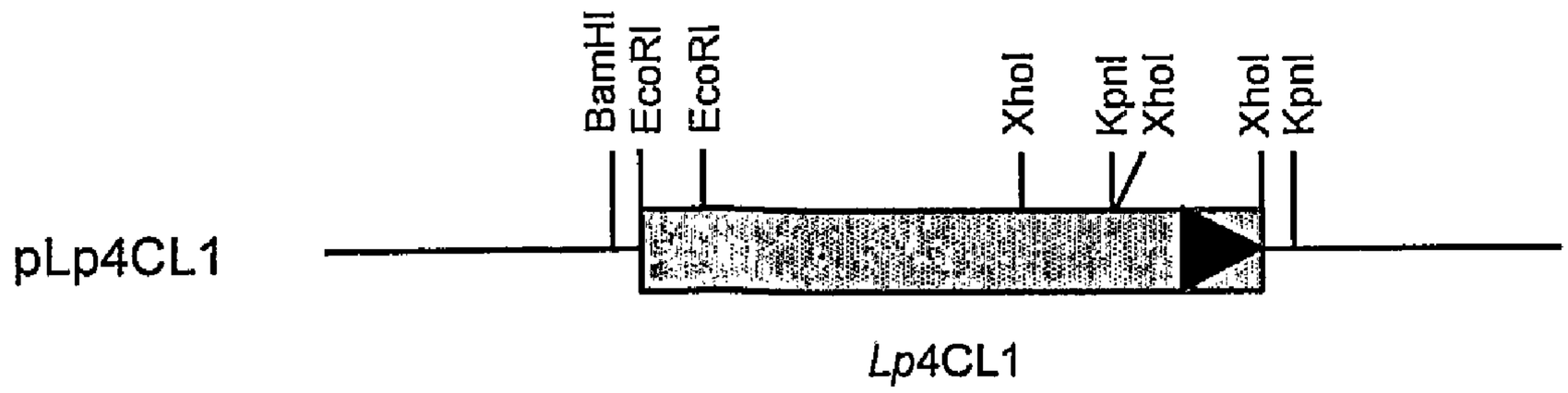
(51) **Int. Cl.**

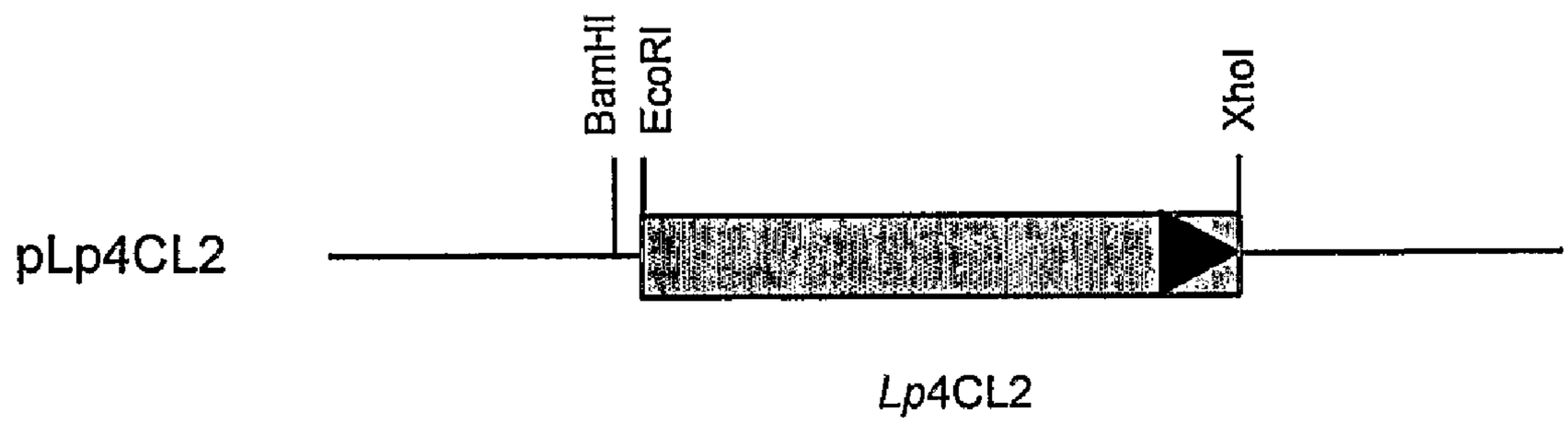
A01H 1/00	(2006.01)
C07H 21/04	(2006.01)
C12N 15/63	(2006.01)
C12N 5/10	(2006.01)
A01H 5/00	(2006.01)

(52) **U.S. Cl.** **800/278**; 536/24.5; 435/320.1; 435/419; 800/298

(57) **ABSTRACT**

The present invention relates to the modification of lignin biosynthesis in plants, to enzymes involved in the lignin biosynthetic pathway and nucleic acids encoding such enzymes and, more particularly, to methods of modifying lignin biosynthesis via sense suppression and to related nucleic acids and constructs.


BamHI
EcoRI
EcoRI
XhoI
KpnI
XhoI
XhoI
KpnI
pLp4CL1
Lp4CL1


BamHI
EcoRI
XhoI
pLp4CL2
Lp4CL2

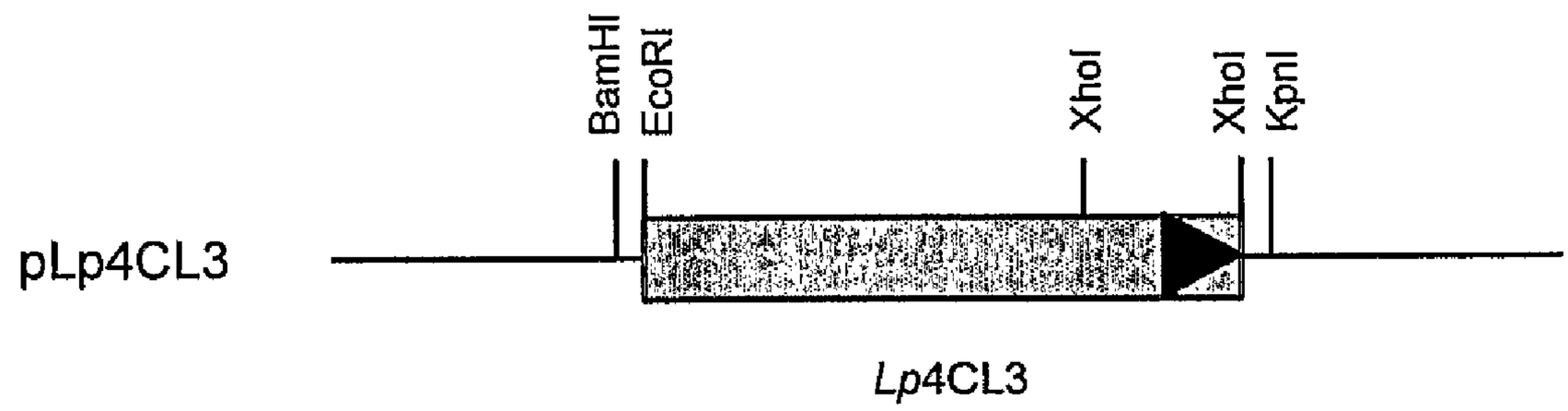

BamHI
EcoRI
XhoI
XhoI
KpnI
pLp4CL3
Lp4CL3

FIGURE 1

```
     CGGCACGAGTGGACTTTCCGACGCCGGAGTCGCCGATGATGACCGCCTTGAGGAGGTAGT
  1  ---------+---------+---------+---------+---------+---------+   60

     CGTAGTCGTCCTCCGCCCTGTACGCGCCGCTGCCCGCCATTTCCTTCCTCGCCTCGCGGG
 61  ---------+---------+---------+---------+---------+---------+  120

     TCCTCCTCCCCGACCTGCGCTAGGCTCTGGATCTCGCGGGGTTTGGGCGCGGCGTCCTCG
121  ---------+---------+---------+---------+---------+---------+  180

     CTGTGAGCTCGTGCCGAATTCGGCACGAGCCACCTTCGAGGCGTGCACTGGTACGAGCTC
181  ---------+---------+---------+---------+---------+---------+  240

     GCGAGCCATTGTCAGTGCAGTGTAGGCTCTGCTACTCGTTGGCCATTCCAAGAAGCTCTC
241  ---------+---------+---------+---------+---------+---------+  300

     TGCTCCCTGAAACCAGAGGATCATGATCACGGTGGCGGCGCCCGAGGTGCAGCAGCCGCA
301  ---------+---------+---------+---------+---------+---------+  360
                            M  I  T  V  A  A  P  E  V  Q  Q  P  Q

     GATCGCGGCGGCTGCTGCGGCCGTGGAGGCGGCGGCACCGGAGGCGACGACGATCTTCCG
361  ---------+---------+---------+---------+---------+---------+  420
      I  A  A  A  A  A  A  V  E  A  A  A  P  E  A  T  T  I  F  R

     GTCCAGGCTCCCGGACATCGACATCCCGACCCACATGCCCCTGCACGACTATTGCTTCGC
421  ---------+---------+---------+---------+---------+---------+  480
      S  R  L  P  D  I  D  I  P  T  H  M  P  L  H  D  Y  C  F  A

     GACGGCAGCCTCGGCCCCGGACGCGCCGTGCCTCATCACCGCGGCCACGGGGAAGACCTA
481  ---------+---------+---------+---------+---------+---------+  540
      T  A  A  S  A  P  D  A  P  C  L  I  T  A  A  T  G  K  T  Y

     CACGTTCGCCGAGACGCACCTGCTGTGCCGCAAGGCCGCGGCGGCGCTGCACGGGCTCGG
541  ---------+---------+---------+---------+---------+---------+  600
      T  F  A  E  T  H  L  L  C  R  K  A  A  A  A  L  H  G  L  G

     CGTGCGCCACGGGGACCGGATCATGCTGCTGCTCCAGAACTCCGTGGAGTTCGCGCTCGC
601  ---------+---------+---------+---------+---------+---------+  660
      V  R  H  G  D  R  I  M  L  L  L  Q  N  S  V  E  F  A  L  A

     CTTCTTCGGCGCGTCCATGCTCGGCGCCGTCAGCACGGCGGCGAACCCGTTCTGCACGCC
661  ---------+---------+---------+---------+---------+---------+  720
      F  F  G  A  S  M  L  G  A  V  S  T  A  A  N  P  F  C  T  P

     GCAGGAGATCCACAAGCAGCTCGTGGCCTCCGGCGCGAAGCTGGTCGTCACGCAGTCCGC
721  ---------+---------+---------+---------+---------+---------+  780
      Q  E  I  H  K  Q  L  V  A  S  G  A  K  L  V  V  T  Q  S  A
```

FIGURE 2

```
     CTACGTCGACAAGCTCCGGCACGAGGCCTTCCCCCGAATCGGCGAGGCCCTCACCGTGAT
781  ---------+---------+---------+---------+---------+---------+   840
      Y  V  D  K  L  R  H  E  A  F  P  R  I  G  E  A  L  T  V  I

     CACCATCGACGAGGACGACGGCACCCCGGACGGCTGCCAGCCGTTCTGGGCCCTCGTGTC
841  ---------+---------+---------+---------+---------+---------+   900
      T  I  D  E  D  D  G  T  P  D  G  C  Q  P  F  W  A  L  V  S

     AGCCGCCGACGAGAACAGCGTCCCGGAGTCTCCCATCTCGCCGGACGACGCGGTGGCGCT
901  ---------+---------+---------+---------+---------+---------+   960
      A  A  D  E  N  S  V  P  E  S  P  I  S  P  D  D  A  V  A  L

     GCCCTACTCGTCGGGCACGACGGGGCTGCCCAAGGGCGTGGTGCTGACGCACGGGGGCT
961  ---------+---------+---------+---------+---------+---------+   1020
      P  Y  S  S  G  T  T  G  L  P  K  G  V  V  L  T  H  G  G  L

     GGTGTCGAGCGTGGCGCAGCAGGTGGACGGCGAGAACCCGAACCTGCACATGCGGGCGGG
1021 ---------+---------+---------+---------+---------+---------+   1080
      V  S  S  V  A  Q  Q  V  D  G  E  N  P  N  L  H  M  R  A  G

     GGAGGACGTGGTGCTCTGCGTGCTGCCGCTCTTCCACATCTTCTCGCTCAACTCGGTGCT
1081 ---------+---------+---------+---------+---------+---------+   1140
      E  D  V  V  L  C  V  L  P  L  F  H  I  F  S  L  N  S  V  L

     GCTGTGCGCGCTGCGGGCGGGCGCCGCCGTGATGCTGATGCCTAGGTTCGAGATGGGGGC
1141 ---------+---------+---------+---------+---------+---------+   1200
      L  C  A  L  R  A  G  A  A  V  M  L  M  P  R  F  E  M  G  A

     CATGCTGGAGGGCATCGAGCGGTGGCGCGTCACGGTGGCGGCCGTGGTGCCGCCGCTGGT
1201 ---------+---------+---------+---------+---------+---------+   1260
      M  L  E  G  I  E  R  W  R  V  T  V  A  A  V  V  P  P  L  V

     GCTCGCGCTCGCCAAGAACCCCGGGGTGGAGAAGCACGACCTCAGCTCCATTCGGATCGT
1261 ---------+---------+---------+---------+---------+---------+   1320
      L  A  L  A  K  N  P  G  V  E  K  H  D  L  S  S  I  R  I  V

     GCTCTCCGGCGCCGCGCCGCTCGGCAAGGAGCTCGAGGACGCGCTACGTGGCCGCCTGCC
1321 ---------+---------+---------+---------+---------+---------+   1380
      L  S  G  A  A  P  L  G  K  E  L  E  D  A  L  R  G  R  L  P

     GCAGGCCATCTTCGGACAGGGCTACGGGATGACGGAGGCCGGGCCGGTGCTGTCCATGTG
1381 ---------+---------+---------+---------+---------+---------+   1440
      Q  A  I  F  G  Q  G  Y  G  M  T  E  A  G  P  V  L  S  M  C

     CCCGGCGTTCGCGCGGGAGCCGACGCCGGCCAAGTCCGGCTCGTGCGGCACCGTGGTGCG
1441 ---------+---------+---------+---------+---------+---------+   1500
      P  A  F  A  R  E  P  T  P  A  K  S  G  S  C  G  T  V  V  R
```

FIGURE 2 CONTINUED

```
     CAACGCCCAGCTCAAGGTGGTCGACCCCGACACCGGCGTCTCCCTCGGCCGCAACCTCCC
1501 ---------+---------+---------+---------+---------+---------+  1560
      N  A  Q  L  K  V  V  D  P  D  T  G  V  S  L  G  R  N  L  P

     CGGCGAGATCTGCATCCGCGGCCCGCAGATCATGAAAGGATACTTGAATGATCCCGTGGC
1561 ---------+---------+---------+---------+---------+---------+  1620
      G  E  I  C  I  R  G  P  Q  I  M  K  G  Y  L  N  D  P  V  A

     CACGGCCGCGACCATCGACGTCGAGGGGTGGCTCCACACCGGCGACATCGGCTACGTCGA
1621 ---------+---------+---------+---------+---------+---------+  1680
      T  A  A  T  I  D  V  E  G  W  L  H  T  G  D  I  G  Y  V  D

     CGACGACGACGAGGTCTTCATCGTCGACCGCGTCAAGGAGCTCATCAAGTTCAAGGGCTT
1681 ---------+---------+---------+---------+---------+---------+  1740
      D  D  D  E  V  F  I  V  D  R  V  K  E  L  I  K  F  K  G  F

     CCAGGTACCGCCGGCCGAGCTCGAGGCTCTGCTCATCGCGCATCCGTCCATCGCCGACGC
1741 ---------+---------+---------+---------+---------+---------+  1800
      Q  V  P  P  A  E  L  E  A  L  L  I  A  H  P  S  I  A  D  A

     GGCCGTCGTCCCGCAAAAGGATGATGCCGCCGGCGAGGTCCCGGTTGCCTTCGTGGTCCG
1801 ---------+---------+---------+---------+---------+---------+  1860
      A  V  V  P  Q  K  D  D  A  A  G  E  V  P  V  A  F  V  V  R

     CGCCGCCGACTCCGACATCGCCGAGGAGGCCATCAAGGAGTTCGTATCCAAGCAGGTGGT
1861 ---------+---------+---------+---------+---------+---------+  1920
      A  A  D  S  D  I  A  E  E  A  I  K  E  F  V  S  K  Q  V  V

     GTTCTACAAGAGGCTGCACAAGGTCTACTTCACCCACGCGATACCCAAGTCGGCGTCGGG
1921 ---------+---------+---------+---------+---------+---------+  1980
      F  Y  K  R  L  H  K  V  Y  F  T  H  A  I  P  K  S  A  S  G

     GAAGATACTCAGGAAAGAACTCAGAGCTAAACTCGCCGCCCCGGCCACTGCCTGAAGAGT
1981 ---------+---------+---------+---------+---------+---------+  2040
      K  I  L  R  K  E  L  R  A  K  L  A  A  P  A  T  A  *  R  V

     GGTTCATGGCTTCATGCTAATCATTTCGATCAGAAAGGCACTTCTAGCATATATGTTCCA
2041 ---------+---------+---------+---------+---------+---------+  2100
      V  H  G  F  M  L  I  I  S  I  R  K  A  L  L  A  Y  M  F  H

     CCTTTTGTTTCATTTGGAAGATTGTATTCCAGCTAGTGGCCAGTGACTGAGTAAGGGATG
2101 ---------+---------+---------+---------+---------+---------+  2160
      L  L  F  H  L  E  D  C  I  P  A  S  G  Q  *

     GGGATAAAAGTTTTGTCTACGTTTTCTTTTACGCTACTCTCTCCATTGGGGAGTACAATG
2161 ---------+---------+---------+---------+---------+---------+  2220

     TATCAGGGGATTCGTGATTGAAGTTAATCAAGATTGGTTCAATTATAAAAAAAAAAAAAA
2221 ---------+---------+---------+---------+---------+---------+  2280

     AAAA
2281 ----  2284
```

FIGURE 2 CONTINUED

```
     CGGCACGAGCGCCATTCCTCCACCTTCAGCTCCGGCCAAAGATTTCCATCCGGCGAGATC
  1  ---------+---------+---------+---------+---------+---------+   60

     CATGGGCTCCATCGCGGCGGACGCGCCTCCCGCGGAGCTGGTGTTCCGGTCCAAGCTCCC
 61  ---------+---------+---------+---------+---------+---------+  120
      M  G  S  I  A  A  D  A  P  P  A  E  L  V  F  R  S  K  L  P

     GGACATCGAGATCCCGACCCACCTGACGCTGCAGGACTACTGCTTCCAGCGCCTGCCGGA
121  ---------+---------+---------+---------+---------+---------+  180
      D  I  E  I  P  T  H  L  T  L  Q  D  Y  C  F  Q  R  L  P  E

     GCTCTCCGCGCGCGCCTGCCTCATCGACGGCGCCACGGGCGCCGCGCTCACCTACGGCGA
181  ---------+---------+---------+---------+---------+---------+  240
      L  S  A  R  A  C  L  I  D  G  A  T  G  A  A  L  T  Y  G  E

     GGTGGACGCCCTGTCCCGCCGCTGCGCCGCGGGGCTGCGCCGCCTCGGCGTCGGCAAGGG
241  ---------+---------+---------+---------+---------+---------+  300
      V  D  A  L  S  R  R  C  A  A  G  L  R  R  L  G  V  G  K  G

     CGACGTCGTCATGGCGCTCCTCCGCAACTGCCCCGAGTTCGCCTTCGTGTTCCTCGGCGC
301  ---------+---------+---------+---------+---------+---------+  360
      D  V  V  M  A  L  L  R  N  C  P  E  F  A  F  V  F  L  G  A

     GGCCCGGCTCGGCGCCGCCACCACCACCGCCAACCCGTTCTACACGCCCCACGAGATCCA
361  ---------+---------+---------+---------+---------+---------+  420
      A  R  L  G  A  A  T  T  T  A  N  P  F  Y  T  P  H  E  I  H

     CCGCCAGGCCACCGCCGCCGGGGCCAGGGTCATCGTCACCGAGGCCTGCGCCGTCGAGAA
421  ---------+---------+---------+---------+---------+---------+  480
      R  Q  A  T  A  A  G  A  R  V  I  V  T  E  A  C  A  V  E  K

     GGTGCGCGCCTTCGCCGCCGAGAGAGGGATTCCCGTCGTCTCCGTCGACGAGGGCGTCGA
481  ---------+---------+---------+---------+---------+---------+  540
      V  R  A  F  A  A  E  R  G  I  P  V  V  S  V  D  E  G  V  D

     CGGCGGCTGCCTCCCGTTCGCCGAGACTCTGCTCGGGGAAGAAAGCGGGGAGCGGTTCGT
541  ---------+---------+---------+---------+---------+---------+  600
      G  G  C  L  P  F  A  E  T  L  L  G  E  E  S  G  E  R  F  V

     CGACGAGGCGGTCGACCCCGACGACGTGGTGGCGCTGCCGTACTCGTCCGGCACCACCGG
601  ---------+---------+---------+---------+---------+---------+  660
      D  E  A  V  D  P  D  D  V  V  A  L  P  Y  S  S  G  T  T  G

     CCTGCCCAAGGGCGTCATGCTCACCCACCGCAGCCTCGTCACCAGCGTCGCCCAGCAGGT
661  ---------+---------+---------+---------+---------+---------+  720
      L  P  K  G  V  M  L  T  H  R  S  L  V  T  S  V  A  Q  Q  V

     GGACGGTGAGAACCCGAACCTGCACTTCAGCTCGTCGGACGTGCTGCTGTGCGTGCTGCC
721  ---------+---------+---------+---------+---------+---------+  780
      D  G  E  N  P  N  L  H  F  S  S  S  D  V  L  L  C  V  L  P
```

FIGURE 3

```
     GCTGTTCCACATCTACTCGCTCAACTCGGTGCTGCTCGCCGGTCTCCGCGCCGGGTGCGC
781  ---------+---------+---------+---------+---------+---------+   840
      L  F  H  I  Y  S  L  N  S  V  L  L  A  G  L  R  A  G  C  A

     GATCGTGATCATGCGCAAGTTCGACCACGGCGCGCTGGTGGACCTGGTGCGCACGCACGG
841  ---------+---------+---------+---------+---------+---------+   900
      I  V  I  M  R  K  F  D  H  G  A  L  V  D  L  V  R  T  H  G

     CGTCACCGTGGCGCCATTCGTGCCGCCCATCGTGGTGGAGATCGCCAAGAGCGCGCGGGT
901  ---------+---------+---------+---------+---------+---------+   960
      V  T  V  A  P  F  V  P  P  I  V  V  E  I  A  K  S  A  R  V

     GACCGCCGCGGACCTGGCGTCCATCCGGCTGGTCATGTCGGGGGCGGCGCCCATGGGCAA
961  ---------+---------+---------+---------+---------+---------+   1020
      T  A  A  D  L  A  S  I  R  L  V  M  S  G  A  A  P  M  G  K

     GGAGCTGCAGGACGCGTTCATGGCCAAGATCCCCAACGCCGTGCTCGGCCAGGGATATGG
1021 ---------+---------+---------+---------+---------+---------+   1080
      E  L  Q  D  A  F  M  A  K  I  P  N  A  V  L  G  Q  G  Y  G

     GATGACCGAGGCCGGCCCTGTGCTGGCGATGTGCCTGGCCTTCGCCAAGGAGCCGTTCGC
1081 ---------+---------+---------+---------+---------+---------+   1140
      M  T  E  A  G  P  V  L  A  M  C  L  A  F  A  K  E  P  F  A

     GGTCAAGTCCGGTTCCTGCGGCACCGTCGTCAGGAACGCCGAGCTCAAGATCGTCGACCC
1141 ---------+---------+---------+---------+---------+---------+   1200
      V  K  S  G  S  C  G  T  V  V  R  N  A  E  L  K  I  V  D  P

     CGACACCGGCGCCTCCCTCGGCCGCAACCTGCCGGGGGAGATCTGCATCCGCGGCAAGCA
1201 ---------+---------+---------+---------+---------+---------+   1260
      D  T  G  A  S  L  G  R  N  L  P  G  E  I  C  I  R  G  K  Q

     GATCATGAAAGGTTACCTAAATGATCCGGTGGCCACAAAGAACACCATTGACAAGGACGG
1261 ---------+---------+---------+---------+---------+---------+   1320
      I  M  K  G  Y  L  N  D  P  V  A  T  K  N  T  I  D  K  D  G

     TTGGCTGCATACTGGTGACATTGGTTATGTCGATGATGACGACGAGATCTTTATTGTCGA
1321 ---------+---------+---------+---------+---------+---------+   1380
      W  L  H  T  G  D  I  G  Y  V  D  D  D  D  E  I  F  I  V  D

     CAGACTGAAGGAGATAATTAAATATAAGGGATTCCAAGTACCTCCGGCGGAACTTGAAGC
1381 ---------+---------+---------+---------+---------+---------+   1440
      R  L  K  E  I  I  K  Y  K  G  F  Q  V  P  P  A  E  L  E  A

     CCTTCTCATTACACACCCTGAAATCAAGGATGCTGCTGTCGTATCGATGCAAGACGAACT
1441 ---------+---------+---------+---------+---------+---------+   1500
      L  L  I  T  H  P  E  I  K  D  A  A  V  V  S  M  Q  D  E  L

     TGCTGGTGAAGTTCCGGTTGCGTTTGTTGTGCGGACTGAGGGTTCAGAGATCAGCGAAAA
1501 ---------+---------+---------+---------+---------+---------+   1560
      A  G  E  V  P  V  A  F  V  V  R  T  E  G  S  E  I  S  E  N
```

FIGURE 3 CONTINUED

```
      CGAGATCAAGCAGTTCGTTGCAAAAGAGGTTGTTTTCTACAAGAGGATCTGCAAAGTGTT
1561  ---------+---------+---------+---------+---------+---------+   1620
       E  I  K  Q  F  V  A  K  E  V  V  F  Y  K  R  I  C  K  V  F

      CTTCGCGGATTCCATTCCAAAGAGTCCATCTGGCAAGATCCTCAGGAAGGACCTGAGAGC
1621  ---------+---------+---------+---------+---------+---------+   1680
       F  A  D  S  I  P  K  S  P  S  G  K  I  L  R  K  D  L  R  A

      AAAGCTCGCCGCAGGCATTCCCAGCAGTAATACCACACAGTCCAAAAGCTAAGTCAGATA
1681  ---------+---------+---------+---------+---------+---------+   1740
       K  L  A  A  G  I  P  S  S  N  T  T  Q  S  K  S  *

      TATTGTTTCCCAACCTTACACACCTCTGTCCAACACCATGTAATGTTCTTAATATAAACG
1741  ---------+---------+---------+---------+---------+---------+   1800

      GAAATTATTACATATAGAAGGGCTGATTCTTTTTACTAGATGTGTCCAACATATGATATG
1801  ---------+---------+---------+---------+---------+---------+   1860

      CTTGTTAGGCCGATGATGTGTAACCTGTCATGTATAGATACCGCCTTTTTTTGACAAGAA
1861  ---------+---------+---------+---------+---------+---------+   1920

      AGGCTGATTATAATGTATACCGTGAACTGAATATTTGTTCAGGGAGATCAAAAAAAAAAA
1921  ---------+---------+---------+---------+---------+---------+   1980

      AAAAAAAAAAAA
1981  ----------+--  1992
```

FIGURE 3 CONTINUED

```
     CGGCACGAGATCTCCCACGACTAATTTAGAAGAAGATTTACTTAGTCTCTGCTTCTCGCT
  1  ---------+---------+---------+---------+---------+---------+   60

     CGATCGCCGGCCGGTGAGGTAGCTAGCTAGCTACTCGTACTAGACCATTACCATGGGTTC
 61  ---------+---------+---------+---------+---------+---------+  120
                                                      M  G  S

     CGTGCCGGAGGAGTCAGTGGTGGCGGTGGCACCGGCGGAGACGGTGTTCCGGTCGAAGCT
121  ---------+---------+---------+---------+---------+---------+  180
      V  P  E  E  S  V  V  A  V  A  P  A  E  T  V  F  R  S  K  L

     CCCCGACATCGAGATCAACAACGAGCAGACGCTGCAGAGCTACTGCTTCGAGAAGATGGC
181  ---------+---------+---------+---------+---------+---------+  240
      P  D  I  E  I  N  N  E  Q  T  L  Q  S  Y  C  F  E  K  M  A

     CGAGGTCGCGTCCCGCCCCTGCATCATCGACGGCCAGACGGGCGCCTCCTACACCTACAC
241  ---------+---------+---------+---------+---------+---------+  300
      E  V  A  S  R  P  C  I  I  D  G  Q  T  G  A  S  Y  T  Y  T

     GGAGGTCGACTCCCTGACCCGTCGCGCCGCGGCGGGGCTGCGCCGCATGGGCGTGGGGAA
301  ---------+---------+---------+---------+---------+---------+  360
      E  V  D  S  L  T  R  R  A  A  A  G  L  R  R  M  G  V  G  K

     GGGCGACGTGGTGATGAACCTGCTGCGCAACTGCCCGGAGTTCGCCTTCTCCTTCCTGGG
361  ---------+---------+---------+---------+---------+---------+  420
      G  D  V  V  M  N  L  L  R  N  C  P  E  F  A  F  S  F  L  G

     CGCGGCGCGGCTGGGCGCCGCCACCACCACCGCCAACCCGTTCTACACCCCGCACGAGAT
421  ---------+---------+---------+---------+---------+---------+  480
      A  A  R  L  G  A  A  T  T  T  A  N  P  F  Y  T  P  H  E  I

     CCACCGCCAGGCGGAGGCGGCGGGCGCCAAGCTGATCGTGACCGAGGCCTGCGCCGTGGA
481  ---------+---------+---------+---------+---------+---------+  540
      H  R  Q  A  E  A  A  G  A  K  L  I  V  T  E  A  C  A  V  E

     GAAGGTGCTGGAGTTCGCGGCGGGGCGGGGCGTGCCCGTGGTCACCGTCGACGGGAGGCG
541  ---------+---------+---------+---------+---------+---------+  600
      K  V  L  E  F  A  A  G  R  G  V  P  V  V  T  V  D  G  R  R

     CGACGGGTGCGTGGACTTCGCGGAGCTGATCGCCGGCGAGGAGCTGCCCGAGGCGGACGA
601  ---------+---------+---------+---------+---------+---------+  660
      D  G  C  V  D  F  A  E  L  I  A  G  E  E  L  P  E  A  D  E

     GGCCGGGGTCCTCCCCGACGACGTCGTCGCCCTGCCCTACTCCTCCGGCACCACCGGGCT
661  ---------+---------+---------+---------+---------+---------+  720
      A  G  V  L  P  D  D  V  V  A  L  P  Y  S  S  G  T  T  G  L

     CCCCAAGGGCGTCATGCTCACCCACCGCAGCCTCGTCACCAGCGTCGCCCAGCTGGTCGA
721  ---------+---------+---------+---------+---------+---------+  780
      P  K  G  V  M  L  T  H  R  S  L  V  T  S  V  A  Q  L  V  D
```

FIGURE 4

```
     CGGGTCGAACCCTAACGTGTGCTTCAACAAGGACGACGCGCTGCTGTGCCTGCTGCCGCT
781  ---------+---------+---------+---------+---------+---------+   840
      G  S  N  P  N  V  C  F  N  K  D  D  A  L  L  C  L  L  P  L

     CTTCCACATCTACTCGCTGCACACGGTGCTGCTGGCGGGGCTCCGCGTCGGCGCCGCCAT
841  ---------+---------+---------+---------+---------+---------+   900
      F  H  I  Y  S  L  H  T  V  L  L  A  G  L  R  V  G  A  A  I

     CGTCATCATGCGCAAGTTCGACGTCGGCGCGCTGGTGGACCTCGTCCGCGCGCACCGCAT
901  ---------+---------+---------+---------+---------+---------+   960
      V  I  M  R  K  F  D  V  G  A  L  V  D  L  V  R  A  H  R  I

     CACCATCGCGCCATTCGTGCCGCCCATCGTCGTGGAGATCGCCAAGAGCGACCGCGTCGG
961  ---------+---------+---------+---------+---------+---------+   1020
      T  I  A  P  F  V  P  P  I  V  V  E  I  A  K  S  D  R  V  G

     CGCCGACGACCTCGCATCCATCCGCATGGTGCTCTCCGGCGCCGCGCCCATGGGCAAGGA
1021 ---------+---------+---------+---------+---------+---------+   1080
      A  D  D  L  A  S  I  R  M  V  L  S  G  A  A  P  M  G  K  D

     CCTCCAGGACGCCTTCATGGCCAAGATCCCCAACGCCGTGCTCGGACAGGGGTACGGGAT
1081 ---------+---------+---------+---------+---------+---------+   1140
      L  Q  D  A  F  M  A  K  I  P  N  A  V  L  G  Q  G  Y  G  M

     GACCGAGGCTGGGCCGGTGCTGGCCATGTGCCTGGCGTTCGCCAAGGAGCCGTTCAAGGT
1141 ---------+---------+---------+---------+---------+---------+   1200
      T  E  A  G  P  V  L  A  M  C  L  A  F  A  K  E  P  F  K  V

     CAAGTCCGGGTCGTGCGGAACCGTGGTGCGCAACGCCGAGCTCAAGGTCGTCGACCCCGA
1201 ---------+---------+---------+---------+---------+---------+   1260
      K  S  G  S  C  G  T  V  V  R  N  A  E  L  K  V  V  D  P  D

     CACCGGCGCATCCCTCGGCCGGAACCAGCCTGGCGAGATTTGCGTCCGGGGGAAGCAGAT
1261 ---------+---------+---------+---------+---------+---------+   1320
      T  G  A  S  L  G  R  N  Q  P  G  E  I  C  V  R  G  K  Q  I

     CATGATAGGTTACCTGAACGACCCAGAGTCGACCAAGAACACCATCGACAAGGACGGCTG
1321 ---------+---------+---------+---------+---------+---------+   1380
      M  I  G  Y  L  N  D  P  E  S  T  K  N  T  I  D  K  D  G  W

     GCTGCACACCGGAGACATCGGCTTGGTGGATGACGACGACGAGATCTTCATCGTCGACAG
1381 ---------+---------+---------+---------+---------+---------+   1440
      L  H  T  G  D  I  G  L  V  D  D  D  D  E  I  F  I  V  D  R

     GCTCAAGGAGATCATCAAGTACAAGGGCTTCCAAGTGGCGCCGGCGGAGCTCGAGGCCCT
1441 ---------+---------+---------+---------+---------+---------+   1500
      L  K  E  I  I  K  Y  K  G  F  Q  V  A  P  A  E  L  E  A  L

     CCTCCTCACGAACCCGGAGGTCAAGGACGCCGCCGTCGTAGGGGTGAAGGATGATCTCTG
1501 ---------+---------+---------+---------+---------+---------+   1560
      L  L  T  N  P  E  V  K  D  A  A  V  V  G  V  K  D  D  L  C
```

FIGURE 4 CONTINUED

```
      CGGCGAAGTCCCGGTCGCCTTCATTAAGAGGATCGAAGGATCTGAGATCAACGAGAACGA
1561  ---------+---------+---------+---------+---------+---------+   1620
       G  E  V  P  V  A  F  I  K  R  I  E  G  S  E  I  N  E  N  E

      GATCAAGCAATTCGTCTCAAAGGAGGTTGTTTTCTACAAGAGGATCAACAAGGTCTACTT
1621  ---------+---------+---------+---------+---------+---------+   1680
       I  K  Q  F  V  S  K  E  V  V  F  Y  K  R  I  N  K  V  Y  F

      CACCGACTCCATTCCCAAGAACCCTTCCGGCAAGATCCTAAGGAAGGACTTGAGAGCCAG
1681  ---------+---------+---------+---------+---------+---------+   1740
       T  D  S  I  P  K  N  P  S  G  K  I  L  R  K  D  L  R  A  R

      GCTCGCCGCTGGCATCCCCACCGAAGTTGCCGCGCCGAGAAGCTAAGGGCCGCTTCTCAG
1741  ---------+---------+---------+---------+---------+---------+   1800
       L  A  A  G  I  P  T  E  V  A  A  P  R  S  *

      GAACGCAGTCACCCATGGTGCTGTTTAGGTGCTGTTATAGACCACACCAAATGGGGAAAG
1801  ---------+---------+---------+---------+---------+---------+   1860

      AAACTACGGGAGGGGATCATATTATTGTTGCAGGAGATATCAGTTTGTTGATTCGCCCTG
1861  ---------+---------+---------+---------+---------+---------+   1920

      CTTGTGTAATGTTGATAAAATGAAATGATATAATAGATGTGTTGTTTTATTTTTTGACCA
1921  ---------+---------+---------+---------+---------+---------+   1980

      TGTAAGAACAAGGCTGTTTTATACACTACTTATTTTTTGAAAAAAAAAAAAAAAAAAAA
1981  ---------+---------+---------+---------+---------+--------   2038
```

FIGURE 4 CONTINUED

```
                10        20        30        40        50        60
Lp4CL1  MITVAAPEVQQPQIAAAAAAVEAAAPEATTIFRSRLPDIDIPTHMPLHDYCFATAASAPD
Lp4CL2                MGSIAADAPPAEL..VFRSKLPDIEIPTHLTLQDYCFQRLPELSA
Lp4CL3              MGSVPEESVVAVAPAETVFRSKLPDIEINNEQTLQSYCFEKMAEVAS

                70        80        90       100       110       120
Lp4CL1  APCLITAATGKTYTFAETHLLCRKAAAALHGLGVRHGDRIMLLLQNSVEFALAFFGASML
Lp4CL2  RACLIDGATGAALTYGEVDALSRRCAAGLRRLGVGKGDVVMALLRNCPEFAFVFLGAARL
Lp4CL3  RPCIIDGQTGASYTYTEVDSLTRRAAAGLRRMGVGKGDVVMNLLRNCPEFAFSFLGAARL

               130       140       150       160       170       180
Lp4CL1  GAVSTAANPFCTPQEIHKQLVASGAKLVVTQSAYVDKLRHEAFPRIGEALTVITIDEDDG
Lp4CL2  GAATTTANPFYTPHEIHRQATAAGARVIVTEACAVEKVRAFAAERGIPVVSV......DE
Lp4CL3  GAATTTANPFYTPHEIHRQAEAAGAKLIVTEACAVEKVLEFAAGRGVPVVTV......DG

               190       200       210       220       230       240
Lp4CL1  TPDGCQPFWALVSAADENSVPESPIS..PDDAVALPYSSGTTGLPKGVVLTHGGLVSSVA
Lp4CL2  GVDGGCLPFAETLLGEESGERFVDEAVDPDDVVALPYSSGTTGLPKGVMLTHRSLVTSVA
Lp4CL3  RRDGCVDF.AELIAGEELPEADEAGVL.PDDVVALPYSSGTTGLPKGVMLTHRSLVTSVA

               250       260       270       280       290       300
Lp4CL1  QQVDGENPNLHMRAGEDVVLCVLPLFHIFSLNSVLLCALRAGAAVMLMPRFEMGAMLEGI
Lp4CL2  QQVDGENPNLHFSS.SDVLLCVLPLFHIYSLNSVLLAGLRAGCAIVIMRKFDHGALVDLV
Lp4CL3  QLVDGSNPNVCFNK.DDALLCLLPLFHIYSLHTVLLAGLRVGAAIVIMRKFDVGALVDLV

               310       320       330       340       350       360
Lp4CL1  ERWRVTVAAVVPPLVLALAKNPGVEKHDLSSIRIVLSGAAPLGKELEDALRGRLPQAIFG
Lp4CL2  RTHGVTVAPFVPPIVVEIAKSARVTAADLASIRLVMSGAAPMGKELQDAFMAKIPNAVLG
Lp4CL3  RAHRITIAPFVPPIVVEIAKSDRVGADDLASIRMVLSGAAPMGKDLQDAFMAKIPNAVLG

               370       380       390       400       410       420
Lp4CL1  QGYGMTEAGPVLSMCPAFAREPTPAKSGSCGTVVRNAQLKVVDPDTGVSLGRNLPGEICI
Lp4CL2  QGYGMTEAGPVLAMCLAFAKEPFAVKSGSCGTVVRNAELKIVDPDTGASLGRNLPGEICI
Lp4CL3  QGYGMTEAGPVLAMCLAFAKEPFKVKSGSCGTVVRNAELKVVDPDTGASLGRNQPGEICV

               430       440       450       460       470       480
Lp4CL1  RGPQIMKGYLNDPVATAATIDVEGWLHTGDIGYVDDDDEVFIVDRVKELIKFKGFQVPPA
Lp4CL2  RGKQIMKGYLNDPVATKNTIDKDGWLHTGDIGYVDDDDEIFIVDRLKEIIKYKGFQVPPA
Lp4CL3  RGKQIMIGYLNDPESTKNTIDKDGWLHTGDIGLVDDDDEIFIVDRLKEIIKYKGFQVAPA

               490       500       510       520       530       540
Lp4CL1  ELEALLIAHPSIADAAVVPQKDDAAGEVPVAFVVRAADSDIAEEAIKEFVSKQVVFYKRL
Lp4CL2  ELEALLITHPEIKDAAVVSMQDELAGEVPVAFVVRTEGSEISENEIKQFVAKEVVFYKRI
Lp4CL3  ELEALLLTNPEVKDAAVVGVKDDLCGEVPVAFIKRIEGSEINENEIKQFVSKEVVFYKRI

               550       560       570
Lp4CL1  HKVYFTHAIPKSASGKILRKELRAKLAAPATA
Lp4CL2  CKVFFADSIPKSPSGKILRKDLRAKLAAGIPSSNTTQSKS
Lp4CL3  NKVYFTDSIPKNPSGKILRKDLRARLAAGIPTEVAAPRS
```

FIGURE 5

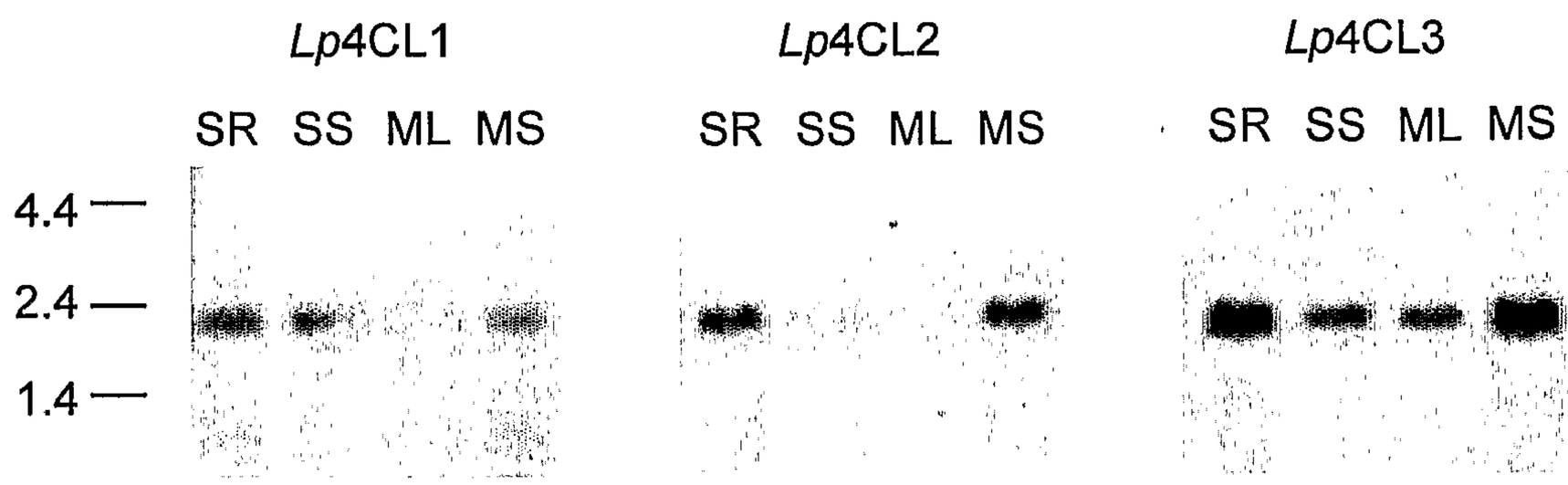


FIGURE 6

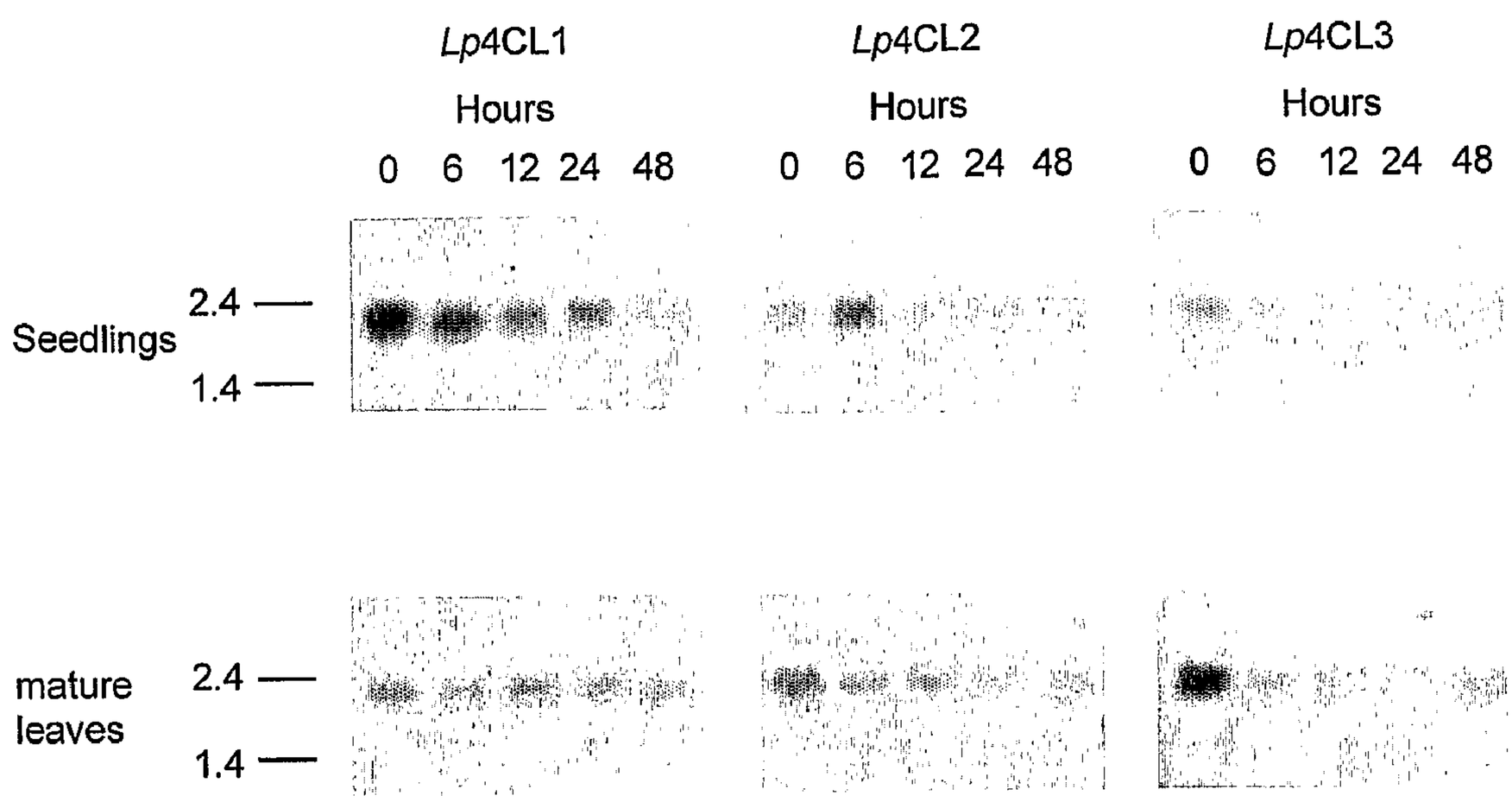

Lp4CL1
Hours
0 6 12 24 48
Lp4CL2
Hours
0 6 12 24 48
Lp4CL3
Hours
0 6 12 24 48
Seedlings
2.4
1.4
mature leaves
2.4
1.4

FIGURE 7

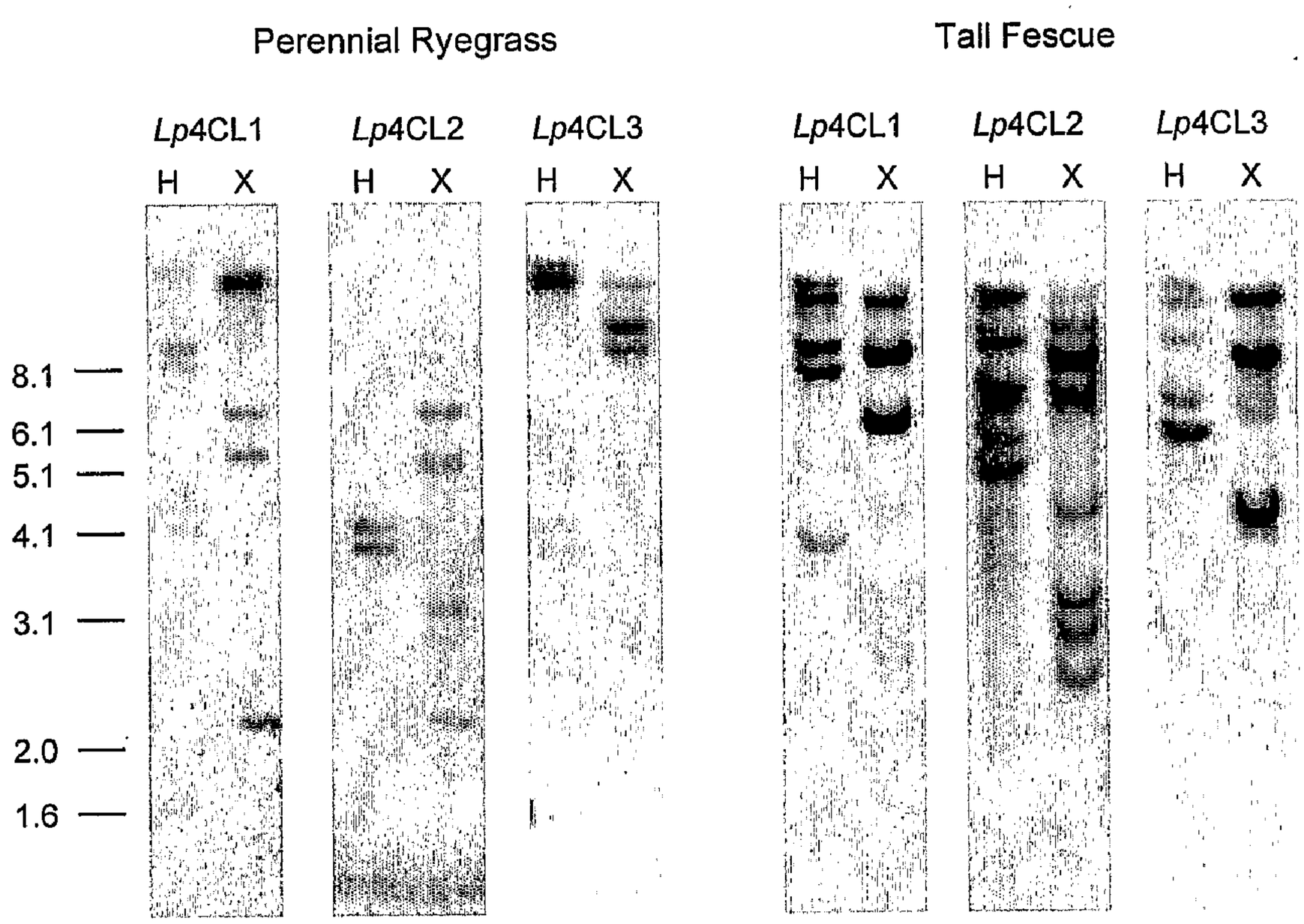

Perennial Ryegrass
Tall Fescue
Lp4CL1
Lp4CL2
Lp4CL3
Lp4CL1
Lp4CL2
Lp4CL3
H X
H X
H X
H X
H X
H X
8.1
6.1
5.1
4.1
3.1
2.0
1.6

FIGURE 8

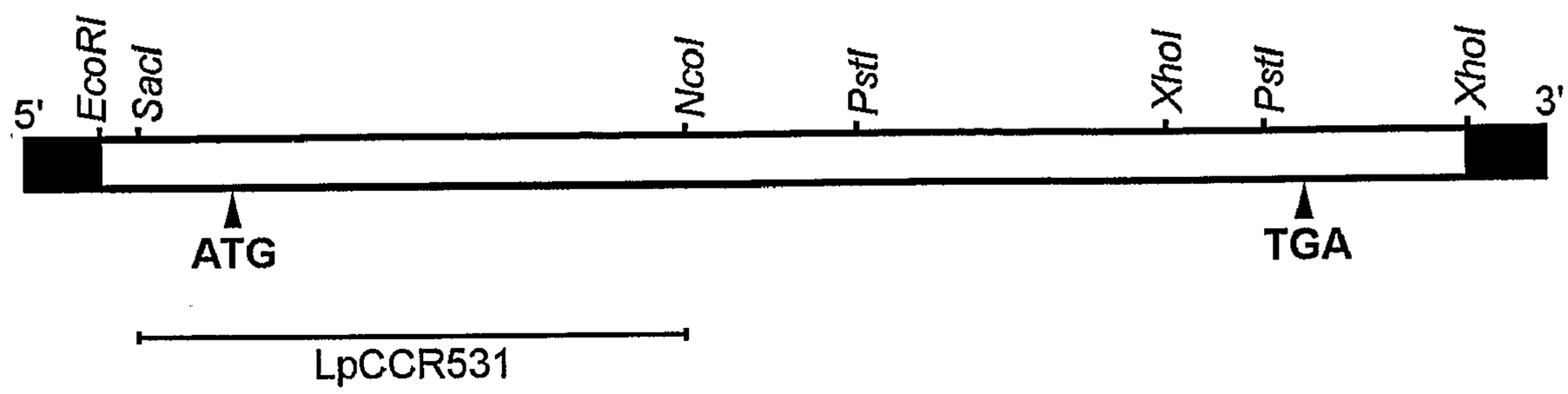


FIGURE 9

```
     GGCACGAGGAATCCTACCAAACCGAGCTACCAGATCCTTCTCTACTAATCGAGCTCCCTA
  1  ---------+---------+---------+---------+---------+---------+   60

     CGCTGCTCCGCCTGTCTTCGTTTCCGCCTCACCGCCGGCCGGTTCTCCGCTCCAAGCTAC
 61  ---------+---------+---------+---------+---------+---------+  120

     GTCCGTCCGTCCACATATATAGCATCGACATGACCATCGCCGAGGTCGTGGCTGCCGGAG
121  ---------+---------+---------+---------+---------+---------+  180
                                   M  T  I  A  E  V  V  A  A  G  D

     ACACCGCCGCCGCGGTGGTGCAGCCCGCCGGGAACGGGCAGACCGTGTGCGTGACCGGCG
181  ---------+---------+---------+---------+---------+---------+  240
       T  A  A  A  V  V  Q  P  A  G  N  G  Q  T  V  C  V  T  G  A

     CCGCCGGGTACATCGCGTCGTGGCTCGTCAAGCTGCTGCTGGAGAAGGGGTACACCGTCA
241  ---------+---------+---------+---------+---------+---------+  300
       A  G  Y  I  A  S  W  L  V  K  L  L  L  E  K  G  Y  T  V  K

     AGGGCACCGTCAGGAACCCAGACGACCCGAAGAACGCGCACCTGAGGGCGCTCGACGGCG
301  ---------+---------+---------+---------+---------+---------+  360
       G  T  V  R  N  P  D  D  P  K  N  A  H  L  R  A  L  D  G  A

     CCGCCGACCGGCTGGTCCTCTGCAAGGCCGACCTCCTCGACTACGACGCCATCCGCCGCG
361  ---------+---------+---------+---------+---------+---------+  420
       A  D  R  L  V  L  C  K  A  D  L  L  D  Y  D  A  I  R  R  A

     CCATCGACGGCTGCCACGGCGTCTTCCACACCGCGTCCCCCGTCACCGACGACCCCGAGC
421  ---------+---------+---------+---------+---------+---------+  480
       I  D  G  C  H  G  V  F  H  T  A  S  P  V  T  D  D  P  E  Q

     AAATGGTGGAGCCGGCGGTGAGGGGCACGCAGTACGTCATAGACGCGGCGGCGGAGGCCG
481  ---------+---------+---------+---------+---------+---------+  540
       M  V  E  P  A  V  R  G  T  Q  Y  V  I  D  A  A  A  E  A  G

     GCACGGTGCGGCGGATGGTGCTCACCTCCTCCATCGGCGCCGTCACCATGGACCCCAACC
541  ---------+---------+---------+---------+---------+---------+  600
       T  V  R  R  M  V  L  T  S  S  I  G  A  V  T  M  D  P  N  R

     GCGGGCCGGACGTGGTCGTCGACGAGTCGTGCTGGAGCGACCTCGACTTCTGCAAGAAAA
601  ---------+---------+---------+---------+---------+---------+  660
       G  P  D  V  V  V  D  E  S  C  W  S  D  L  D  F  C  K  K  T

     CCAGGAACTGGTACTGCTACGGGAAGGCGGTTGCGGAGCAGGCGGCATCGGAGTTGGCGC
661  ---------+---------+---------+---------+---------+---------+  720
       R  N  W  Y  C  Y  G  K  A  V  A  E  Q  A  A  S  E  L  A  R

     GGCAGCGCGGCGTGGACCTTGTGGTGGTGAACCCGGTGCTGGTGATCGGCCCCCTGCTGC
721  ---------+---------+---------+---------+---------+---------+  780
       Q  R  G  V  D  L  V  V  V  N  P  V  L  V  I  G  P  L  L  Q
```

FIGURE 10

```
     AGCCGACGGTGAACGCCAGCATCGGCCACATCCTCAAGTACCTGGACGGGTCGGCCAGCA
781  ---------+---------+---------+---------+---------+---------+   840
       P  T  V  N  A  S  I  G  H  I  L  K  Y  L  D  G  S  A  S  K

     AGTTCGCCAACGCCGTGCAGGCGTACGTGGACGTCCGCGACGTGGCCGACGCCCACCTCC
841  ---------+---------+---------+---------+---------+---------+   900
       F  A  N  A  V  Q  A  Y  V  D  V  R  D  V  A  D  A  H  L  R

     GCGTCTTCGAGTGCGCCGCCGCGTCCGGCCGCCACCTCTGCGCCGAGCGCGTCCTCCACC
901  ---------+---------+---------+---------+---------+---------+   960
       V  F  E  C  A  A  A  S  G  R  H  L  C  A  E  R  V  L  H  R

     GCGAGGACGTCGTGCGCATCCTCGCCAAGCTCTTCCCCGAGTACCCCGTCCCCACCAGGT
961  ---------+---------+---------+---------+---------+---------+   1020
       E  D  V  V  R  I  L  A  K  L  F  P  E  Y  P  V  P  T  R  C

     GCTCTGATGAGACGAACCCGAGGAAGCAGCCATACAAGATGTCGAACCAGAAGCTCCAGG
1021 ---------+---------+---------+---------+---------+---------+   1080
       S  D  E  T  N  P  R  K  Q  P  Y  K  M  S  N  Q  K  L  Q  D

     ACCTCGGACTCGAGTTCAGGCCGGTGAGCCAGTCCCTGTACGAGACGGTGAAGAGCCTCC
1081 ---------+---------+---------+---------+---------+---------+   1140
       L  G  L  E  F  R  P  V  S  Q  S  L  Y  E  T  V  K  S  L  Q

     AGGAGAAGGGCCACCTTCCGGTGCTCAGCGAGCAGGCAGAGGCGGACAAGGAAACCCTAG
1141 ---------+---------+---------+---------+---------+---------+   1200
       E  K  G  H  L  P  V  L  S  E  Q  A  E  A  D  K  E  T  L  A

     CTGCCGAGCTGCAGGCAGGGGTTACCATCCGAGCATGAGGAACAAGAAATCAACCATGTC
1201 ---------+---------+---------+---------+---------+---------+   1260
       A  E  L  Q  A  G  V  T  I  R  A  *

     CATACTGCTACTGTCATGTAAACCAGCTGTTGAATGCCTAAAATCTAAGTTCTTGTAATA
1261 ---------+---------+---------+---------+---------+---------+   1320

     CTGTGTTGTTTCATGTGGACTAGATTGATCGAATAAACATCTCTACACAAGGTTGCTAAA
1321 ---------+---------+---------+---------+---------+---------+   1380

     AAAAAAAAAAAAAAA
1381 ---------+-----  1395
```

FIGURE 10 CONTINUED

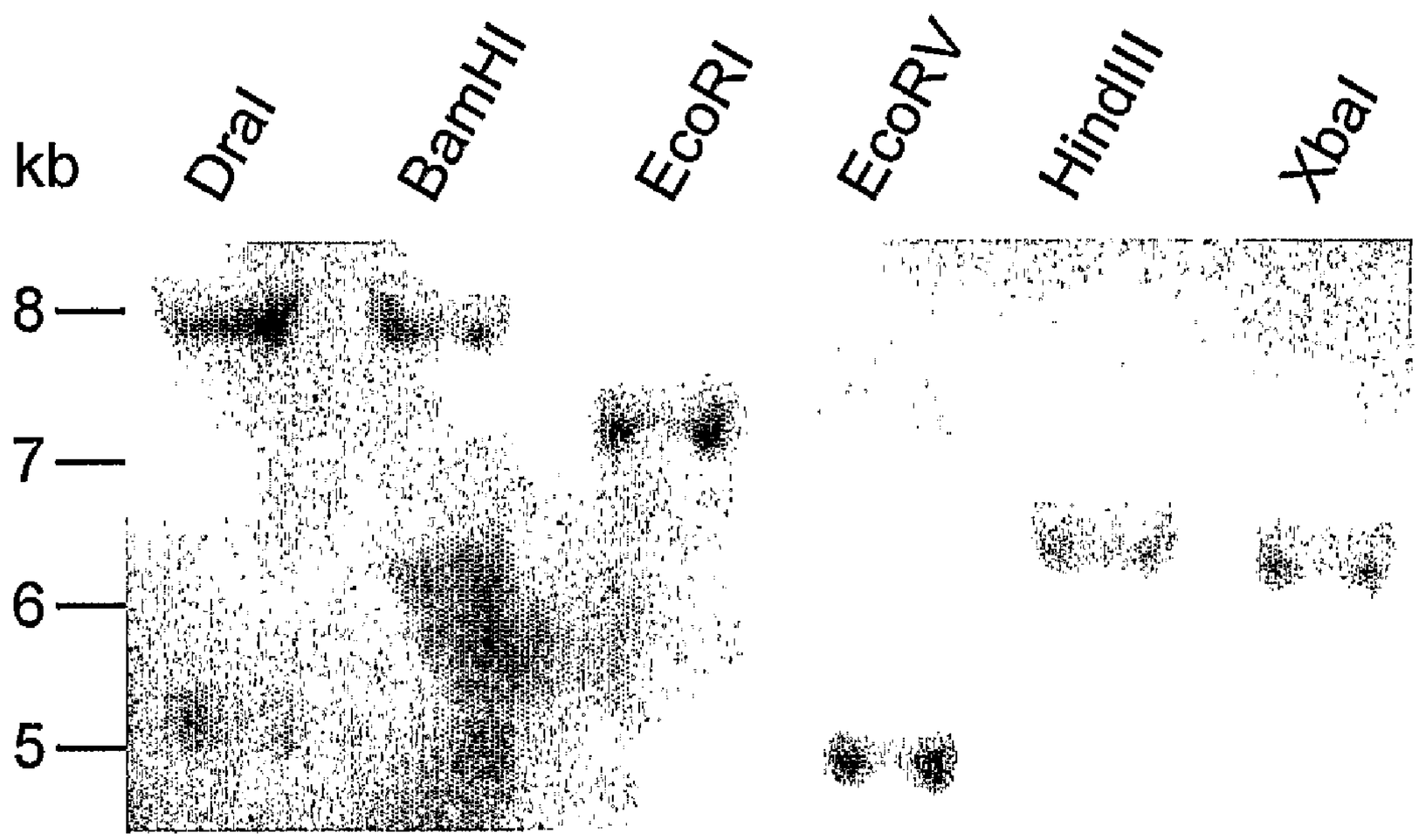


FIGURE 11

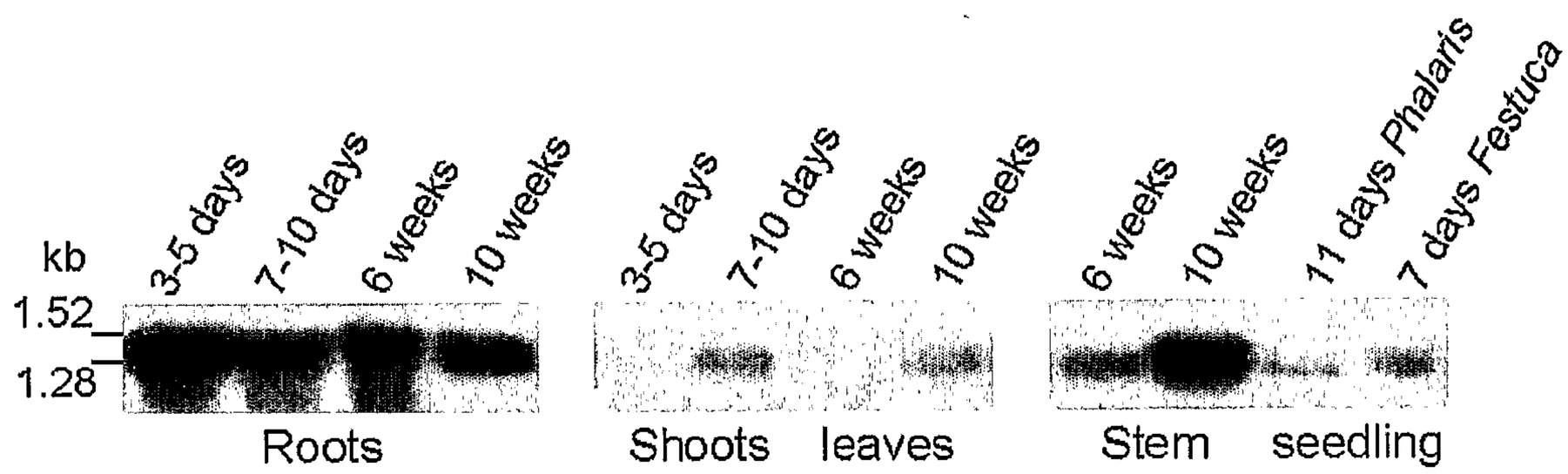

kb
1.52
1.28
3-5 days
7-10 days
6 weeks
10 weeks
Roots
3-5 days
7-10 days
6 weeks
10 weeks
Shoots leaves
6 weeks
10 weeks
11 days Phalaris
7 days Festuca
Stem seedling

FIGURE 12

```
     GGCACGAGCAACAAGTCATCAATGGCGGAAGGCTTGCCGGCGCTCGGTTGGGCTGCGAGG
  1  ---------+---------+---------+---------+---------+---------+   60
                          M  A  E  G  L  P  A  L  G  W  A  A  R

     GACGCCTCCGGTCACCTCTCCCCTTACAGCTTCTCGAGAAGCGTTCCGAAGGACGACgAT
 61  ---------+---------+---------+---------+---------+---------+  120
     D  A  S  G  H  L  S  P  Y  S  F  S  R  S  V  P  K  D  D  D

     GTGACGATCAAGGTGCTCTTCTGCGGGATCTGCCACACTGACCTCCACATCATCAAGAAC
121  ---------+---------+---------+---------+---------+---------+  180
     V  T  I  K  V  L  F  C  G  I  C  H  T  D  L  H  I  I  K  N

     GACTGGGGCAACGCCCTCTACCCCATCGTCCCAGGGCATGAGATCGTGGGCGTCGTCGCC
181  ---------+---------+---------+---------+---------+---------+  240
     D  W  G  N  A  L  Y  P  I  V  P  G  H  E  I  V  G  V  V  A

     AGCGTCGGCAGCGGCGTCAGCAGCTTCAAGGCCGGCgACACGGTGGGCGTGGGCTACTTC
241  ---------+---------+---------+---------+---------+---------+  300
     S  V  G  S  G  V  S  S  F  K  A  G  D  T  V  G  V  G  Y  F

     CTCGACTCCTGCCGCACCTGCTACAGCTGCAGCAAGGGGTACGAGAACTTCTGCCCCACC
301  ---------+---------+---------+---------+---------+---------+  360
     L  D  S  C  R  T  C  Y  S  C  S  K  G  Y  E  N  F  C  P  T

     CTGACGCTCACCTCCAACGGCGTCGACGGCGGCGGCGCCACCACCCAGGGCGGCTTCTCC
361  ---------+---------+---------+---------+---------+---------+  420
     L  T  L  T  S  N  G  V  D  G  G  G  A  T  T  Q  G  G  F  S

     GACGTCCTCGTCGTCAACAAGGACTACGTCATCCGCGTCCCGGACAACCTGCCCCTGGCC
421  ---------+---------+---------+---------+---------+---------+  480
     D  V  L  V  V  N  K  D  Y  V  I  R  V  P  D  N  L  P  L  A

     GGCGCGGCACCTCTCCTCTGCGCCGGCGTCACAGTCTACAGCCCTATGGTGGAGTACGGC
481  ---------+---------+---------+---------+---------+---------+  540
     G  A  A  P  L  L  C  A  G  V  T  V  Y  S  P  M  V  E  Y  G

     CTCAACGCCCCcgGGAAGCACyTCGGcGTCGTCGGCCTGGGCGGGCTCGGCCACGTCGcC
541  ---------+---------+---------+---------+---------+---------+  600
     L  N  A  P  G  K  H  X  G  V  V  G  L  G  G  L  G  H  V  A

     GTCAAGTTCGGCAAGGCCTTCGGGATGACCGTCACCGTCATCAGCTCCTCGGACAGGAAG
601  ---------+---------+---------+---------+---------+---------+  660
     V  K  F  G  K  A  F  G  M  T  V  T  V  I  S  S  S  D  R  K

     CGCGACGAGGCGCTCGGCCGCCTCGGCGCCGACGCcTTCCTCGTCAGCAGCGACCCCGAG
661  ---------+---------+---------+---------+---------+---------+  720
     R  D  E  A  L  G  R  L  G  A  D  A  F  L  V  S  S  D  P  E
```

FIGURE 13

```
     CAGATGAAGGCGGCGGCGGGCACCATGGACGGCATCATCGACACGGTGTCCGCGGGCCAC
721  ---------+---------+---------+---------+---------+---------+   780
     Q  M  K  A  A  A  G  T  M  D  G  I  I  D  T  V  S  A  G  H

     CCGATCGTGCCGCTGCTCGACCTGCTCAAGCCCATGGGGCAGATGGTCGTGGTGGGCGCG
781  ---------+---------+---------+---------+---------+---------+   840
     P  I  V  P  L  L  D  L  L  K  P  M  G  Q  M  V  V  V  G  A

     CCCAGCAAGCCGCTCGAGCTCCCGGCCTTCGCCATCATCGGCGGCGGCAAGCGCCTCGCC
841  ---------+---------+---------+---------+---------+---------+   900
     P  S  K  P  L  E  L  P  A  F  A  I  I  G  G  G  K  R  L  A

     GGGAGCGGCACCGGCAGCGTCGCACACTGCCagGCCATGCTCGACTTCGCGGGCAAGCAC
901  ---------+---------+---------+---------+---------+---------+   960
     G  S  G  T  G  S  V  A  H  C  Q  A  M  L  D  F  A  G  K  H

     GGCATCACCGCCGACGTCGAGGTCGTCAAGATGGACTACgGTCAACACCGCCATCGAGCG
961  ---------+---------+---------+---------+---------+---------+   1020
     G  I  T  A  D  V  E  V  V  K  M  D  Y  G  Q  H  R  H  R  A

     GCTAGAGAAGAACGACGTCAGGTACCGCTTCGTCATCGACGTCGCCGGCAGCCACCTGCA
1021 ---------+---------+---------+---------+---------+---------+   1080
     A  R  E  E  R  R  Q  V  P  L  R  H  R  R  R  R  Q  P  P  A

     GGGCACCGCCGCTTAACTTGTGCTACACAATGTGGACGCGCGCTCGTTTGGTCCAGAAAA
1081 ---------+---------+---------+---------+---------+---------+   1140
     G  H  R  R  L  T  C  A  T  Q  C  G  R  A  L  V  W  S  R  K

     AGGTTCGCCGGCTCACAGCCACATGAACAAGTCAATGAGTCGTTGGTGTGTTGTTTATCT
1141 ---------+---------+---------+---------+---------+---------+   1200
     R  F  A  G  S  Q  P  H  E  Q  V  N  E  S  L  V  C  C  L  S

     TCATTCCACATATGGGACGCAGTTCCAGATTTTCATGTCAAATAATTGCGTCGTGTGCGG
1201 ---------+---------+---------+---------+---------+---------+   1260
     S  F  H  I  W  D  A  V  P  D  F  H  V  K

     TTGTCAAGACTCAAATAGGAGAAAAAAAGACTCGTGATTTCGTTTTGCAAAAAAAAAAAA
1261 ---------+---------+---------+---------+---------+---------+   1320

     AAAAA
1321 -----  1325
```

FIGURE 13 CONTINUED

```
     GGCACGAGTCGCCTCCAACGTCTTCCCTTAACCGGCCGTCCCTACGCtTGCACCACCACC
  1  ---------+---------+---------+---------+---------+---------+   60

     ACGCACAGACAGAGCAGTTTCCCAGCCCCCGCCGGAACCGGATGGCACCCACGGCGGCGG
 61  ---------+---------+---------+---------+---------+---------+  120
                                                M  A  P  T  A  A  E

     AGCAGACGGAGCACCACCAGCACACCAGGAAGGCGGTGGGGCTGGCGGCGCGCGACGACG
121  ---------+---------+---------+---------+---------+---------+  180
       Q  T  E  H  H  Q  H  T  R  K  A  V  G  L  A  A  R  D  D  A

     CCGGCCACCTCTCCCCGCTCGCCATCACACGGAGGAGCACAGGAGACGACGATGTGGTGA
181  ---------+---------+---------+---------+---------+---------+  240
       G  H  L  S  P  L  A  I  T  R  R  S  T  G  D  D  D  V  V  I

     TAAAGATTTTGTACTGCGGAATCTGCCACTCTGACCTGCACGCCCTGAAGAACGACTGGA
241  ---------+---------+---------+---------+---------+---------+  300
       K  I  L  Y  C  G  I  C  H  S  D  L  H  A  L  K  N  D  W  K

     AGAACTCAAGGTACCCGATGATCCCCGGGCACGAGATCGCCGGCGAGGTCACGGAGGTGG
301  ---------+---------+---------+---------+---------+---------+  360
       N  S  R  Y  P  M  I  P  G  H  E  I  A  G  E  V  T  E  V  G

     GCAAGAACGTGAGCAAGTTCAAGGCCGGCGACCGCGTGGGCGTCGGGTGCATGGTGAACT
361  ---------+---------+---------+---------+---------+---------+  420
       K  N  V  S  K  F  K  A  G  D  R  V  G  V  G  C  M  V  N  S

     CGTGCCGGTCGTGCGAGAGCTGCGACAAGGGCTTCGAGAACCACTGCCCGGGCATGATCC
421  ---------+---------+---------+---------+---------+---------+  480
       C  R  S  C  E  S  C  D  K  G  F  E  N  H  C  P  G  M  I  L

     TCACCTACAACTCGGTCGACGTCGACGGCACCGTCACCTACGGCGGCTACTCCAGCATGG
481  ---------+---------+---------+---------+---------+---------+  540
       T  Y  N  S  V  D  V  D  G  T  V  T  Y  G  G  Y  S  S  M  V

     TGGTGGTGCACGAGCGGTTCGTGGTCCGGTTCCCCGACGCCATGCCGCTGGACAAGGGCG
541  ---------+---------+---------+---------+---------+---------+  600
       V  V  H  E  R  F  V  V  R  F  P  D  A  M  P  L  D  K  G  A

     CGCCGCTGCTGTGCGCCGGCATCACCGTGTACAGCCCCATGAAGTACCACGGGCTCAACG
601  ---------+---------+---------+---------+---------+---------+  660
       P  L  L  C  A  G  I  T  V  Y  S  P  M  K  Y  H  G  L  N  V

     TTCCCGGGCTGCACCTCGGCGTGCTGGGGCTGGGCGGGCTGGGCCACGTTGCGGTCAAGT
661  ---------+---------+---------+---------+---------+---------+  720
       P  G  L  H  L  G  V  L  G  L  G  G  L  G  H  V  A  V  K  F

     TCGGCAAGGCCTTCGGAATGAAAGTGACGGTGATCAGCTCGTCGCCGGGGAAGAAGGAGG
721  ---------+---------+---------+---------+---------+---------+  780
       G  K  A  F  G  M  K  V  T  V  I  S  S  S  P  G  K  K  E  E
```

FIGURE 14

```
     AGGCCCTGGGGCGGCTGGGCGCCGACGCGTTCATCGTCAGCAAGGACGCCGACGAGATGA
781  ---------+---------+---------+---------+---------+---------+   840
      A  L  G  R  L  G  A  D  A  F  I  V  S  K  D  A  D  E  M  K

     AGGCTGTGATAGCACCATGGATGGCATCANTAAACACGGTATCTGCAAACATCCCCCTGA
841  ---------+---------+---------+---------+---------+---------+   900
      A  V  I  A  P  W  M  A  S  X  N  T  V  S  A  N  I  P  L  T

     CCCCTCTCTTCGGGCTGCTCAAGCCCAACGGCAAGATGATCATGGTCGGCCTCCCCGAGA
901  ---------+---------+---------+---------+---------+---------+   960
      P  L  F  G  L  L  K  P  N  G  K  M  I  M  V  G  L  P  E  K

     AGCCCATCGAGATTCCTCCCTTCGCTCTAGTTGCCACGAATAAGACCCTGGCCGGGAGCA
961  ---------+---------+---------+---------+---------+---------+   1020
      P  I  E  I  P  P  F  A  L  V  A  T  N  K  T  L  A  G  S  I

     TCATCGGCGGCATGAGCGACACGCAGGAGATGCTGGACCTCGCGGCGAAGCACGGCGTGA
1021 ---------+---------+---------+---------+---------+---------+   1080
      I  G  G  M  S  D  T  Q  E  M  L  D  L  A  A  K  H  G  V  T

     CGGCCGACATCGAGGTGGTCGGCGCGGAGTATGTGAACACGGCCTTGGAGCGCCTTGCCA
1081 ---------+---------+---------+---------+---------+---------+   1140
      A  D  I  E  V  V  G  A  E  Y  V  N  T  A  L  E  R  L  A  K

     AGAACGACGTCAGGTATCGCTTCGTCATCGACATCGGCAACACCCTCGACAATGTTGCGG
1141 ---------+---------+---------+---------+---------+---------+   1200
      N  D  V  R  Y  R  F  V  I  D  I  G  N  T  L  D  N  V  A  A

     CCACCACCGAGTGAACGTACTCAGCACTGCTTACGATCTACGTTGTTCCACTGTTAGTGC
1201 ---------+---------+---------+---------+---------+---------+   1260
      T  T  E  *

     TCCGTAGTAAACAATAAACGATCAAAACTCTTGTCATCTGGTGCATTGGTGTAGACATGG
1261 ---------+---------+---------+---------+---------+---------+   1320

     TTGTTTGCGAGGAAACTGAGTTGAAGGATGGATGGATAAAAAAAAAAAAAAAAAAAAA
1321 ---------+---------+---------+---------+---------+--------   1378
```

FIGURE 14 CONTINUED

FIGURE 15

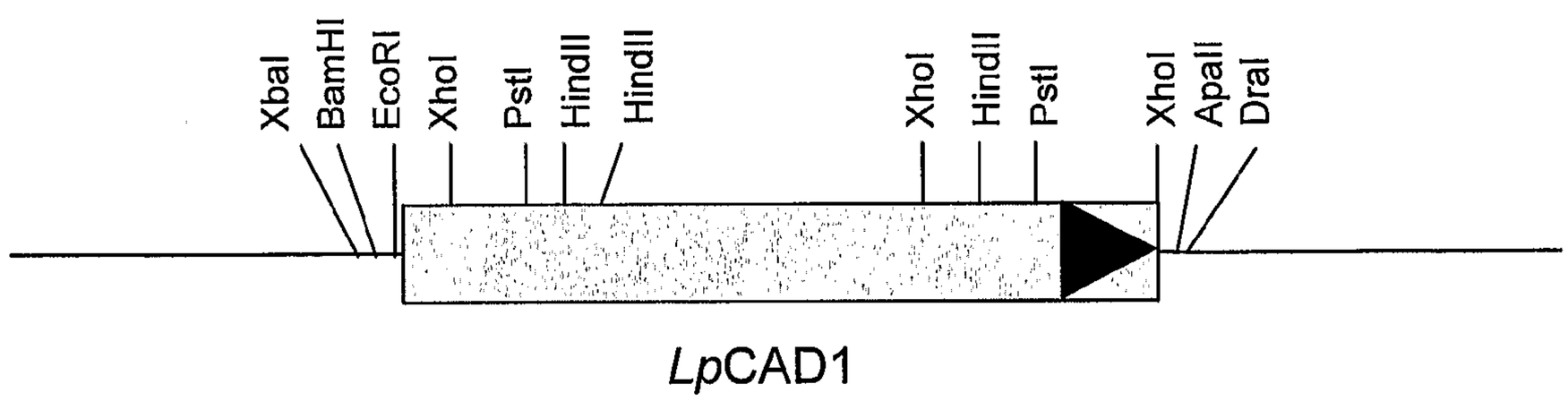
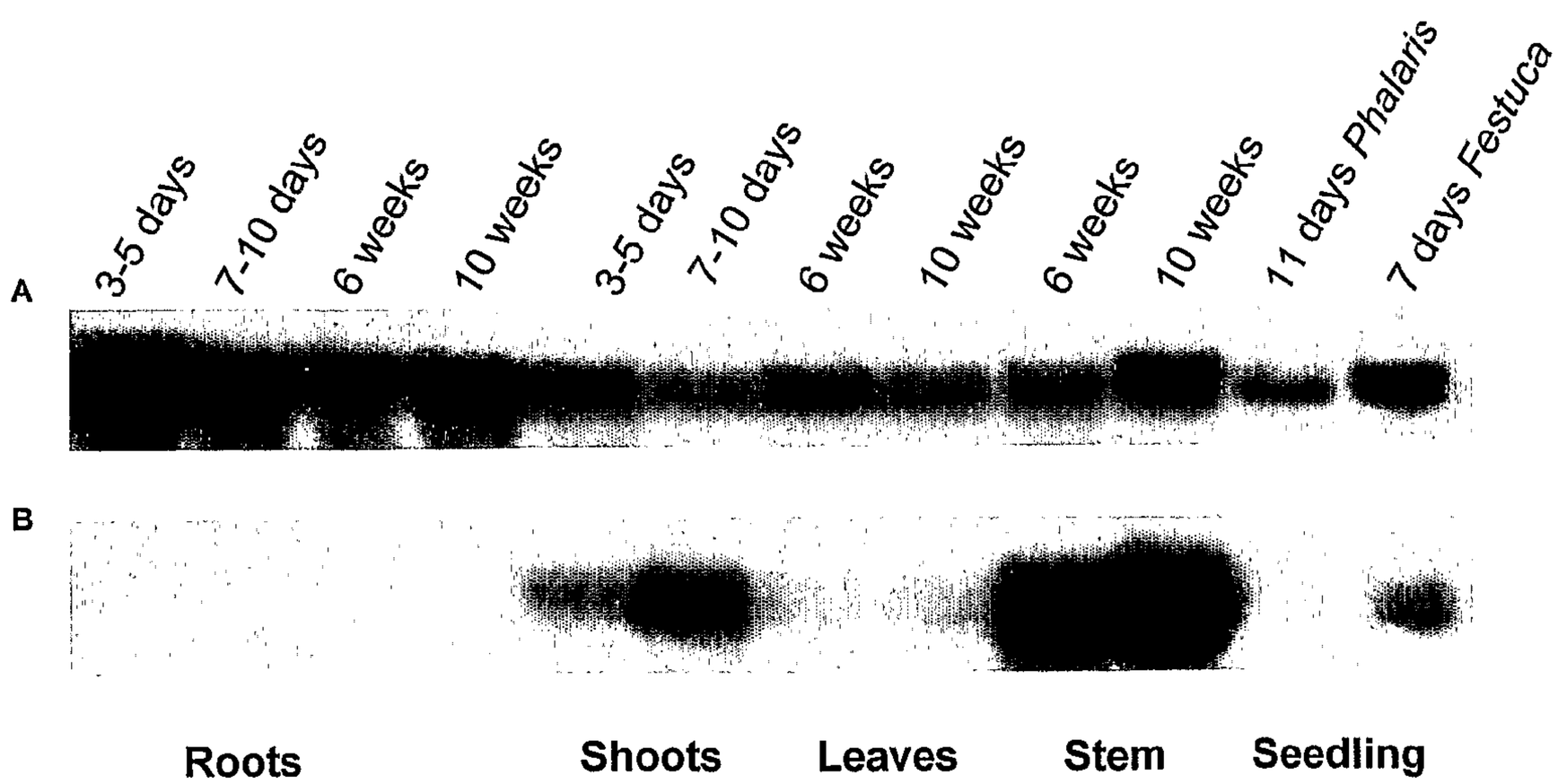


FIGURE 16



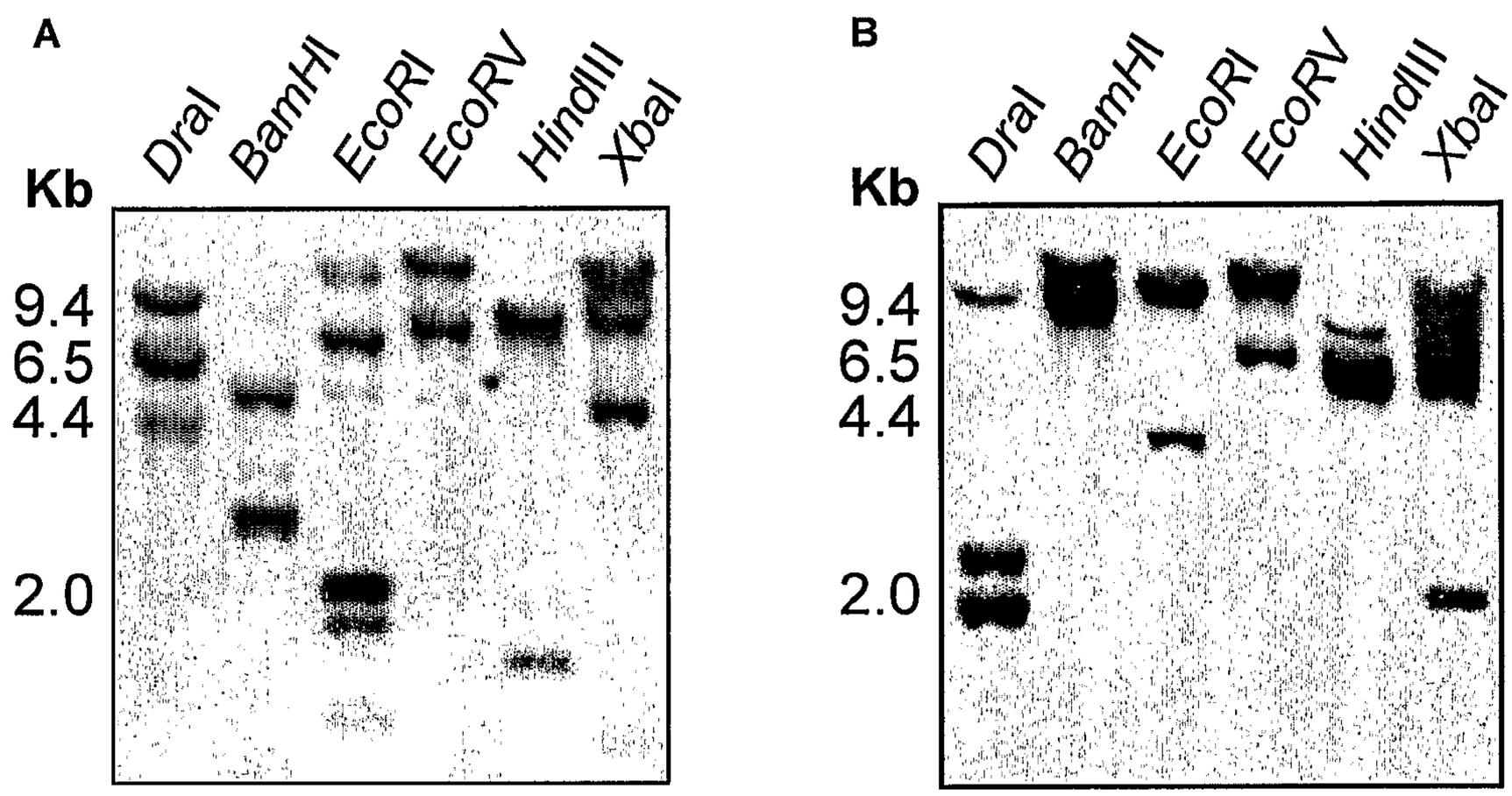

A
DraI
BamHI
EcoRI
EcoRV
HindIII
XbaI
Kb
9.4
6.5
4.4
2.0
B
DraI
BamHI
EcoRI
EcoRV
HindIII
XbaI
Kb
9.4
6.5
4.4
2.0

FIGURE 17

```
                          pBluescript
        GCGGCCGCTCTAAAACTAGTGGATCCCCCGGGCTGCAGGAATTCGATATCAAGCTTATnG
              SalI
        ATACCGTCGACAGCGGTTnCAAATCGCCGGTCCTGGGGTGGAAGTGnAGCAGTGGGAAGA
        +---------+---------+---------+---------+---------+---------   -4581

        TGTGTGCGAGGGGTTGTGTTTTGGATGnAAGACAGGCGGGCCAGTGGAGAACAAGAGAGA
-4580   +---------+---------+---------+---------+---------+---------   -4521

        ACGCGAGAGGCCAAAGTATCCGCAGCCCCGCAAACAAGGCCTAGATTTGGGTTAAGTTTG
-4520   +---------+---------+---------+---------+---------+---------   -4461

        GGTCGTCTCAGACACCGCGGCCATCCTTTTAGGTGGTCCGCGCGCTGGACCGTATTTTTA
-4460   +---------+---------+---------+---------+---------+---------   -4401

        TCTGAGTTGACCCATTCAGACGCGCAGACACGAGATGGATGGTGCAGTwAgAGATGACCT
-4400   +---------+---------+---------+---------+---------+---------   -4341

                                                              HindIII
        AAGTACAArAACCTCTCCCCGA.GCTGCCGCCATCcGTCACTTACCGAGCGAcAAAGcTT
-4340   +---------+---------+---------+---------+---------+---------   -4281

        CCCACTTCATCACACTCAGCCCAGCAAGCATACTGATGGTGAGCGCACTCGCGGCTGTGC
-4280   +---------+---------+---------+---------+---------+---------   -4221

        CCACCGACCCCACGCCATCCAAAACCAACTCTACTTTTCACCmCACCAACAAAAGACAAA
-4220   +---------+---------+---------+---------+---------+---------   -4161

        ATATGGTGGATTTTGTGATGAGATGGAAGCGGAGCTTGTCAGAATGGGAAACGCATAAAT
-4160   +---------+---------+---------+---------+---------+---------   -4101

        CGAGAACACGTATACAGTGCTGGAAATTGGATGACTAAGCCCCAAGGGTTAGAAAAAAAA
-4100   +---------+---------+---------+---------+---------+---------   -4041

                XbaI
        TnAGACCATGTCTAGATGGAATTAGACATTTTTTGATATAATAGAAGCGGGACTTGGCGC
-4040   +---------+---------+---------+---------+---------+---------   -3981

        GACAATTTCAAACTTCGTCCCTAACAGGTATCGAACtTTCGAtAGTTAGCGTGTGCTACT
-3980   +---------+---------+---------+---------+---------+---------   -3921

        GCggAcCCCCAACCACtTGTGTTAAGCCCACATCgGTTAAGGCCCAAGGGTTAGATGAAA
-3920   +---------+---------+---------+---------+---------+---------   -3861

        GTACCAATCTCACTCATTTGCGACTAGCTACAAAACTTGCTTTTCACATGTACGGTCATA
-3860   +---------+---------+---------+---------+---------+---------   -3801

        CTACAATTTTGACCTTGGTAACGTAAGTATGGACTGTATGGTGTGCTAAGGTGTGTTGGC
-3800   +---------+---------+---------+---------+---------+---------   -3741

        AGCTCAAATAAACCCAAAAATTTCAACACACGTCAACCATGAACTGAGATTCACACCAAC
-3740   +---------+---------+---------+---------+---------+---------   -3681

        GGCTGAGCCGTCTCCTTTAAAAGATAGAGGGAGAAAACCATAATCACCATTGGTGGTCAT
-3680   +---------+---------+---------+---------+---------+---------   -3621

        GTGTGAGTGTGCAAGCAAAAAAAAATGGAGAAGCCAAAACCCGTTGAGAGAGTGCGAGAG
-3620   +---------+---------+---------+---------+---------+---------   -3561

        CATACAAGAACACCACAACAAAGTGTGAAGGAGAAAAAGAATATGAGATAAGATTTCGGA
-3560   +---------+---------+---------+---------+---------+---------   -3501
```

FIGURE 18

```
       AATACTTTTGCACACCCATGCATGGGTGTGGGTGTTTCCGTCACCGTCTATGTATTTCTC
-3500  +---------+---------+---------+---------+---------+---------   -3441

       GAAATTCATGCCCACCATGGTAGATAAAAATATTTTTTCTCTCTCCTCTTTTTATTCAA
-3440  +---------+---------+---------+---------+---------+---------   -3381

       ATCTCAAAGCAtAAkrArTGGTGACAGAACGATAAGATTCCTACCTAGCTTTCTGAGATC
-3380  +---------+---------+---------+---------+---------+---------   -3321

       CCACTAGTTTATCTTCAAGCTGGTGATTGAAGGATTAACCATGCTTGAATTAGATTGGCT
-3320  +---------+---------+---------+---------+---------+---------   -3261

       TCAAACTTGGTAGTAGCTTGTTTCATACTTTGATTACTTTGGTATGGTTAGTTGGTTTGA
-3260  +---------+---------+---------+---------+---------+---------   -3201

       GATTTTGGTCAATGTAGAATCAGATTTGAGAGCGATTGTCAGCTTGAATTGCCGCAGTTT
-3200  +---------+---------+---------+---------+---------+---------   -3141

       TAGCACATACTAGTTTGGATAGATGAACAGTTTGGAGAGACAAATAATGTCTATACGAGC
-3140  +---------+---------+---------+---------+---------+---------   -3081

       TCATCGGATAATATTAGTCTATGGCTTTTGCTTCGGTGTCCCCTCTGCAAACTTTACCCC
-3080  +---------+---------+---------+---------+---------+---------   -3021

       TCTGTAGATGGTAGGATTTTCTGATATCCTTTCATGGTTTAAGGGTGTGCGTGTAAGGAA
-3020  +---------+---------+---------+---------+---------+---------   -2961

       CGGGAGATACCGGATCACACCTTTTCGTCTACACTTTACAAGCATGTAACACCTAAGATT
-2960  +---------+---------+---------+---------+---------+---------   -2901

       GATTGATATCTAGGCTTACACCCCAATGGAGGTAAACTAATATTATTGAAATGCGACTTT
-2900  +---------+---------+---------+---------+---------+---------   -2841

       TCAAAAGTCCCAATATAACCTTGACGATGATCTTACAACTACTCGCGCCAGTCTTGTATG
-2840  +---------+---------+---------+---------+---------+---------   -2781

                                 KpnI
       ATATCAGATTGGCCGAGGATCGTGGGTACCTTGTAGTGGACTATGATGCTCATGGAGGTT
-2780  +---------+---------+---------+---------+---------+---------   -2721

       GTATGGACATGTTGTAATGCTGGTTTTCTCTAGGTTTTTTCTAATCAACTTGGCATTCTT
-2720  +---------+---------+---------+---------+---------+---------   -2661

       CTCCTTAACACATAATAAGAGGGAATACCTCCATACATTATTCTGAAAAAAGCATGGCCA
-2660  +---------+---------+---------+---------+---------+---------   -2601

       ACAATGAAACAGAAACAAGTACGACAGTCTATACCCGACCCAAACAATGGCTCAGGTCTT
-2600  +---------+---------+---------+---------+---------+---------   -2541

       TCACGATGCATAGTTTGTTAGCATGTATTTTATAGTAGGAACTAAAATTTAAAGACAACT
-2540  +---------+---------+---------+---------+---------+---------   -2481

       TGCnAAAACAATTTTGTCTCTTGAGTGTTTTTTAAGGATGCGGCATTTATCGATTATACA
-2480  +---------+---------+---------+---------+---------+---------   -2421

       TTACATATGTGATTGGATtAGCCAACTTTTTGTCTTCCgATGATCATATGAAAGGGTTGT
-2420  +---------+---------+---------+---------+---------+---------   -2361

       ATCTTAGGGCATCTCCAATGGGnAGACTCAAATGCAAAAAAATnGTCCGTTTGGGTCTTC
-2360  +---------+---------+---------+---------+---------+---------   -2301

       CnGGACAAAACCTGCTCCCAACGGGGCAACCCAACTTAAAAACGGACAGGTGCAGCGTCC
-2300  +---------+---------+---------+---------+---------+---------   -2241
```

FIGURE 18 CONTINUED

```
        GGCnTGACCCAAAACTGACGCAAATTTGGnAnATTTTTGGGGCnAGCCAGACGAACGCGG
-2240   +---------+---------+---------+---------+---------+---------   -2181

        GCGTCCACTGTATCCGACTATGTCCGCATCCTGGCCCATCTGACAGTGACACAAAATACA
-2180   +---------+---------+---------+---------+---------+---------   -2121

        ACCACATGCGCCCCCCACCCTTCTCTCTCCTCCGTTCGCCTTTTCCCATGGAAnCnGTCC
-2120   +---------+---------+---------+---------+---------+---------   -2061

        TCGCTCCTCGCCGGAATTGATCTCGCCTAACCATGCTCCGCCGCCACCcTCGcCTkAAGG
-2060   +---------+---------+---------+---------+---------+---------   -2001

        CCCCAgCCGCCGCTACcTCCTTTTTGTCAGCCCTATTgGAAGTCGCCGgAGTTGAAACGA
-2000   +---------+---------+---------+---------+---------+---------   -1941

        GCGCCGCCAGCCTcGACACCGCCGAGCAAGACGAAGACTGGGCGGAGCTCGCCGAGACGG
-1940   +---------+---------+---------+---------+---------+---------   -1881

        GACGGGGACGGAGCTCGCCATGCGTGCCTCGCAGGGGCGCGATGGGGGCGGAGCTCGCCG
-1880   +---------+---------+---------+---------+---------+---------   -1821

                PstI
        TGGCTGGCTGCAGCACCTCGGGCCGCTGCTAGCCGTGCCACGACGCGAGCATGCGCCTCG
-1820   +---------+---------+---------+---------+---------+---------   -1761

        ACGCCGCCCCGTGCTACCTCGTCGCGCGCCCAGGGCCGCCCCGCCCCTGCCGACCGgCGg
-1760   +---------+---------+---------+---------+---------+---------   -1701

        CGgAgACGCGAcCTTCGCGgACGTGCCCGGCGGCAGAGACGCGTCCTTCGCGACAGCGCC
-1700   +---------+---------+---------+---------+---------+---------   -1641

        CTCCTCGATCTCCGTCGAGCCGCATACGCgGcTAgGAgGGACGCGGGCGTCCCCGGTGTC
-1640   +---------+---------+---------+---------+---------+---------   -1581

        GGCCTCCGTTGTGGCGCATCGCGGGCGCGGCCTCCGTCGAGGcGCATCGCGGGCGTGGCC
-1580   +---------+---------+---------+---------+---------+---------   -1521

        TCGTGGCGCAGCCTGCCCTGATTCGGTCTGAGGCGCGGCGCGGAGCTTCCTCGCGGCGGC
-1520   +---------+---------+---------+---------+---------+---------   -1461

        GCGGGCGGAGCCTCCTCGCTGCGGCGCGACCTGCTCTGCCGCGGTCCGAGACGCGGCGCG
-1460   +---------+---------+---------+---------+---------+---------   -1401

        GGCAGAGCTTCCTCGCGGCGGCTCGGGCGCGGCTTCCTCGCGGCGATGGCGCTTCCAGGC
-1400   +---------+---------+---------+---------+---------+---------   -1341

        TCGCACGCGGCCTCCGGCGTGGCGCAGCGAGAGCGCAGCCTCCGGTGAGTTAGGCACAGG
-1340   +---------+---------+---------+---------+---------+---------   -1281

        CGCGACACGACATCCCCGGCCTCGGCCTCCGGCGTGGCGCAGCGCGAGCGCGACGTAGCC
-1280   +---------+---------+---------+---------+---------+---------   -1221

        TAGGTTGGCAACTAGTaCTACGAGGAAGAAAGAGGAGAAACAATTATTTGGGTCACAGCG
-1220   +---------+---------+---------+---------+---------+---------   -1161

        TTGGGCGTACTGTGCGATCCAAACGGACACCCgGACGCGAaACGATGTCAGCGTGTCCGC
-1160   +---------+---------+---------+---------+---------+---------   -1101

        GTGGcGACCCAAACGACCCGAAACGGACGTCcGTTTGGGTCGGTGCGTTGGAGATGCCCT
-1100   +---------+---------+---------+---------+---------+---------   -1041

        TACTCCCCATCCTCAAATGAGTCTAATTATATATCTTGTTGTAAGTTTTAAAAAAGTTAA
-1040   +---------+---------+---------+---------+---------+---------   -981
```

FIGURE 18 CONTINUED

```
      ACTTTGATCAACATTAGTAATGATAGTAGCAACGAATACAAAATTAAATTGTAAAAATAT
-980  +---------+---------+---------+---------+---------+---------   -921

      ATTATGAAACTTTATTTTAAGATGGATCTAGTTATACTAATTTTCTGCGGATGGAGGAAG
-920  +---------+---------+---------+---------+---------+---------   -861

      TAGCTAAATATTGTTAATTTCTAAATAAAAAATTAAAACTTTAACTTAAAACAAAAGTTA
-860  +---------+---------+---------+---------+---------+---------   -801

                                        Putative Myb Binding domains
      CAAGCATAATTATCTGtGGATGGAGGAAGtAGCTAAGATACACCAATCCTCTCTCTACAT
-800  +---------+---------+---------+---------+---------+---------   -741

      TACCTAGCATGCCACATCAGGAAACTATTTAGGATAAGCTCCAAGGAACCACCCAGAACA
-740  +---------+---------+---------+---------+---------+---------   -681

      ACAATTTACATGGCCTGGCTAACCTAATGACAATTTCCGAGCAACTGGTGGTGGTGGTAC
-680  +---------+---------+---------+---------+---------+---------   -621

      GCGTTCCTTGTTCAATTGTCTCTATTACAAGAGTGGCCCTGTATAGGTAAAAAAAAATAA
-620  +---------+---------+---------+---------+---------+---------   -561

       HindIII                          PstI
      CAAGCTTCCAAGGACGGCCATGTTCCTTGTTCCTGCAGGCTGCACGTACTCACGACGAAG
-560  +---------+---------+---------+---------+---------+---------   -501

      TGTATCTCGTGTTCTGGACATTTGTCTCGCGCATTTTGTAACCATGAAATTAAAAATGTG
-500  +---------+---------+---------+---------+---------+---------   -441

      GTGGCCTGCTATATCTGTATGGGGGTATCATGCACTCCTTCGCAGAGGAATCCAGACGAC
-440  +---------+---------+---------+---------+---------+---------   -381

      GATTTACACGTGTTTCCACCTTAGCTTTTTTTAAGTGTGTGTGTAAGGAACGATCATATA
-380  +---------+---------+---------+---------+---------+---------   -321

                                                   XhoI
      ACTGCCCCTGAATGCTGCATATATATAAACCGACTCCATCATGTACTCGAGACAAGGTCG
-320  +---------+---------+---------+---------+---------+---------   -261

      TCAAGAAAAACAAACTATGCCTATCTCACTAGCAATGATTTGAGAGTACAGCTTTTCCGG
-260  +---------+---------+---------+---------+---------+---------   -201

      TGCCATATTTTTTCCTATATATCTTTTTCTGAAGAACAAGAAAAAAAAAACAGTTGGTGT
-200  +---------+---------+---------+---------+---------+---------   -141

      GGTGGTTGGTGAAGCGAGAAAGCCCCATATAAGCCCTGCTCACCCTCCCCGCAAAGCACA
-140  +---------+---------+---------+---------+---------+---------   -81

                                                                PvuI
      ACTCATAGCTCGGGTCTCTCGCTCACACCAAAATCGCCCACCAGCACCAGCATCTCTCGA
-80   +---------+---------+---------+---------+---------+---------   -21

      TCGGCAGACGCATAGATCGATGGGCTCCACCGCCGCCGACATGGCCGCGTCCGCGGACGA
-20   +---------+---------+---------+---------+---------+---------   39
                        M  G  S  T  A  A  D  M  A  A  S  A  D  E

      GGACGCGTGCATGTTCGCCCTCCAGCTCGCTTCCTCGTCGGTCCTCCCGATGACGCTGAA
 40   +---------+---------+---------+---------+---------+---------   99
       D  A  C  M  F  A  L  Q  L  A  S  S  S  V  L  P  M  T  L  K

      GAACGCCATCGAGCTTGGCCTCCTGGAGATCCTGGTGGCCGCCGGCGGCAAGTCGCTGAC
100   +---------+---------+---------+---------+---------+---------   159
       N  A  I  E  L  G  L  L  E  I  L  V  A  A  G  G  K  S  L  T

      CCCGACCGAGGTGGCCGCCAAGCTCCCGTCCGCGGCGAACCCGGAAGCGCCGGACATGGT
160   +---------+---------+---------+---------+---------+---------   219
       P  T  E  V  A  A  K  L  P  S  A  A  N  P  E  A  P  D  M  V
```

FIGURE 18 CONTINUED

```
     GGACCGCATACTCCGGCTGCTCGCGTCGTACAACGTCGTGACGTGCCTGGTGGAGGAGGG
220  +---------+---------+---------+---------+---------+---------   279
      D  R  I  L  R  L  L  A  S  Y  N  V  V  T  C  L  V  E  E  G

     CAAGGACGGCCGCCTCTCCCGGAGCTACGGCGCCGCGCCCGTGTGCAAGTTCCTCACCCC
280  +---------+---------+---------+---------+---------+---------   339
      K  D  G  R  L  S  R  S  Y  G  A  A  P  V  C  K  F  L  T  P

     CAACGAGGACGGCGTCTCCATGGCGGCGCTCGCGCTCATGAACCAGGACAAGGTCCTCAT
340  +---------+---------+---------+---------+---------+---------   399
      N  E  D  G  V  S  M  A  A  L  A  L  M  N  Q  D  K  V  L  M

            Intron/exon boundary
     GGAGAGCTG↓GTGAGTCTCTCAGTGGAGCTAGTTACTGTAGATCCGAATTCGTTCCCTTTA
400  +---------+---------+---------+---------+---------+---------   459
      E  S
                                 SalI          pBluescript
     GTGAGGGTTAATTCCGCGGCCGCGTCGACCTCGAGGGGGGGCCCGGTACCCAATTCGCCC
460  +---------+---------+---------+

     TATAGTGAGTCGTATTACGCGCGCTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAA
     AACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGT
     AATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAA
     TGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTTGTG    744
```

FIGURE 18 CONTINUED

pLpOMTGUS

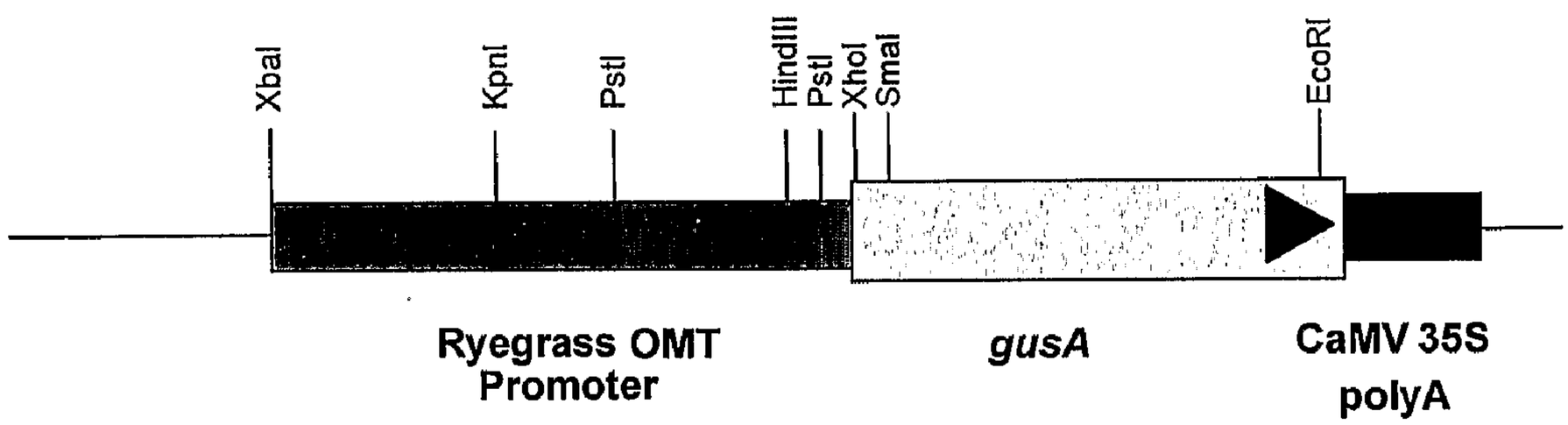
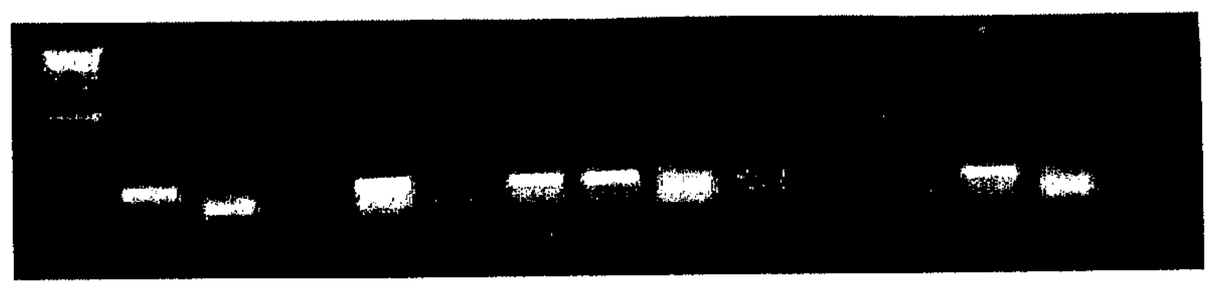
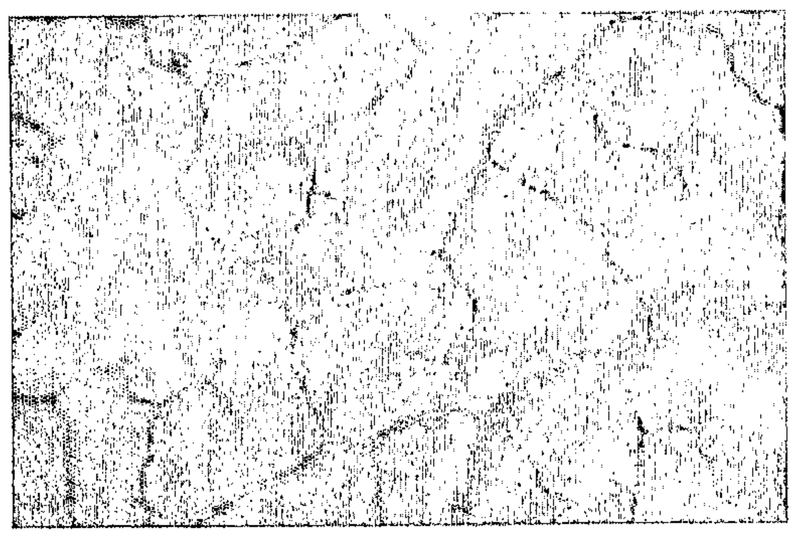


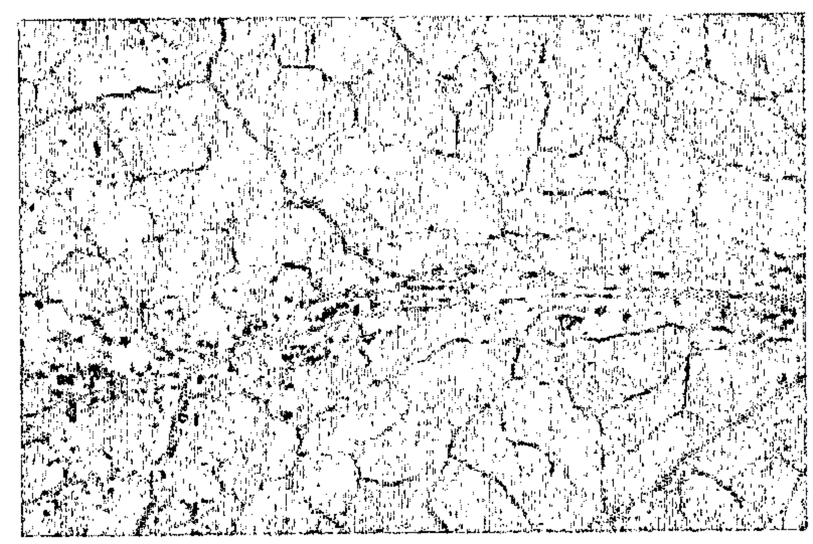
FIGURE 19



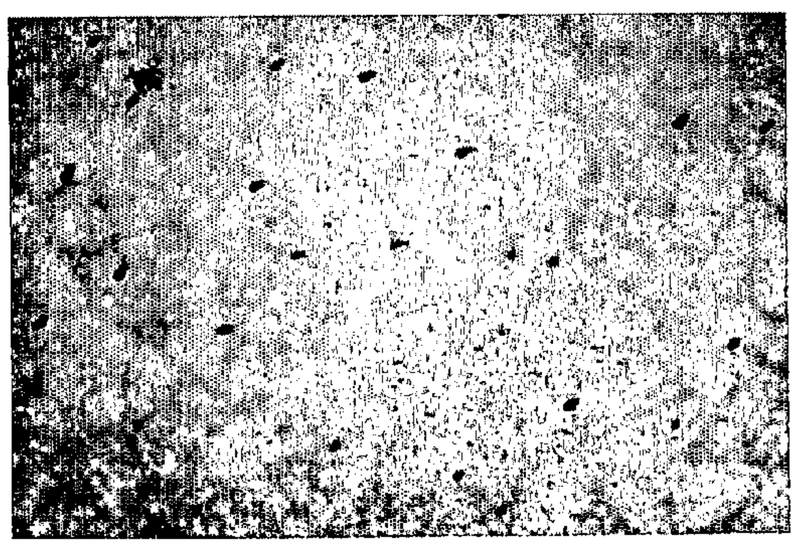
A



B



C



D

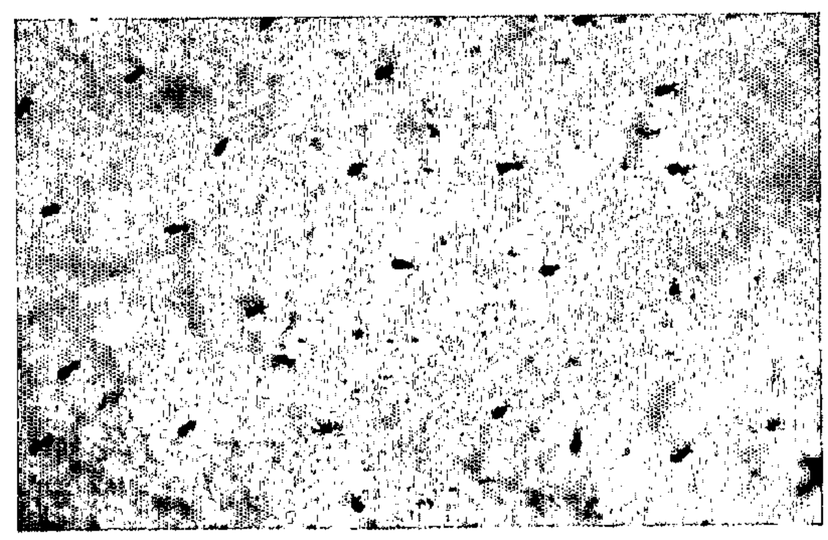
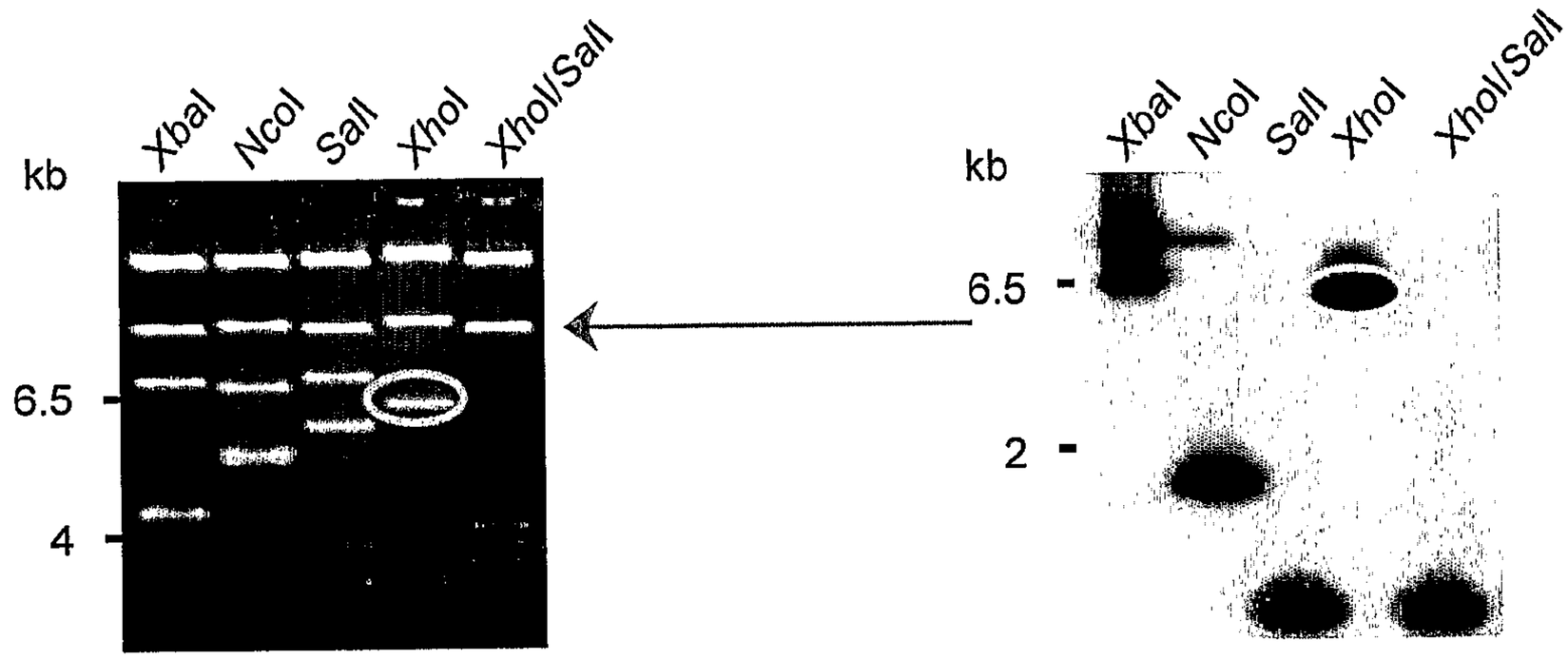
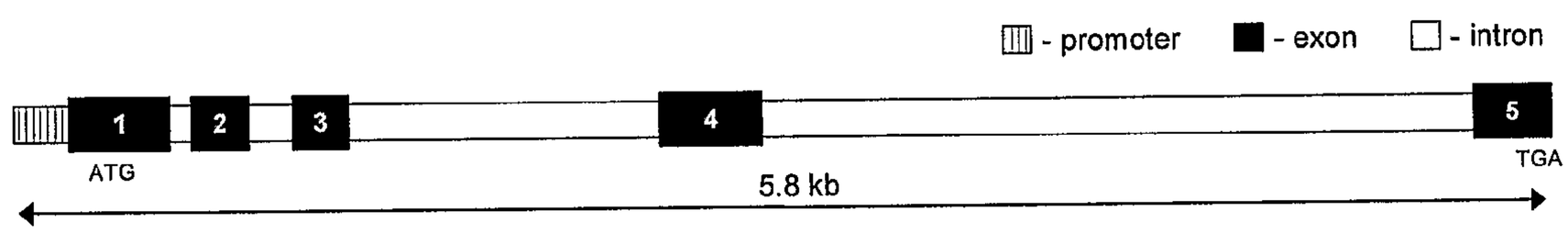


FIGURE 20

A



B

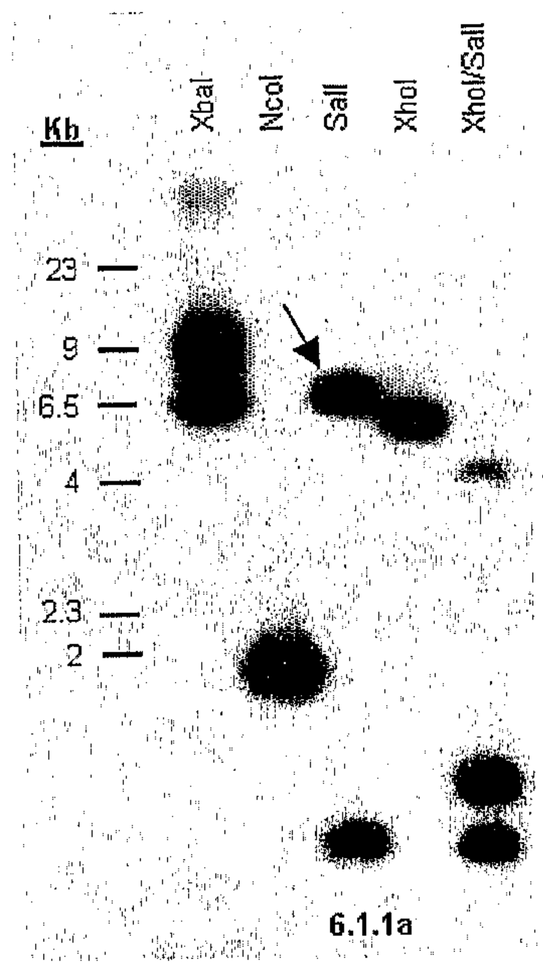


C

Exon	*Lp* CCR1	*Eg* CCR1	*Es* CCR1	*Pb* CCR1
1	173 bp	133 bp	133 bp	139 bp
2	155 bp	155 bp	155 bp	155 bp
3	189 bp	186 bp	186 bp	186 bp
4	353 bp	353 bp	353 bp	353 bp
5	220 bp	218 bp	184 bp	184 bp

FIGURE 21

A


Kb
23
9
6.5
4
2.3
2
XbaI
NcoI
SalI
XhoI
XhoI/SalI
6.1.1a

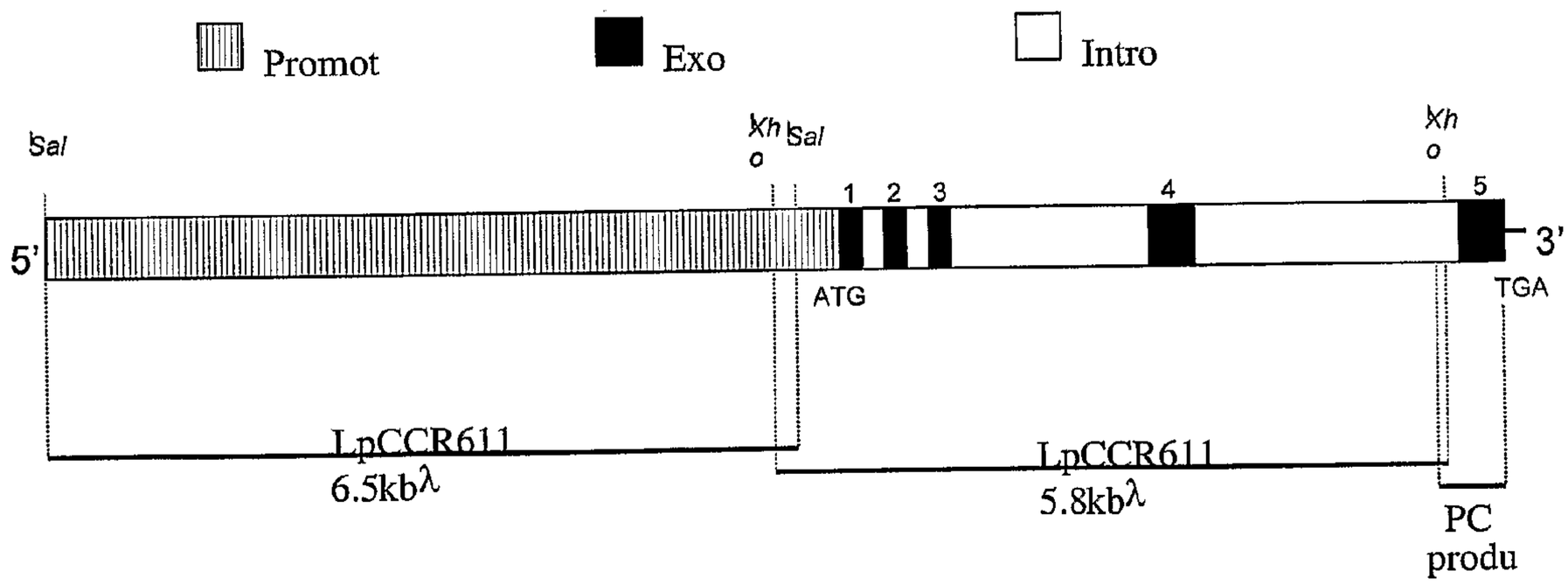

Promot
Exo
Intro
SalI
XhoI
SalI
XhoI
1
2
3
4
5
5'
3'
ATG
TGA
LpCCR611
6.5kbλ
LpCCR611
5.8kbλ
PC
produ

FIGURE 22

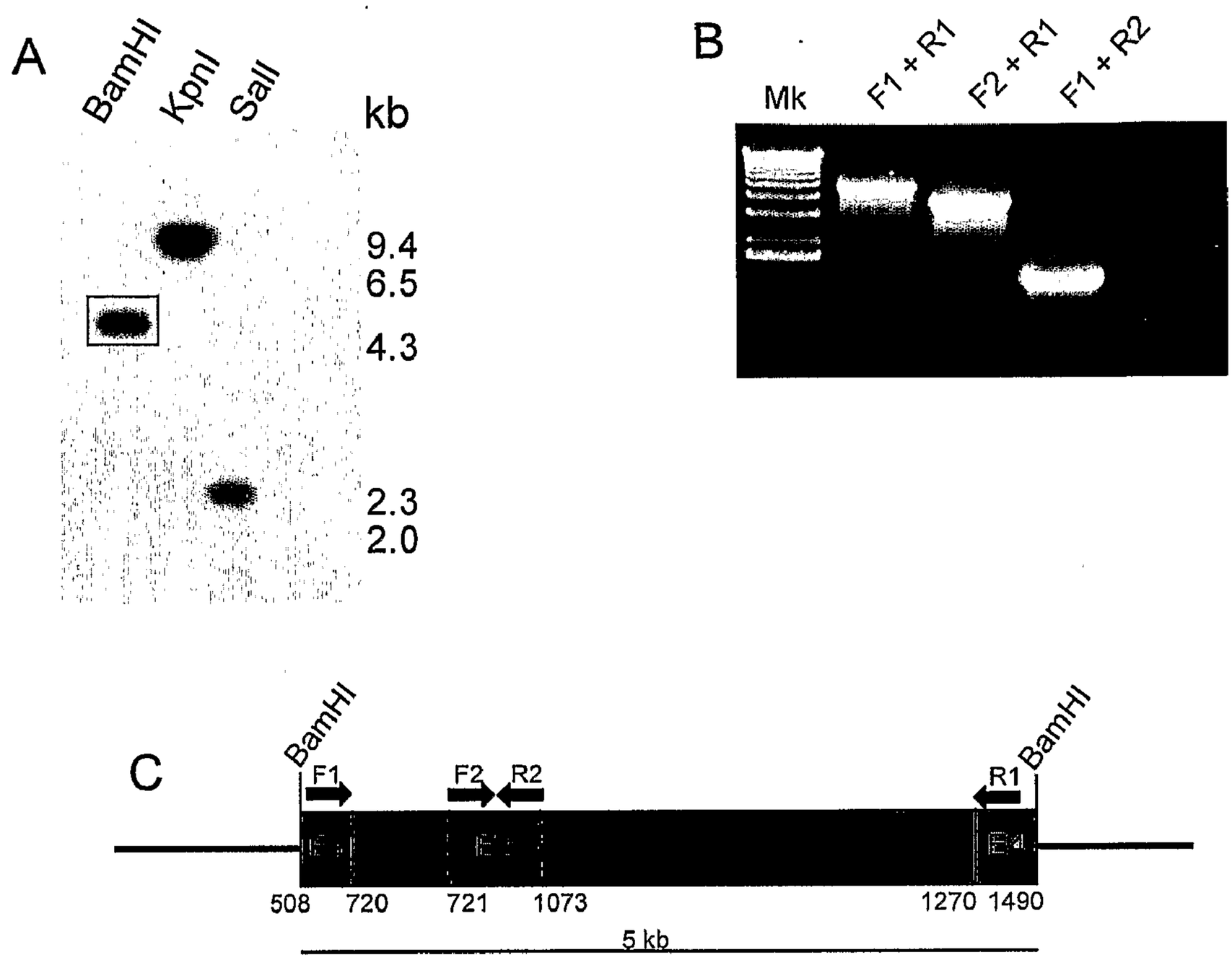

A
BamHI
KpnI
SalI
kb
9.4
6.5
4.3
2.3
2.0
B
Mk
F1 + R1
F2 + R1
F1 + R2
C
BamHI
F1
F2
R2
R1
BamHI
508
720
721
1073
1270
1490
5 kb

FIGURE 23

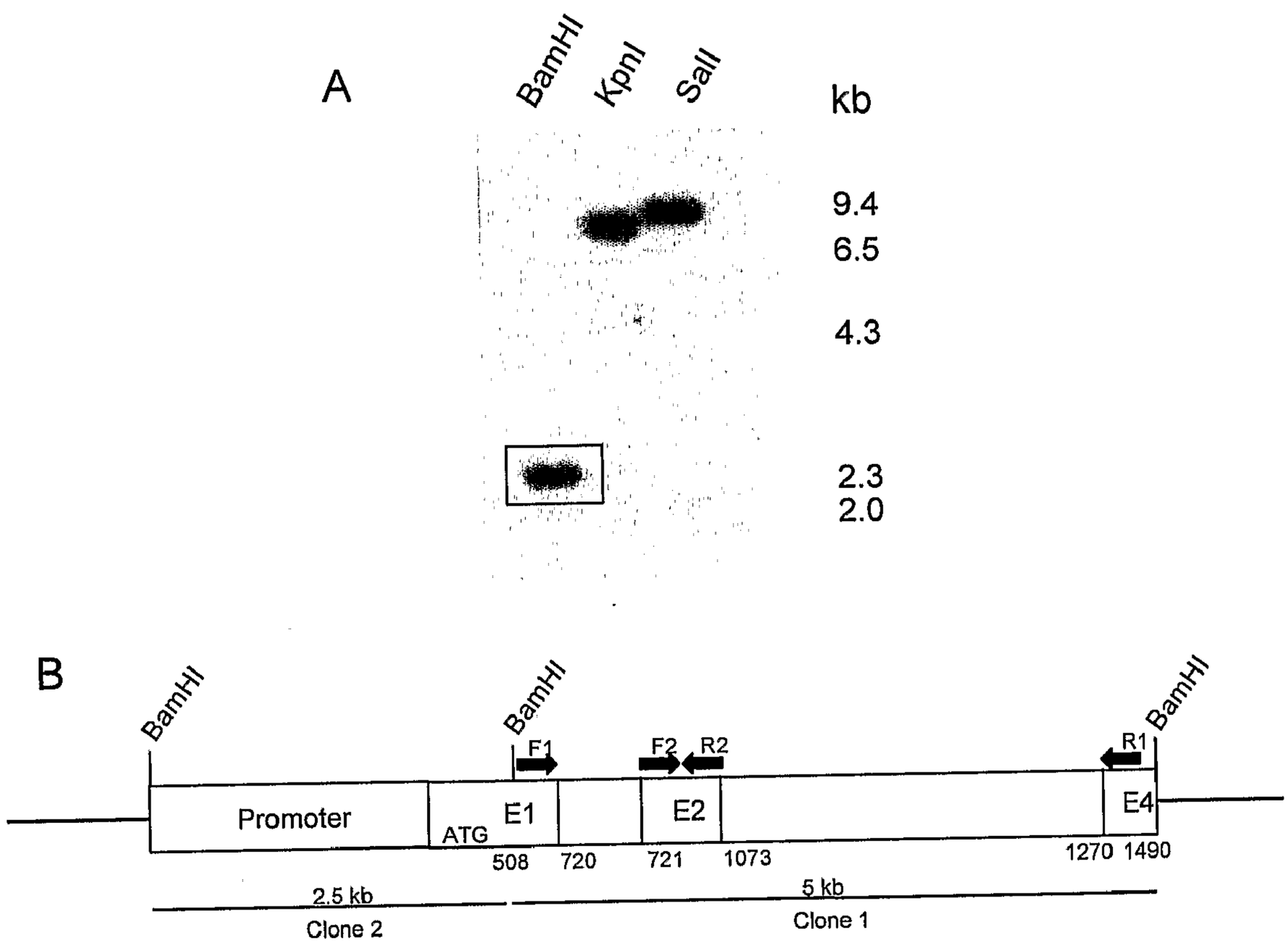

A
BamHI
KpnI
SalI
kb
9.4
6.5
4.3
2.3
2.0
B
BamHI
BamHI
BamHI
F1
F2
R2
R1
Promoter
ATG
E1
E2
E4
508
720
721
1073
1270
1490
2.5 kb
Clone 2
5 kb
Clone 1

FIGURE 24

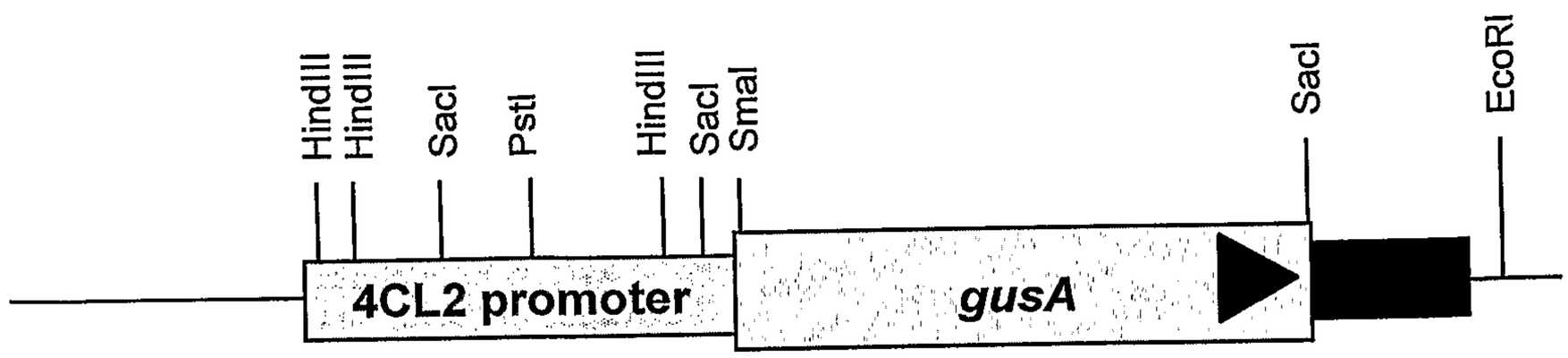

HindIII
HindIII
SacI
PstI
HindIII
SacI
SmaI
SacI
EcoRI
4CL2 promoter
gusA

FIGURE 25

```
     TCCCGTATCTTCAACGTGACACCCCTACACTTCCTGCTTGTCTTGGAGATTTACACACAC
  1  ---------+---------+---------+---------+---------+---------+   60

     ACGGCAATTACCAGGAGTATCTTCCTAGATTATTTTTTTCGATAAGGATCTTCCAGATAT
 61  ---------+---------+---------+---------+---------+---------+  120

     AGCATGTGAATCTCTGTACTACTACTGTTTGTCAAGCAAAATTAACATTGACATCAGTGT
121  ---------+---------+---------+---------+---------+---------+  180

     TTTTGTTGGGGGCAGCGGAATCTTTGACGCCTCTTCTTGCCTCTCAAGACATGTCACCCT
181  ---------+---------+---------+---------+---------+---------+  240

     CACTAGTTAGTGTGCCAGCTGGTAGTACTACGTACGATGCTCCCTCCCTCCGTAATTATT
241  ---------+---------+---------+---------+---------+---------+  300

     CAACCTTTTTGCTCTCTCTTTTTATAAAGTCAAACCTTTTAAATCTGACCAGATATCTGC
301  ---------+---------+---------+---------+---------+---------+  360

     TAAAAAATTAGCAGACATGCATACATCAAAGCAGTAGTCCTCCCTCCGTTTAAAATTACC
361  ---------+---------+---------+---------+---------+---------+  420

     TGGGTTTATTCAAATAAAGTCAAACTCTGTAAAATTCAATTAAATATTTAGAAAAATCTA
421  ---------+---------+---------+---------+---------+---------+  480

     ACAGCACCTGTAGTATAAAAGTATGCTCCCTCTGTTTGTAAAAAAGCTAAGCAACTTTTT
481  ---------+---------+---------+---------+---------+---------+  540

     TGAGATACGGATAAATCTTTAGCTAAAACATGTCTATATACCTTTGTATCTAGATAAAGT
541  ---------+---------+---------+---------+---------+---------+  600

     TGGAAAGCTTTTTTAGAAACAGACAAAGTATGTGTTTGACATTATGAATGTTGAGTATTT
601  ---------+---------+---------+---------+---------+---------+  660

     TTCCTCTAATCTTGATCAAATTTTACAAATTTTGGCTTGAATAGAGGGACCATTATTAGT
661  ---------+---------+---------+---------+---------+---------+  720

     ATGAAACTACATAAATTTGTAAAACACTCAACATAATTTACGATGGGTCAGTGATAGCAC
721  ---------+---------+---------+---------+---------+---------+  780

     TAACTTAGCTTTTCATAAATGCCACTGCTTTTCAATAGAGCATGAAGCAGGACAAATTTA
781  ---------+---------+---------+---------+---------+---------+  840

     TTCGTGTGACTTGAATAGAGGGAGCCTGTTCTGGTTCAACTCACCCTGCATGTGTGTCTT
841  ---------+---------+---------+---------+---------+---------+  900

     CATCCCTTTTGCTCTTCCTATCTGTGGTGTCAATTGAGTGTCCCACGTGCATGTGGGCGA
901  ---------+---------+---------+---------+---------+---------+  960
```

FIGURE 26

```
      AACTTGAACCTAGAAATTGACATGCTCCCACTGCCCGGAGCGGAGTATCTTTGTGCTTTG
 961  ---------+---------+---------+---------+---------+---------+   1020

      TTACCCTTATTGTTGCTACGTACTACAGTGTTTAGATTGGAACTTCATAATCAAAAGAAC
1021  ---------+---------+---------+---------+---------+---------+   1080

      TTAGTTTCCTACAATTTTTTGCTAAGCAATATAATGAGCAATCAAACTTCTATATCTGTG
1081  ---------+---------+---------+---------+---------+---------+   1140

      GCAAATAACTAATCCATTATAGTTACAGTTTAGATGCAGACGCCAGTGTTTCTTCCCCTT
1141  ---------+---------+---------+---------+---------+---------+   1200

      TTCGGAAAAAAGCTATTCCATAATAAGTGTTGGAAATTTAATAAATGGGTACTACGAATT
1201  ---------+---------+---------+---------+---------+---------+   1260

      TGAAAAAAAAAGTGTCAAAAATTCACTAAGAAAGTACGTAGTACAAATTTAAACTAAGAT
1261  ---------+---------+---------+---------+---------+---------+   1320

      TCCGACACTTATTAGGATCGGAGAGAGTAAGTAGCAAACTACTACTCCATCCACCTAAAA
1321  ---------+---------+---------+---------+---------+---------+   1380

      CACGTGATTTAACTTTGTCTAGATACGGATAGAAAGTTGGGATACATCCGTATCTTAAAA
1381  ---------+---------+---------+---------+---------+---------+   1440

      AAAAACGCACTTATTTTAGACGAAGGAGGGAGTATTTCAACCTTGATTTTAAACGGAATC
1441  ---------+---------+---------+---------+---------+---------+   1500

      TACAAAGGGAATACATGGATTGTACAAGTGGGCTGACCGTATCCATTATGTACTCGTACT
1501  ---------+---------+---------+---------+---------+---------+   1560

      TTGCAGTTTGAAAGCAAAGGCTAGTGTAATTTGTAGGTGGTTCTAGGCGTCTAGCTGTTT
1561  ---------+---------+---------+---------+---------+---------+   1620

      CATGGCGTTATCACAGCCGTGCCAGTGTGCTCAGGGCCGTACATAAGTTGCTTGGTGTAT
1621  ---------+---------+---------+---------+---------+---------+   1680

      GTGTCGATCTAGGATTTGCCGTCTTACAATTTTGCTTTCCAACTTATTTTCTGTAAAGAG
1681  ---------+---------+---------+---------+---------+---------+   1740

      ATCGATGTGAACTTCTCTGTCGAGTAAACTGAAATTGTCTGAATAAATATAACTCGGCAG
1741  ---------+---------+---------+---------+---------+---------+   1800

      ATTATGTTTTATCGTTTGCATGCGTAACAGGCTACACAAATTGCTCGAGTCAGCAGCGAG
1801  ---------+---------+---------+---------+---------+---------+   1860

      TTGAGCTCACAACGAATCCATCAGCAAAAATACTATACTATAGTAGCACATCGTTTCTTT
1861  ---------+---------+---------+---------+---------+---------+   1920
```

FIGURE 26 CONTINUED

```
     TTTCATGACGTTTCTGTTTCTTCCTAACTTTCCAGGAGCACCGGAGACGACGATGTGGTG
1921 ---------+---------+---------+---------+---------+---------+  1980
                                   R  S  T  G  D  D  D  V  V

     ATAAAGATTTTGTACTGCGGAATCTGCCACTCTGACCTGCACGCCCTGAAGAACGACTGG
1981 ---------+---------+---------+---------+---------+---------+  2040
     I  K  I  L  Y  C  G  I  C  H  S  D  L  H  A  L  K  N  D  W

     AAGAACTCAAGGTACCCGATGATCCCCGGGCACGAGATCGCCGGCGAGGTCACGGAGGTG
2041 ---------+---------+---------+---------+---------+---------+  2100
     K  N  S  R  Y  P  M  I  P  G  H  E  I  A  G  E  V  T  E  V

     GGCAAGAACGTGAGCAAGTTCAAGGCCGGCGACCGCGTGGGCGTCGGGTGCATGGTGAAC
2101 ---------+---------+---------+---------+---------+---------+  2160
     G  K  N  V  S  K  F  K  A  G  D  R  V  G  V  G  C  M  V  N

     TCGTGCCGGTCGTGCGAGAGCTGCGACAAGGGCTTCGAGAACCACTGCCCGGGCATGATC
2161 ---------+---------+---------+---------+---------+---------+  2220
     S  C  R  S  C  E  S  C  D  K  G  F  E  N  H  C  P  G  M  I

     CTCACCTACAACTCGGTCGACGTCGACGGCACCGTCACCTACGGCGGCTACTCCAGCATG
2221 ---------+---------+---------+---------+---------+---------+  2280
     L  T  Y  N  S  V  D  V  D  G  T  V  T  Y  G  G  Y  S  S  M

     GTGGTGGTGCACGAGCGGTTCGTGGTCCGGTTCCCCGACGCCATGCCGCTGGACAAGGGC
2281 ---------+---------+---------+---------+---------+---------+  2340
     V  V  V  H  E  R  F  V  V  R  F  P  D  A  M  P  L  D  K  G

     GCGCCGCTGCTGTGCGCCGGCATCACCGTGTACAGCCCCATGAAGTACCACGGGCTCAAC
2341 ---------+---------+---------+---------+---------+---------+  2400
     A  P  L  L  C  A  G  I  T  V  Y  S  P  M  K  Y  H  G  L  N

     GTTCCCGGGCTGCACCTCGGCGTGCTGGGGCTGGGCGGGCTGGGCCACGTTGCGGTCAAG
2401 ---------+---------+---------+---------+---------+---------+  2460
     V  P  G  L  H  L  G  V  L  G  L  G  G  L  G  H  V  A  V  K

     TTCGGCAAGGCCTTCGGAATGAAAGTGACGGTGATCAGCTCGTCGCCGGGGAAGAAGGAG
2461 ---------+---------+---------+---------+---------+---------+  2520
     F  G  K  A  F  G  M  K  V  T  V  I  S  S  S  P  G  K  K  E

     GAGGCCCTGGGGCGGCTGGGCGCCGACGCGTTCATCGTCAGCAAGGACGCCGACGAGATG
2521 ---------+---------+---------+---------+---------+---------+  2580
     E  A  L  G  R  L  G  A  D  A  F  I  V  S  K  D  A  D  E  M

     AAGGTAGGCGGACCCGCTGGTTCAGGTTACTTCCCCTGTCCGGTGCAGAAGAAAGAGGAA
2581 ---------+---------+---------+---------+---------+---------+  2640

     K
```

FIGURE 26 CONTINUED

G at 851 bp (coding sequence) missing from cDNA in cv Ellett

```
     CTTGAGGGTTCATGTTTGTTTTGCGTTGGTGATGTCTTTGCAGGCTGTGATGAGCACCAT
2641 ---------+---------+---------+---------+---------+---------+   2700
                                                A  V  M  S  T  M

     GGATGGCATCATAAACACGGTATCTGCAAACATCCCCCTGACCCCTCTCTTCGGGCTGCT
2701 ---------+---------+---------+---------+---------+---------+   2760
      D  G  I  I  N  T  V  S  A  N  I  P  L  T  P  L  F  G  L  L

     CAAGCCCAACGGCAAGATGATCATGGTCGGCCTCCCCGAGAAGCCCATCGAGATTCCTCC
2761 ---------+---------+---------+---------+---------+---------+   2820
      K  P  N  G  K  M  I  M  V  G  L  P  E  K  P  I  E  I  P  P

     CTTCGCTCTAGTTGCCAGTAAGTCTTAGGATCTCTTGCAATAAGGAGAAATCATGCACTG
2821 ---------+---------+---------+---------+---------+---------+   2880
      F  A  L  V  A

     ATCGATCAGAGAAATGAGATAGCATCCTGATGAACATTGTACGTGTGTGCAGCGAATAAG
2881 ---------+---------+---------+---------+---------+---------+   2940
                                                          N  K

     ACCCTGGCCGGGAGCATCATCGGCGGCATGAGCGACACGCAGGAGATGCTGGACCTCGCG
2941 ---------+---------+---------+---------+---------+---------+   3000
     T  L  A  G  S  I  I  G  G  M  S  D  T  Q  E  M  L  D  L  A

     GCGAAGCACGGCGTGACGGCCGACATCGAGGTGGTCGGCGCGGAGTATGTGAACACGGCC
3001 ---------+---------+---------+---------+---------+---------+   3060
     A  K  H  G  V  T  A  D  I  E  V  V  G  A  E  Y  V  N  T  A

     TTGGAGCGCCTTGCCAAGAACGACGTCAGGTATCGCTTCGTCATCGACATCGGCAACACC
3061 ---------+---------+---------+---------+---------+---------+   3120
     L  E  R  L  A  K  N  D  V  R  Y  R  F  V  I  D  I  G  N  T

     CTCGACAAGGTTGCGGCCACCACCGAGTGAACGTACTCAGCACTGCTTACGATCTACGTT
3121 ---------+---------+---------+---------+---------+---------+   3180
     L  D  K  V  A  A  T  T  E  *

     GTTCCACTGTTAGTGCTCCGTAGTAAACAATAAACGATCAAAACTCTTGTCATCTGGTGC
3181 ---------+---------+---------+---------+---------+---------+   3240

     ATTGGTGTAGACATGGTTGTTTGCGAGGAAACTGAGTTGAAGGATGGATGGATAAGTTTG
3241 ---------+---------+---------+---------+---------+---------+   3300

     CTTCTTGCCGTGTTAATGGATTACCTACTTAGCTTCACTGCAATTAACAAATTAAGAAAC
3301 ---------+---------+---------+---------+---------+---------+   3360

     GACACACCCAAAAGACTTTCGTCAGTTTTCTTGGATTATACAAGTCGTTATGGTTGGGTG
3361 ---------+---------+---------+---------+---------+---------+   3420
```

FIGURE 26 CONTINUED

```
     TCAGTGTGTCACAGATAATCATACTATGGTATTTAACCTGGAAGATCGTTTTTTTGGCGG
3421 ---------+---------+---------+---------+---------+---------+  3480

     CAACTCAGTGGGTTTTCCCACTATGTATATTTATAAATATTCAACAAGTCATGAGGTACA
3481 ---------+---------+---------+---------+---------+---------+  3540

     AAGGGTTGTTGCTAGAGGATAGCAACAAGAAGCTAGCCAAAAGATCATAGGCTTAAAAAA
3541 ---------+---------+---------+---------+---------+---------+  3600

     GAGAGAAAAGAAAACAAAACTGCTATAGTTATCGAAATCTCTCAGCTCAAATTTTAAAAC
3601 ---------+---------+---------+---------+---------+---------+  3660

     CAGCATAAGACTTTCTAGAAGCCTTATGAACAAGAAGAGCTAGCTCATCTTTAAACCTTT
3661 ---------+---------+---------+---------+---------+---------+  3720

     TCCTGCATCTGTAAAGATTGAGGGTGCAACCCTTGAATATAAAATCATTCCTGTCATCCA
3721 ---------+---------+---------+---------+---------+---------+  3780

     GATAGACTATGTAGTCAAAATAGTCATTTCCATGAAGAAGGGCACTTTTAATACATTTTT
3781 ---------+---------+---------+---------+---------+---------+  3840

     GAGACTTGGTATGATACTCTGAATGTCAACACCCTGGAAGATCTTTTCACTCCTATGGAA
3841 ---------+---------+---------+---------+---------+---------+  3900

     GGACAAGAAAGCATTTCAACTCCTTTTACTAAGGAAGAGATTGACAAGGTGATTCAGAGA
3901 ---------+---------+---------+---------+---------+---------+  3960

     ATTCCTTTAGACACTATAGAAAGTCACAAGGTGCCAACGGCGCAATCCTGTGCCGACGGC
3961 ---------+---------+---------+---------+---------+---------+  4020

     TTTTTATCGGGGAAGCCAGCATCGGTACCGAGACCGGCAGCCCACCAACTAGGCCGTCGG
4021 ---------+---------+---------+---------+---------+---------+  4080

     CACACATCCTCCAGTGTCGGCGGCCAACATCGGCATAAGTTGGCCCGTTGGGCATCAACT
4081 ---------+---------+---------+---------+---------+---------+  4140

     CCCCCGTCGGAACAGGTCTAGCGCATGGACCGTCGTGATGGCGGCGGCAACGACGTCATC
4141 ---------+---------+---------+---------+---------+---------+  4200

     CTATGCCGACGGCCTAGCCGTCGGCCTAGCTTGCCAGCGCTATGCCGACGTCACATTGCC
4201 ---------+---------+---------+---------+---------+---------+  4260

     ATCGGCACATGCTAGTTTTTTTTTCTTTTTTCTACATGCCAAATTGTATATGTATATATA
4261 ---------+---------+---------+---------+---------+---------+  4320

     CTCATTTACTTATTACTTCCAATTATTTTAATGTGTATATATTTTGCTCACCAATTGTAC
4321 ---------+---------+---------+---------+---------+---------+  4380
```

FIGURE 26 CONTINUED

```
     GAATTTGTACCCTCCGAGAAATTGCTAAAATGATGGAGTGACCTACAACGAGCCTTGGAT
4381 ---------+---------+---------+---------+---------+---------+   4440

     ATGTGAGTTCTTCTTGCCCCATTGCACAAAAATTGTAAATATTAGGGTTTACTGGATCCA
4441 ---------+---------+---------+---------+---------+---------+   4500

     CTAGTTCTAGAGCGGCCGCCACCGCGGGGAGCTCCAGCTTTTGTTCCCTTTAGTA
A)   ---------+---------+---------+---------+---------+-----  4555
```

FIGURE 26 CONTINUED

```
     GGCACGAGTCGCCTCCAACGTCTTCCCTTAACCGGCCGTCCCTACGCtTGCACCACCACC
  1  ---------+---------+---------+---------+---------+---------+   60

     ACGCACAGACAGAGCAGTTTCCCAGCCCCCGCCGGAACCGGATGGCACCCACGGCGGCGG
 61  ---------+---------+---------+---------+---------+---------+  120
                                                M  A  P  T  A  A  E

     AGCAGACGGAGCACCACCAGCACACCAGGAAGGCGGTGGGGCTGGCGGCGCGCGACGACG
121  ---------+---------+---------+---------+---------+---------+  180
       Q  T  E  H  H  Q  H  T  R  K  A  V  G  L  A  A  R  D  D  A

     CCGGCCACCTCTCCCCGCTCGCCATCACACGGAGGAGCACAGGAGACGACGATGTGGTGA
181  ---------+---------+---------+---------+---------+---------+  240
       G  H  L  S  P  L  A  I  T  R  R  S  T  G  D  D  D  V  V  I

     TAAAGATTTTGTACTGCGGAATCTGCCACTCTGACCTGCACGCCCTGAAGAACGACTGGA
241  ---------+---------+---------+---------+---------+---------+  300
       K  I  L  Y  C  G  I  C  H  S  D  L  H  A  L  K  N  D  W  K

     AGAACTCAAGGTACCCGATGATCCCCGGGCACGAGATCGCCGGCGAGGTCACGGAGGTGG
301  ---------+---------+---------+---------+---------+---------+  360
       N  S  R  Y  P  M  I  P  G  H  E  I  A  G  E  V  T  E  V  G

     GCAAGAACGTGAGCAAGTTCAAGGCCGGCGACCGCGTGGGCGTCGGGTGCATGGTGAACT
361  ---------+---------+---------+---------+---------+---------+  420
       K  N  V  S  K  F  K  A  G  D  R  V  G  V  G  C  M  V  N  S

     CGTGCCGGTCGTGCGAGAGCTGCGACAAGGGCTTCGAGAACCACTGCCCGGGCATGATCC
421  ---------+---------+---------+---------+---------+---------+  480
       C  R  S  C  E  S  C  D  K  G  F  E  N  H  C  P  G  M  I  L

     TCACCTACAACTCGGTCGACGTCGACGGCACCGTCACCTACGGCGGCTACTCCAGCATGG
481  ---------+---------+---------+---------+---------+---------+  540
       T  Y  N  S  V  D  V  D  G  T  V  T  Y  G  G  Y  S  S  M  V

     TGGTGGTGCACGAGCGGTTCGTGGTCCGGTTCCCCGACGCCATGCCGCTGGACAAGGGCG
541  ---------+---------+---------+---------+---------+---------+  600
       V  V  H  E  R  F  V  V  R  F  P  D  A  M  P  L  D  K  G  A

     CGCCGCTGCTGTGCGCCGGCATCACCGTGTACAGCCCCATGAAGTACCACGGGCTCAACG
601  ---------+---------+---------+---------+---------+---------+  660
       P  L  L  C  A  G  I  T  V  Y  S  P  M  K  Y  H  G  L  N  V

     TTCCCGGGCTGCACCTCGGCGTGCTGGGGCTGGGCGGGCTGGGCCACGTTGCGGTCAAGT
661  ---------+---------+---------+---------+---------+---------+  720
       P  G  L  H  L  G  V  L  G  L  G  G  L  G  H  V  A  V  K  F

     TCGGCAAGGCCTTCGGAATGAAAGTGACGGTGATCAGCTCGTCGCCGGGGAAGAAGGAGG
721  ---------+---------+---------+---------+---------+---------+  780
       G  K  A  F  G  M  K  V  T  V  I  S  S  S  P  G  K  K  E  E
```

FIGURE 27

```
     AGGCCCTGGGGCGGCTGGGCGCCGACGCGTTCATCGTCAGCAAGGACGCCGACGAGATGA
781  ---------+---------+---------+---------+---------+---------+   840
       A  L  G  R  L  G  A  D  A  F  I  V  S  K  D  A  D  E  M  K
```

G at 851 bp (coding sequence) missing from cDNA in cv Ellett

```
     AGGCTGTGATGAGCACCATGGATGGCATCATAAACACGGTATCTGCAAACATCCCCCTGA
841  ---------+---------+---------+---------+---------+---------+   900
       A  V  M  S  T  M  D  G  I  I  N  T  V  S  A  N  I  P  L  T

     CCCCTCTCTTCGGGCTGCTCAAGCCCAACGGCAAGATGATCATGGTCGGCCTCCCCGAGA
901  ---------+---------+---------+---------+---------+---------+   960
       P  L  F  G  L  L  K  P  N  G  K  M  I  M  V  G  L  P  E  K

     AGCCCATCGAGATTCCTCCCTTCGCTCTAGTTGCCACGAATAAGACCCTGGCCGGGAGCA
961  ---------+---------+---------+---------+---------+---------+   1020
       P  I  E  I  P  P  F  A  L  V  A  T  N  K  T  L  A  G  S  I

     TCATCGGCGGCATGAGCGACACGCAGGAGATGCTGGACCTCGCGGCGAAGCACGGCGTGA
1021 ---------+---------+---------+---------+---------+---------+   1080
       I  G  G  M  S  D  T  Q  E  M  L  D  L  A  A  K  H  G  V  T

     CGGCCGACATCGAGGTGGTCGGCGCGGAGTATGTGAACACGGCCTTGGAGCGCCTTGCCA
1081 ---------+---------+---------+---------+---------+---------+   1140
       A  D  I  E  V  V  G  A  E  Y  V  N  T  A  L  E  R  L  A  K

     AGAACGACGTCAGGTATCGCTTCGTCATCGACATCGGCAACACCCTCGACAATGTTGCGG
1141 ---------+---------+---------+---------+---------+---------+   1200
       N  D  V  R  Y  R  F  V  I  D  I  G  N  T  L  D  N  V  A  A

     CCACCACCGAGTGAACGTACTCAGCACTGCTTACGATCTACGTTGTTCCACTGTTAGTGC
1201 ---------+---------+---------+---------+---------+---------+   1260
       T  T  E  *

     TCCGTAGTAAACAATAAACGATCAAAACTCTTGTCATCTGGTGCATTGGTGTAGACATGG
1261 ---------+---------+---------+---------+---------+---------+   1320

     TTGTTTGCGAGGAAACTGAGTTGAAGGATGGATGGATAAAAAAAAAAAAAAAAAAAAA
A)   ---------+---------+---------+---------+---------+--------   1378

     GGCACGAGTCGCCTCCAACGTCTTCCCTTAACCGGCCGTCCCTACGCtTGCACCACCACC
1    ---------+---------+---------+---------+---------+---------+   60

     ACGCACAGACAGAGCAGTTTCCCAGCCCCCGCCGGAACCGGATGGCACCCACGGCGGCGG
61   ---------+---------+---------+---------+---------+---------+   120
                                                M  A  P  T  A  A  E

     AGCAGACGGAGCACCACCAGCACACCAGGAAGGCGGTGGGGCTGGCGGCGCGCGACGACG
121  ---------+---------+---------+---------+---------+---------+   180
       Q  T  E  H  H  Q  H  T  R  K  A  V  G  L  A  A  R  D  D  A
```

FIGURE 27 CONTINUED

```
     CCGGCCACCTCTCCCCGCTCGCCATCACACGGAGGAGCACAGGAGACGACGATGTGGTGA
181  ---------+---------+---------+---------+---------+---------+   240
       G  H  L  S  P  L  A  I  T  R  R  S  T  G  D  D  D  V  V  I

     TAAAGATTTTGTACTGCGGAATCTGCCACTCTGACCTGCACGCCCTGAAGAACGACTGGA
241  ---------+---------+---------+---------+---------+---------+   300
       K  I  L  Y  C  G  I  C  H  S  D  L  H  A  L  K  N  D  W  K

     AGAACTCAAGGTACCCGATGATCCCCGGGCACGAGATCGCCGGCGAGGTCACGGAGGTGG
301  ---------+---------+---------+---------+---------+---------+   360
       N  S  R  Y  P  M  I  P  G  H  E  I  A  G  E  V  T  E  V  G

     GCAAGAACGTGAGCAAGTTCAAGGCCGGCGACCGCGTGGGCGTCGGGTGCATGGTGAACT
361  ---------+---------+---------+---------+---------+---------+   420
       K  N  V  S  K  F  K  A  G  D  R  V  G  V  G  C  M  V  N  S

     CGTGCCGGTCGTGCGAGAGCTGCGACAAGGGCTTCGAGAACCACTGCCCGGGCATGATCC
421  ---------+---------+---------+---------+---------+---------+   480
       C  R  S  C  E  S  C  D  K  G  F  E  N  H  C  P  G  M  I  L

     TCACCTACAACTCGGTCGACGTCGACGGCACCGTCACCTACGGCGGCTACTCCAGCATGG
481  ---------+---------+---------+---------+---------+---------+   540
       T  Y  N  S  V  D  V  D  G  T  V  T  Y  G  G  Y  S  S  M  V

     TGGTGGTGCACGAGCGGTTCGTGGTCCGGTTCCCCGACGCCATGCCGCTGGACAAGGGCG
541  ---------+---------+---------+---------+---------+---------+   600
       V  V  H  E  R  F  V  V  R  F  P  D  A  M  P  L  D  K  G  A

     CGCCGCTGCTGTGCGCCGGCATCACCGTGTACAGCCCCATGAAGTACCACGGGCTCAACG
601  ---------+---------+---------+---------+---------+---------+   660
       P  L  L  C  A  G  I  T  V  Y  S  P  M  K  Y  H  G  L  N  V

     TTCCCGGGCTGCACCTCGGCGTGCTGGGGCTGGGCGGGCTGGGCCACGTTGCGGTCAAGT
661  ---------+---------+---------+---------+---------+---------+   720
       P  G  L  H  L  G  V  L  G  L  G  G  L  G  H  V  A  V  K  F

     TCGGCAAGGCCTTCGGAATGAAAGTGACGGTGATCAGCTCGTCGCCGGGGAAGAAGGAGG
721  ---------+---------+---------+---------+---------+---------+   780
       G  K  A  F  G  M  K  V  T  V  I  S  S  S  P  G  K  K  E  E

     AGGCCCTGGGGCGGCTGGGCGCCGACGCGTTCATCGTCAGCAAGGACGCCGACGAGATGA
781  ---------+---------+---------+---------+---------+---------+   840
       A  L  G  R  L  G  A  D  A  F  I  V  S  K  D  A  D  E  M  K
```

G missing at 851 bp in the cDNA isolated from cv Ellett resulted in a premature stop codon (truncated CAD2)

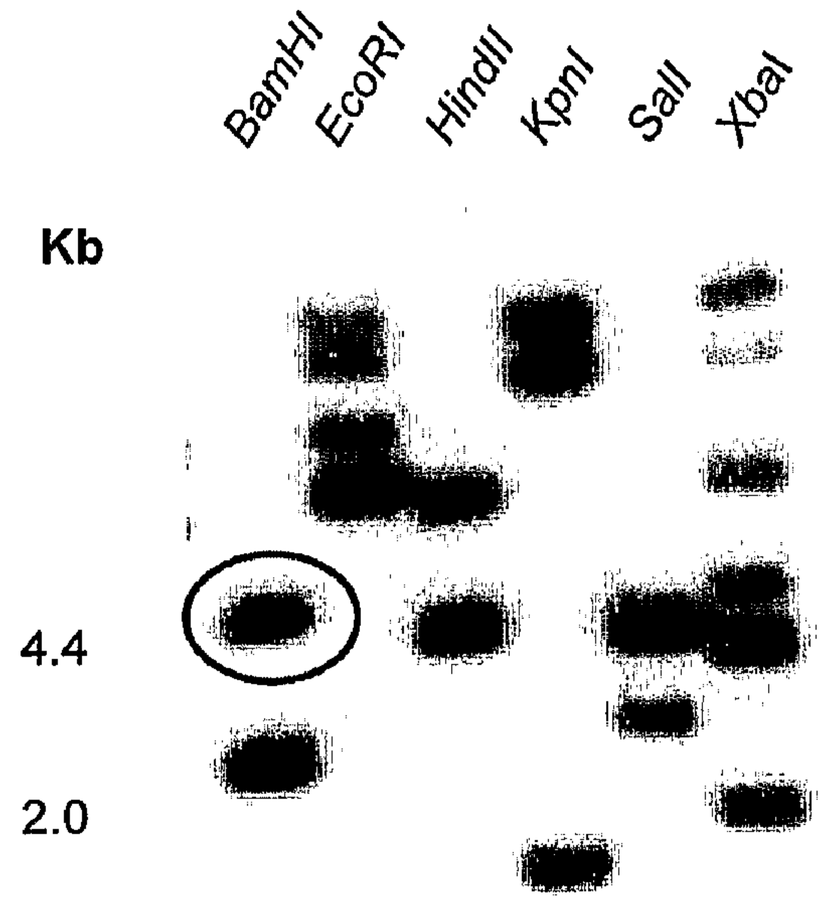
```
     AGGCTGTGATAGCACCATGGATGGCATCATAAACACGGTATCTGCAAACATCCCCCTGAC
841  ---------+---------+---------+---------+---------+---------+   900
       A  V  I  A  P  W  M  A  S  *
```

FIGURE 27 CONTINUED

```
     CCCTCTCTTCGGGCTGCTCAAGCCCAACGGCAAGATGATCATGGTCGGCCTCCCCGAGAA
901  ---------+---------+---------+---------+---------+---------+   960

     GCCCATCGAGATTCCTCCCTTCGCTCTAGTTGCCACGAATAAGACCCTGGCCGGGAGCAT
961  ---------+---------+---------+---------+---------+---------+   1020

     CATCGGCGGCATGAGCGACACGCAGGAGATGCTGGACCTCGCGGCGAAGCACGGCGTGAC
1021 ---------+---------+---------+---------+---------+---------+   1080

     GGCCGACATCGAGGTGGTCGGCGCGGAGTATGTGAACACGGCCTTGGAGCGCCTTGCCAA
1081 ---------+---------+---------+---------+---------+---------+   1140

     GAACGACGTCAGGTATCGCTTCGTCATCGACATCGGCAACACCCTCGACAATGTTGCGGC
1141 ---------+---------+---------+---------+---------+---------+   1200

     CACCACCGAGTGAACGTACTCAGCACTGCTTACGATCTACGTTGTTCCACTGTTAGTGCT
1201 ---------+---------+---------+---------+---------+---------+   1260

     CCGTAGTAAACAATAAACGATCAAAACTCTTGTCATCTGGTGCATTGGTGTAGACATGGT
1261 ---------+---------+---------+---------+---------+---------+   1320

     TGTTTGCGAGGAAACTGAGTTGAAGGATGGATGGATAAAAAAAAAAAAAAAAAAAAAA
A)   ---------+---------+---------+---------+---------+-------   1377
```

FIGURE 27 CONTINUED

A


BamHI
EcoRI
HindII
KpnI
SalI
XbaI
Kb
4.4
2.0

B

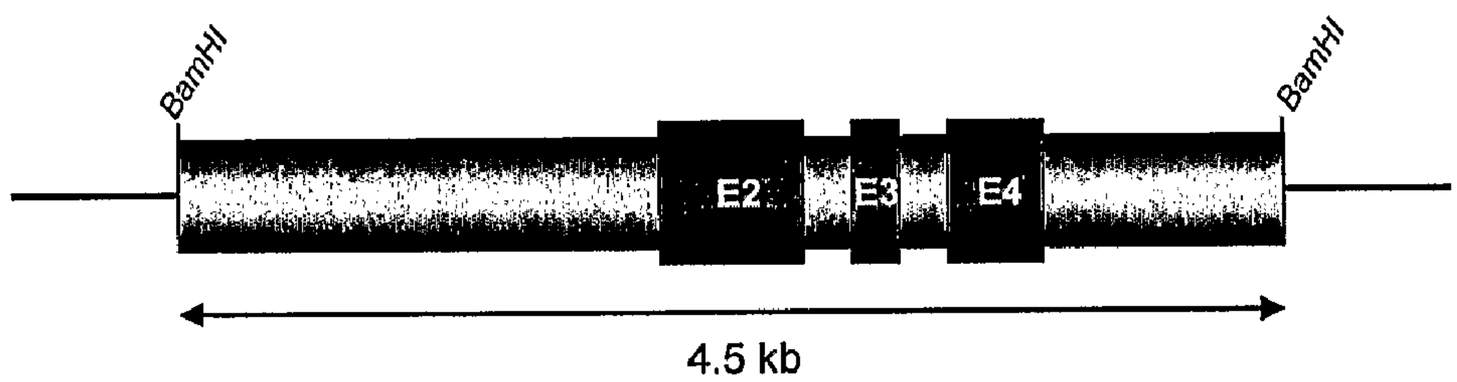
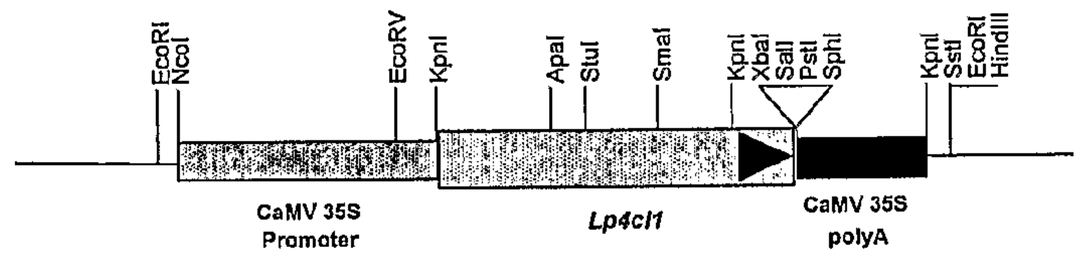

BamHI
BamHI
E2
E3
E4
4.5 kb

FIGURE 28

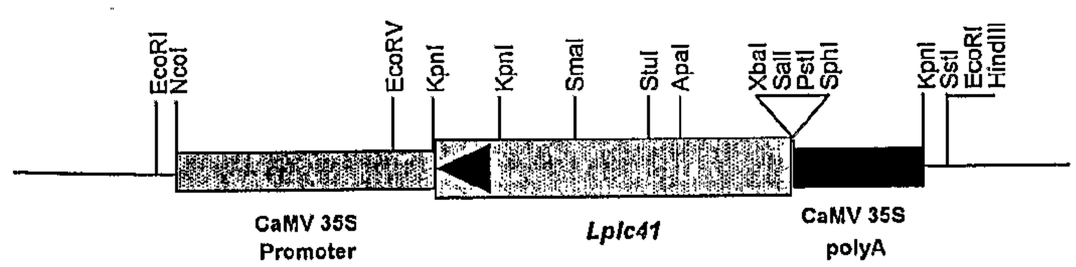
FIGURE 29A

A)

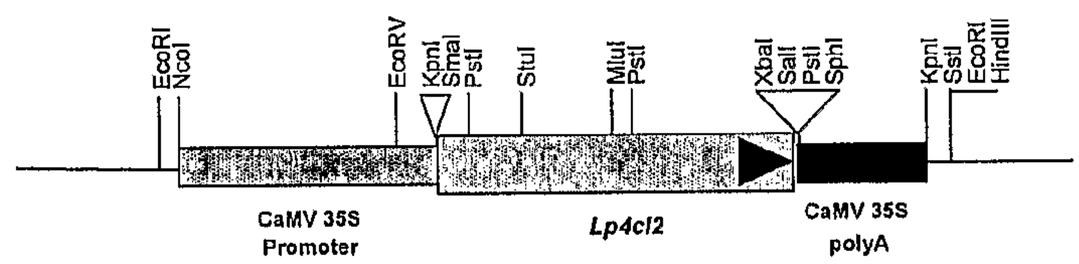
p35S4cl1



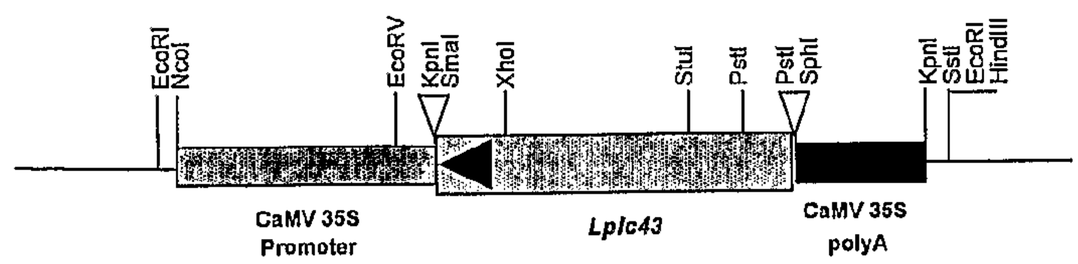
p35Slc41



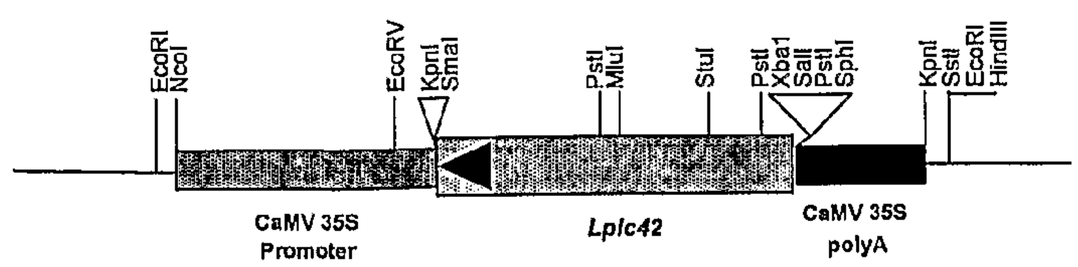
p35S4cl2



p35Slc43



p35Slc42



p35S4cl3

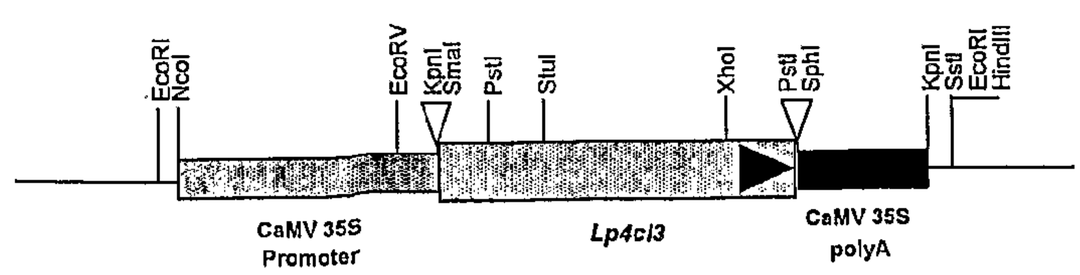
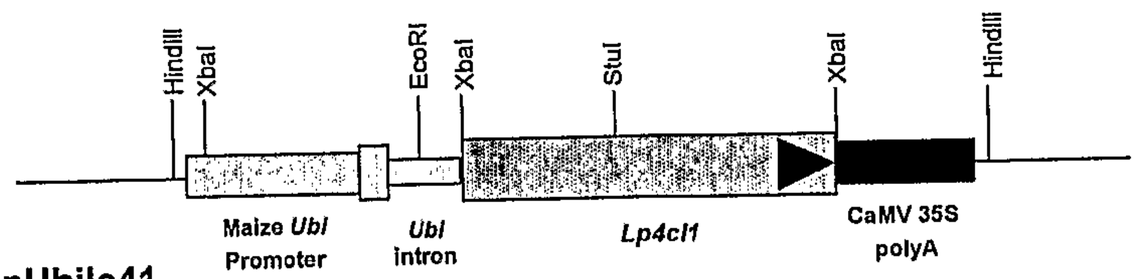


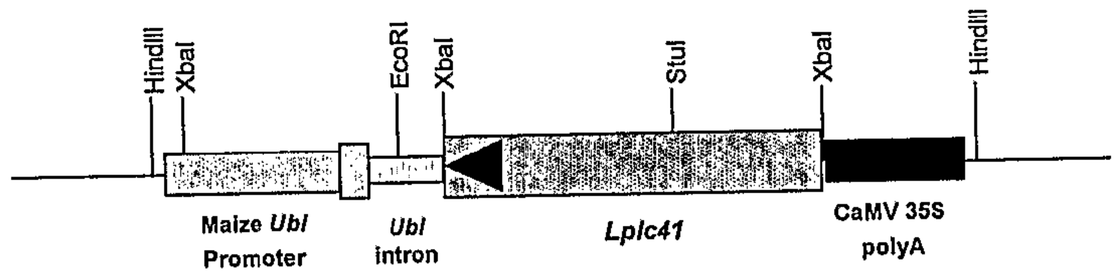
FIGURE 29B

B)

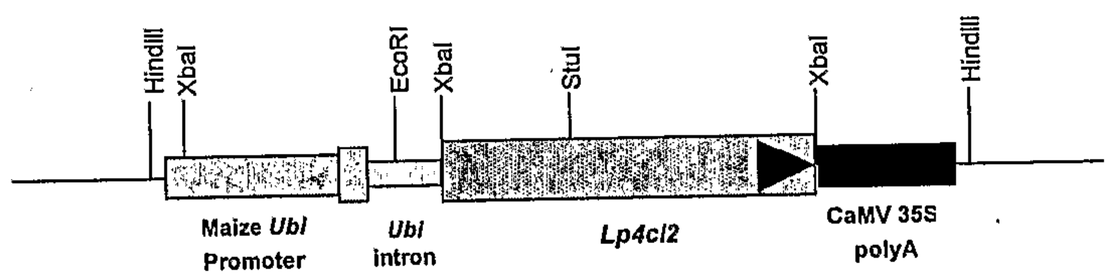
pUbi4cl1



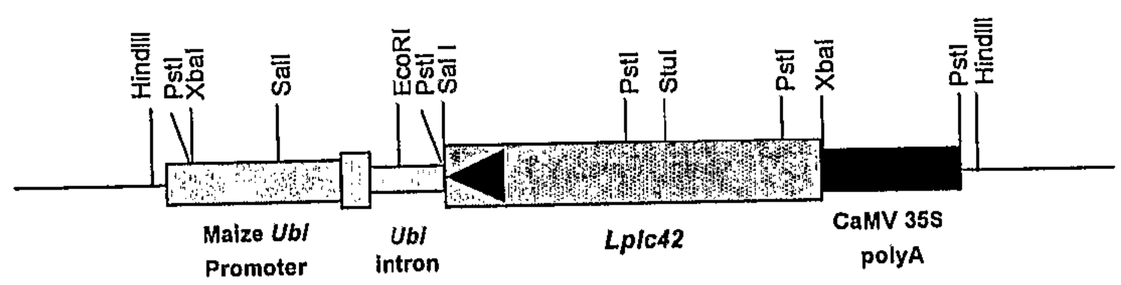
pUbilc41



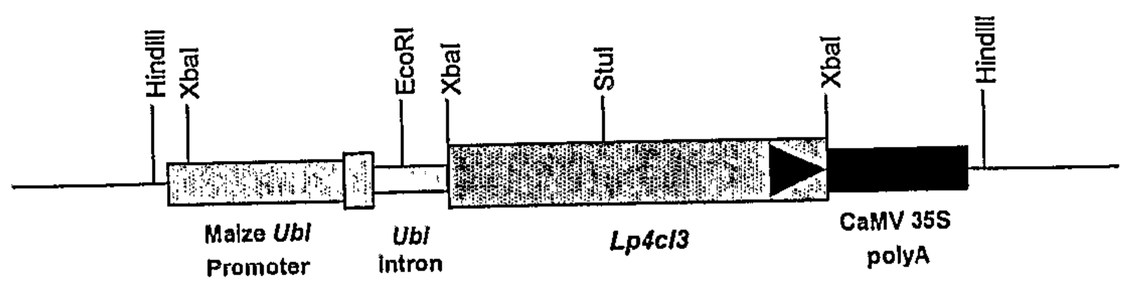
pUbi4cl2



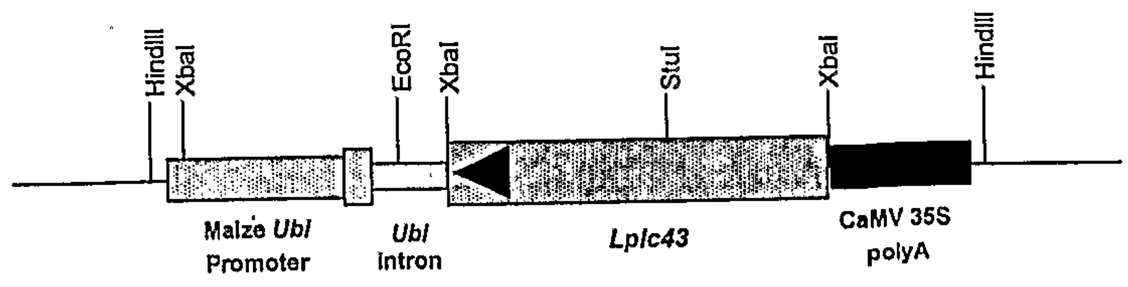
pUbilc42



pUbi4cl3

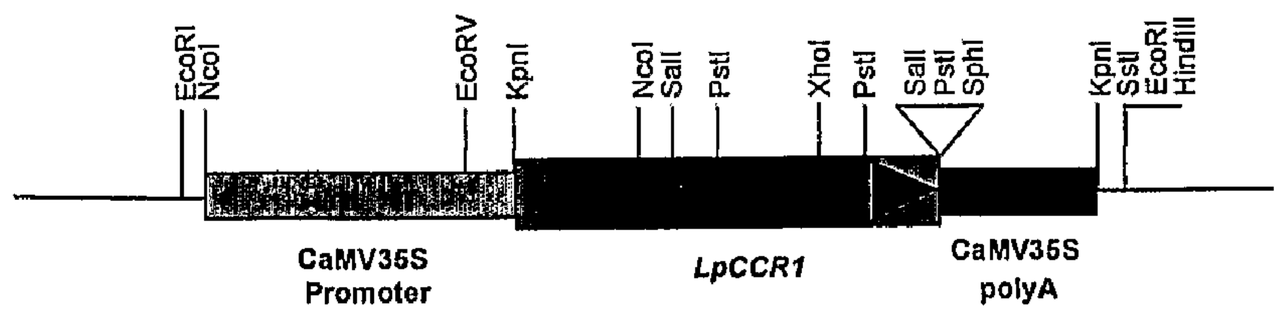


pUbilc43

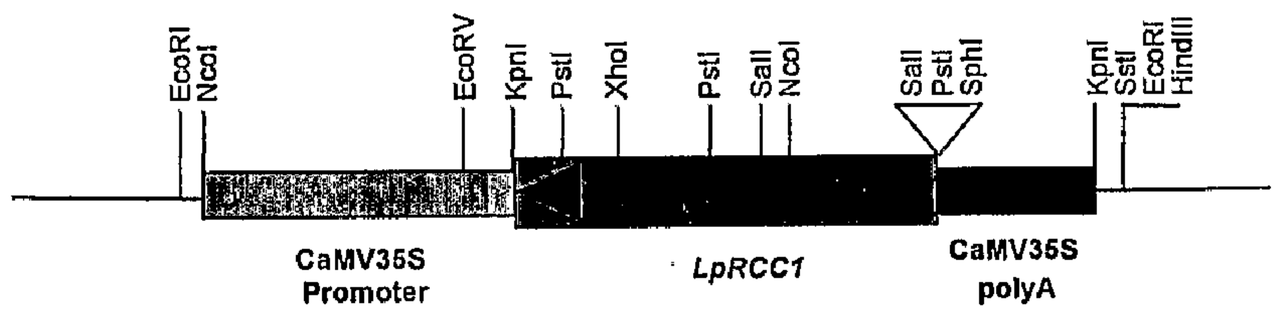


A)

p35SCCR1

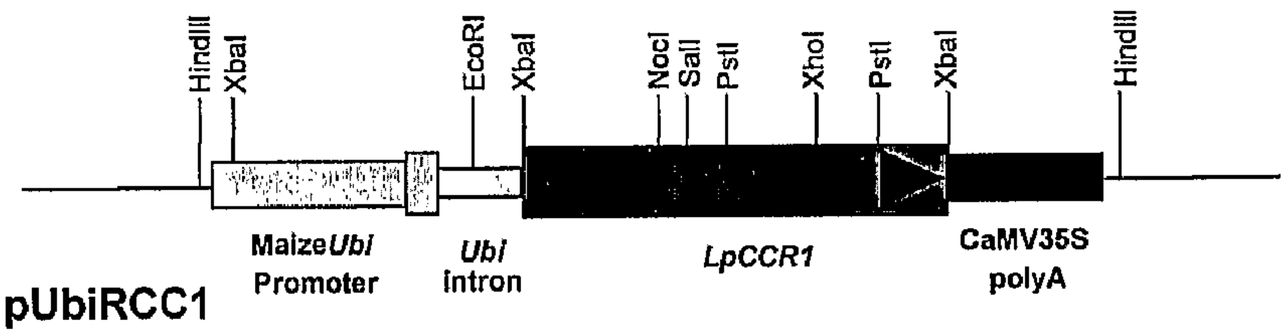


p35SRCC1



B)

pUbiCCR1



pUbiRCC1

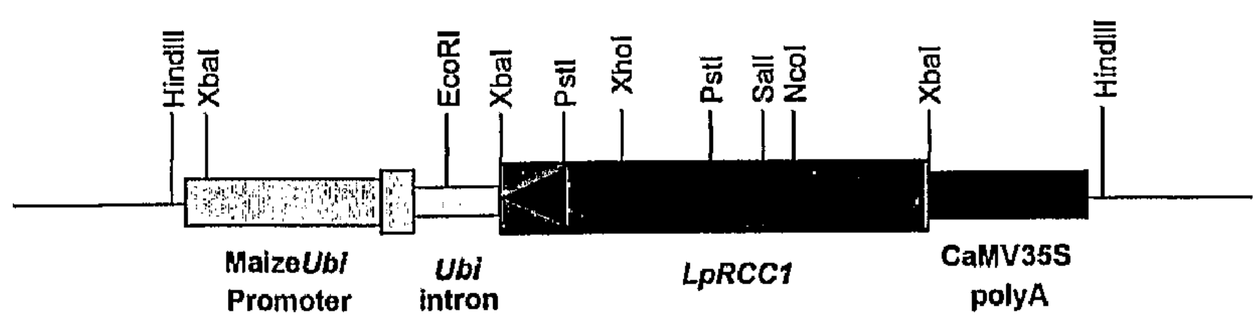
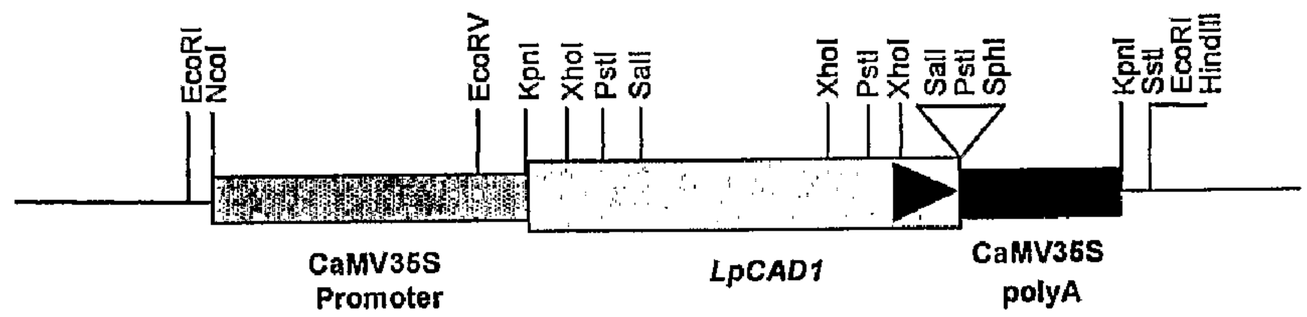


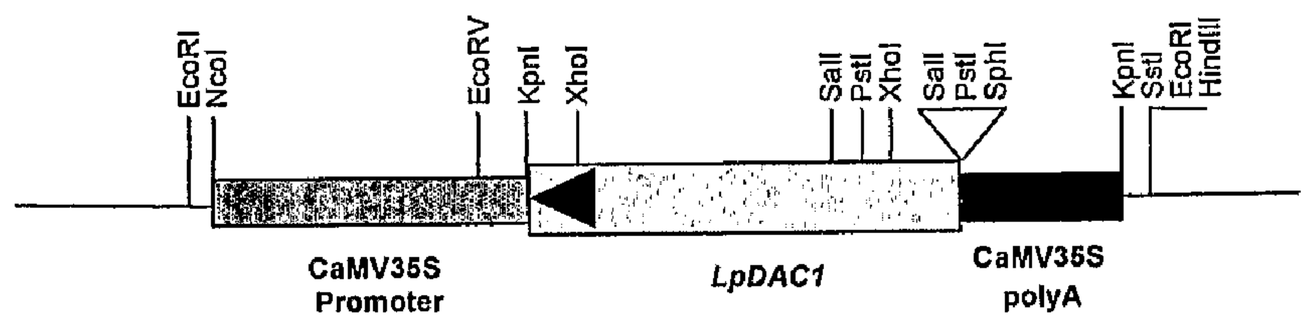
FIGURE 30

A)

p35SCAD1

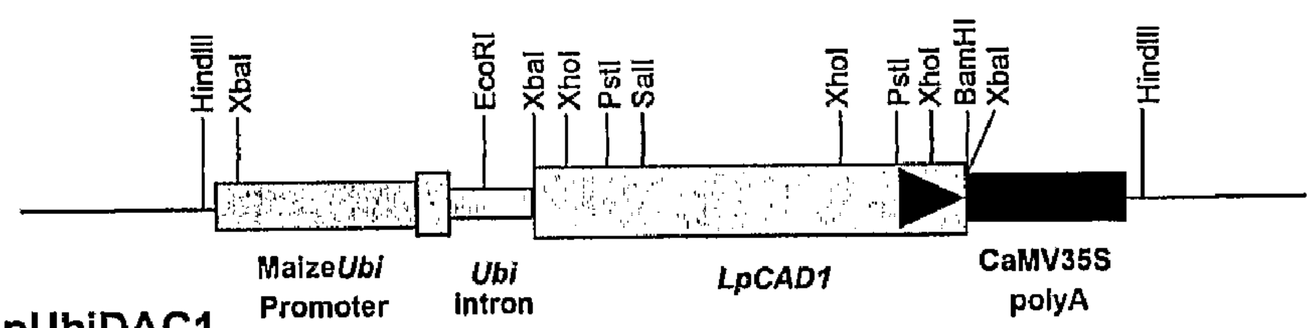


p35SDAC1



B)

pUbiCAD1



pUbiDAC1

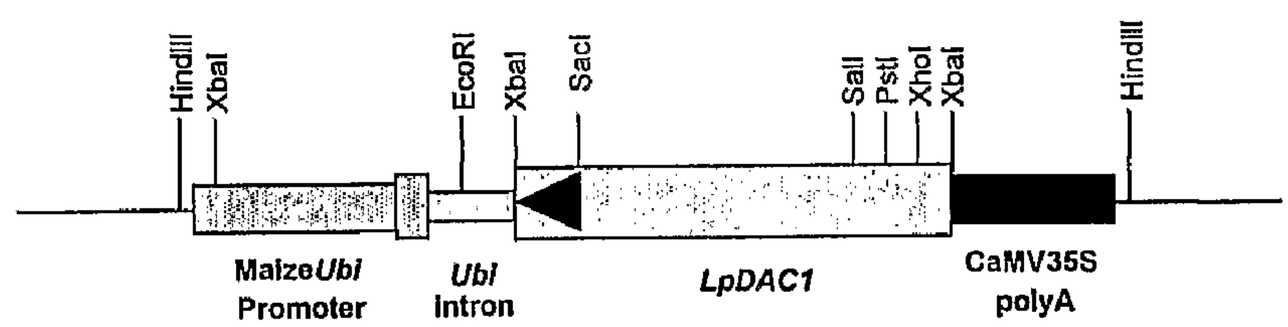


FIGURE 31

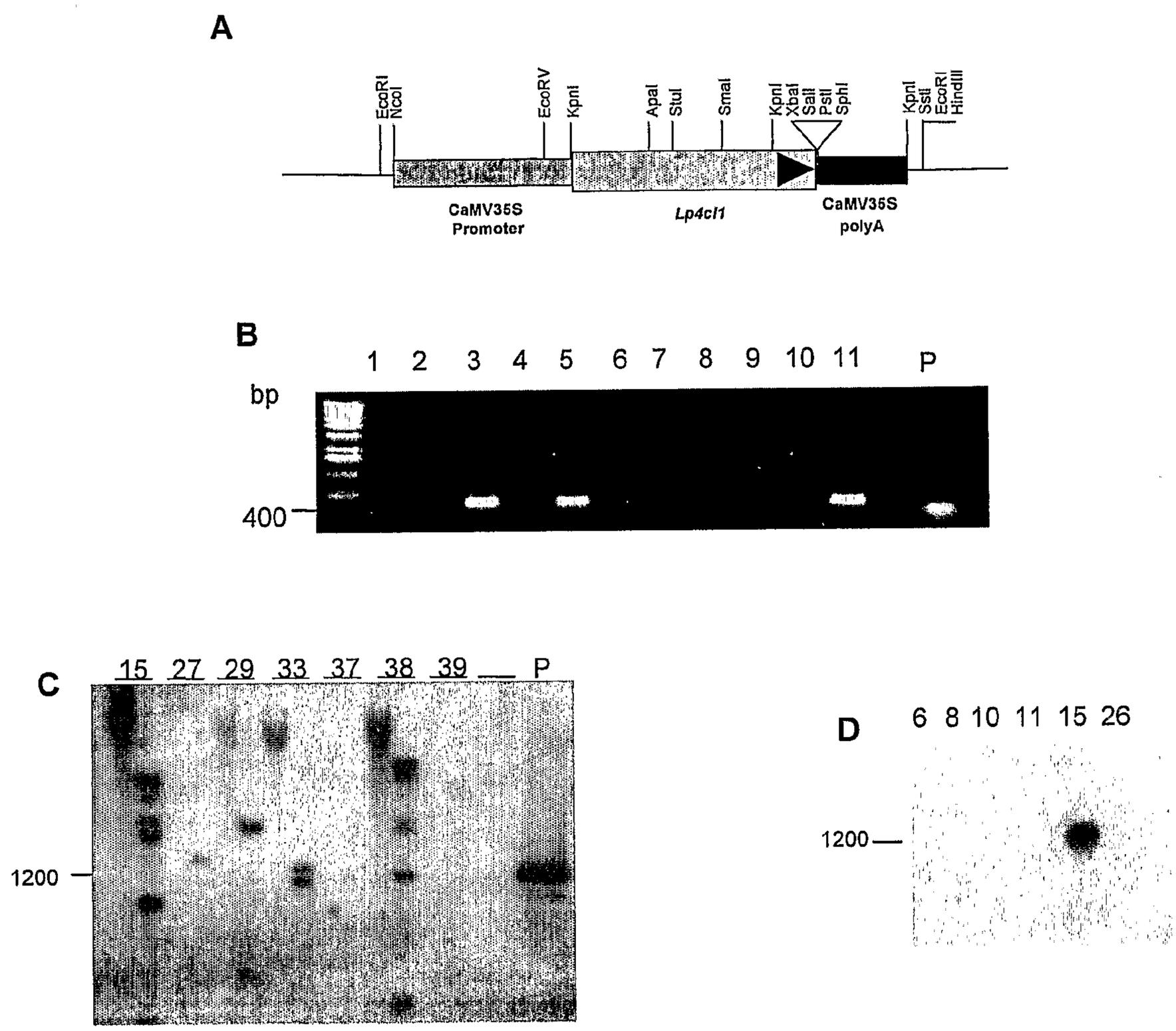
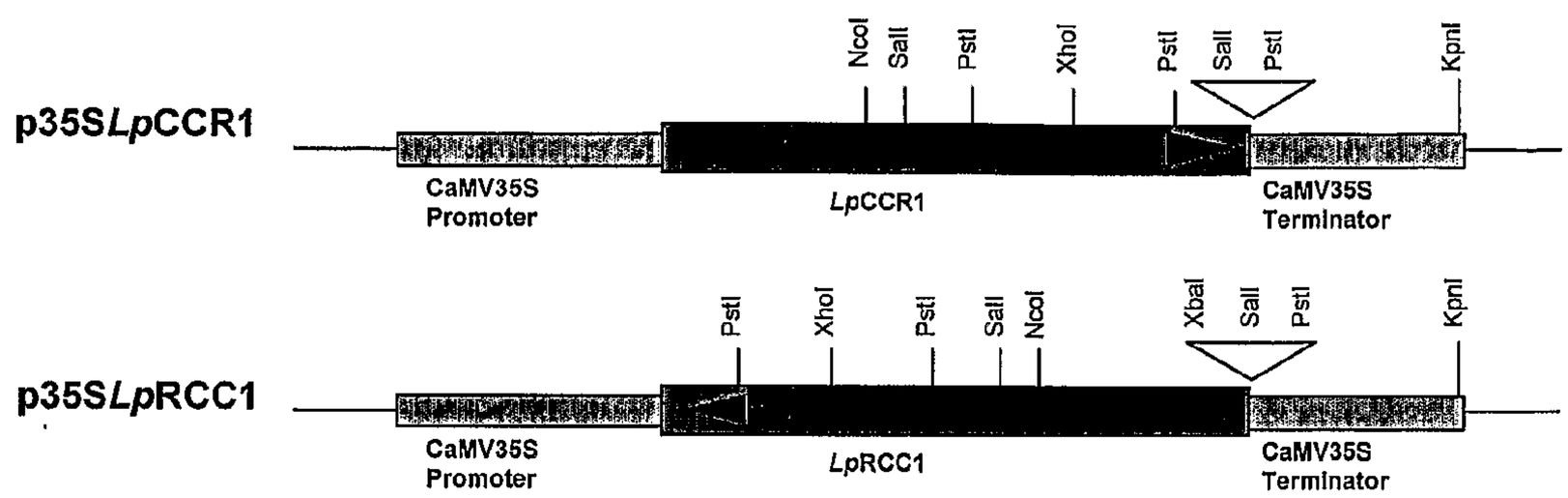

A
EcoRI
NcoI
EcoRV
KpnI
ApaI
StuI
SmaI
KpnI
XbaI
SalI
PstI
SphI
KpnI
SstI
EcoRI
HindIII
CaMV35S Promoter
Lp4cl1
CaMV35S polyA
B
1 2 3 4 5 6 7 8 9 10 11 P
bp
400
C
15 27 29 33 37 38 39 P
1200
D
6 8 10 11 15 26
1200

FIGURE 32

A


p35SLpCCR1
NcoI
SalI
PstI
XhoI
PstI
SalI
PstI
KpnI
CaMV35S Promoter
LpCCR1
CaMV35S Terminator
p35SLpRCC1
PstI
XhoI
PstI
SalI
NcoI
XbaI
SalI
PstI
KpnI
CaMV35S Promoter
LpRCC1
CaMV35S Terminator

B

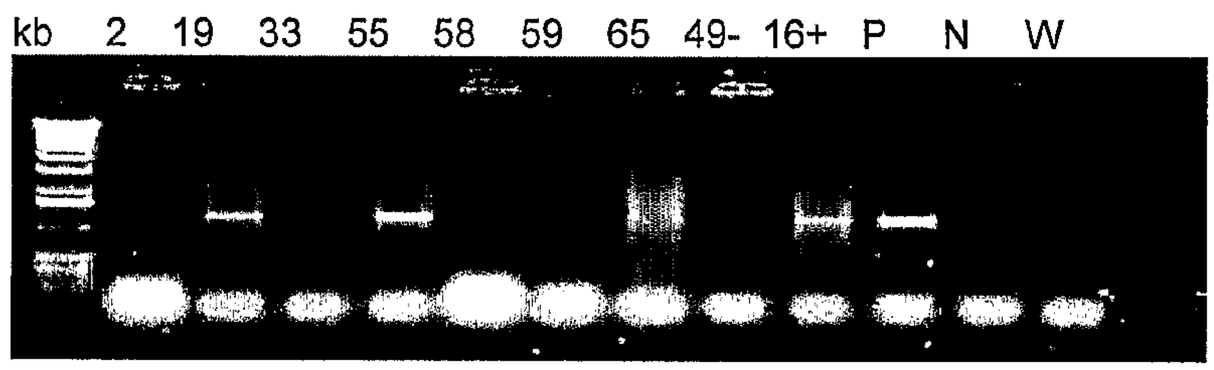

kb 2 19 33 55 58 59 65 49- 16+ P N W

FIGURE 33

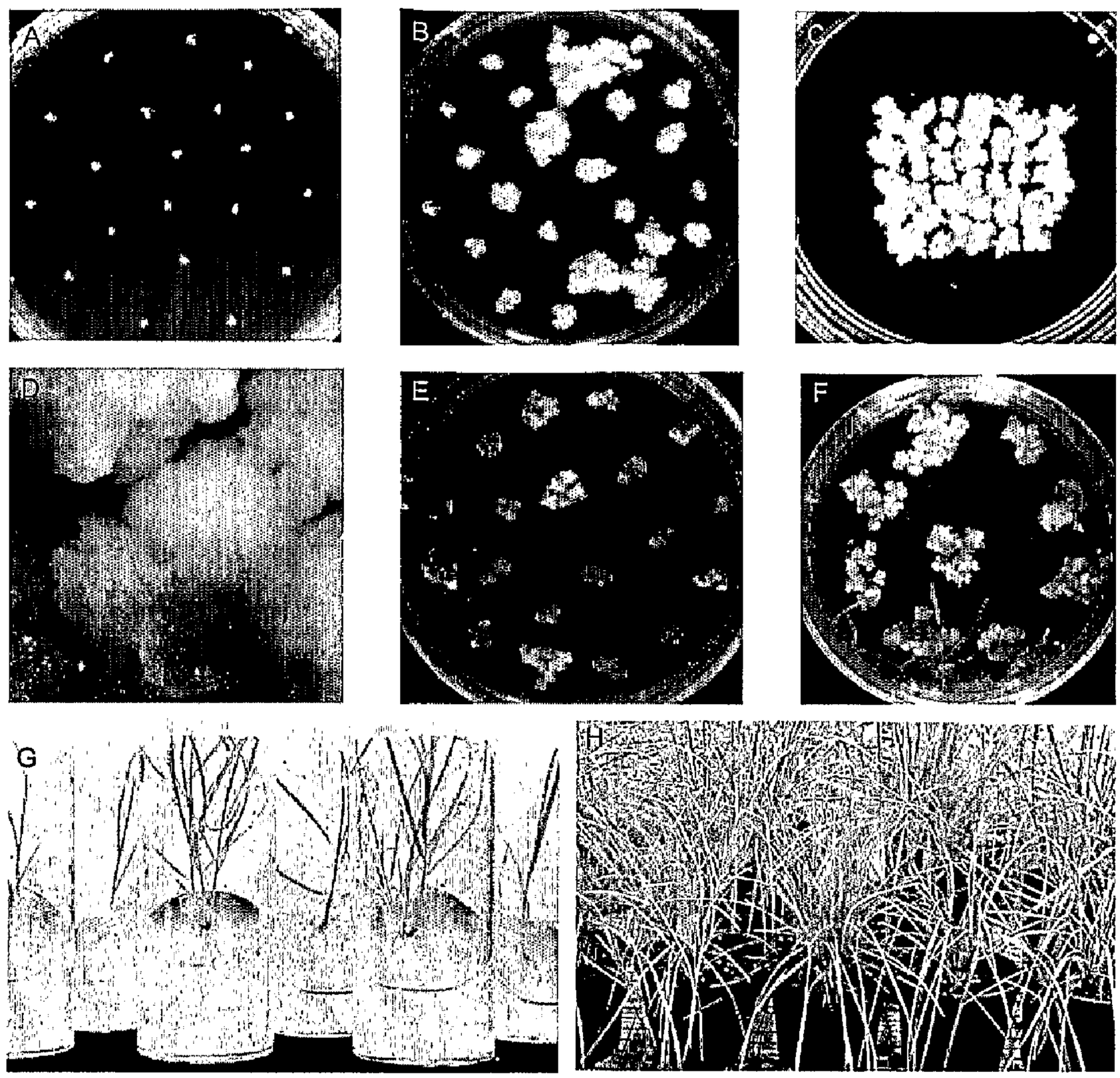
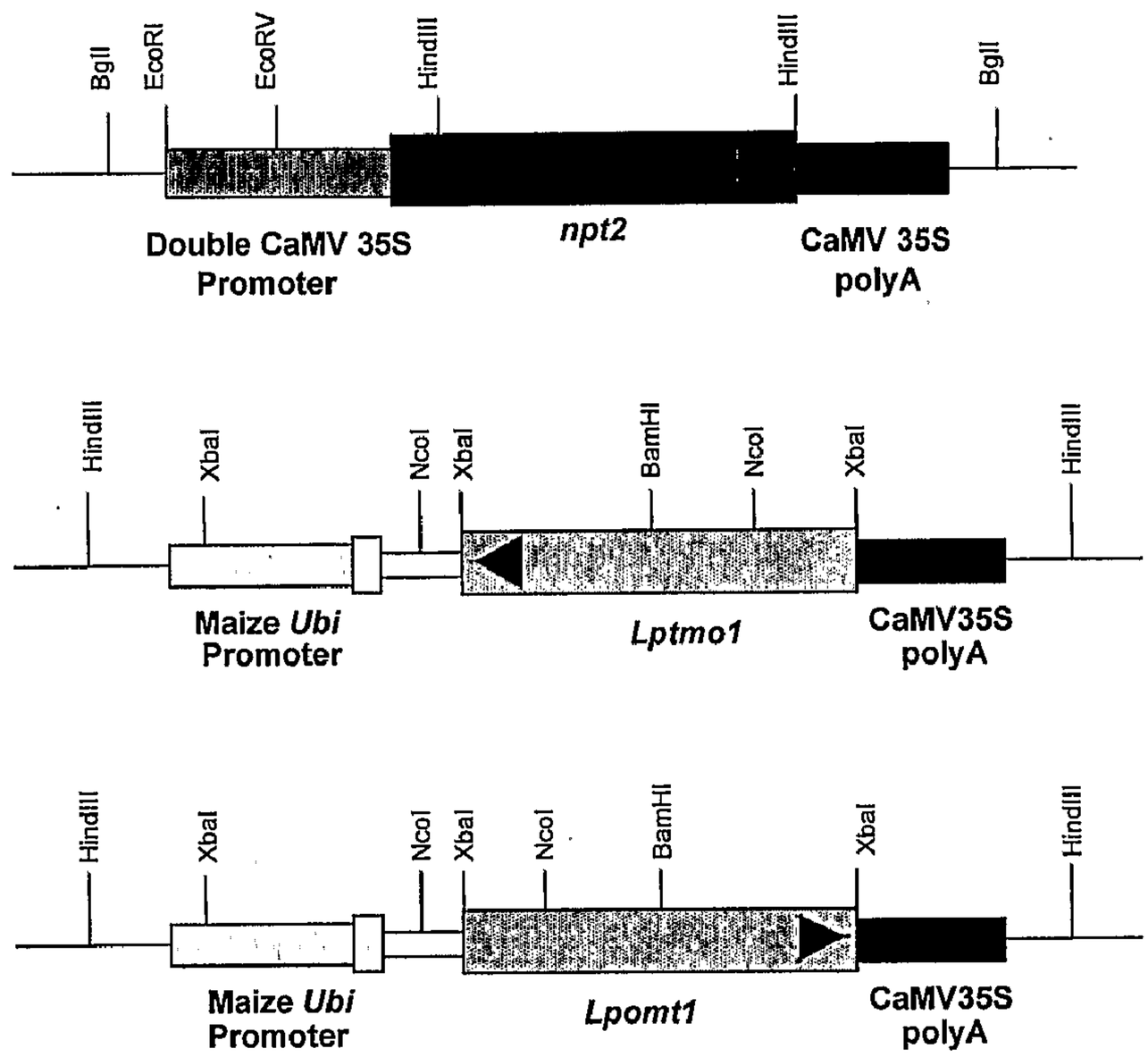


FIGURE 34

A


BglI
EcoRI
EcoRV
HindIII
HindIII
BglI
Double CaMV 35S Promoter
npt2
CaMV 35S polyA
HindIII
XbaI
NcoI
XbaI
BamHI
NcoI
XbaI
HindIII
Maize Ubi Promoter
Lptmo1
CaMV35S polyA
HindIII
XbaI
NcoI
XbaI
NcoI
BamHI
XbaI
HindIII
Maize Ubi Promoter
Lpomt1
CaMV35S polyA

B

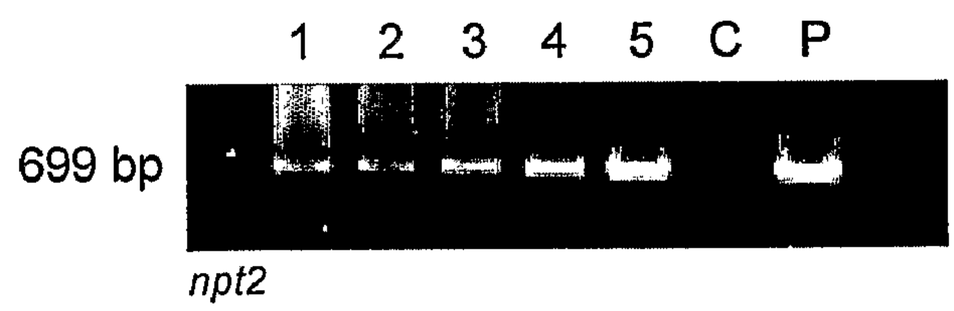
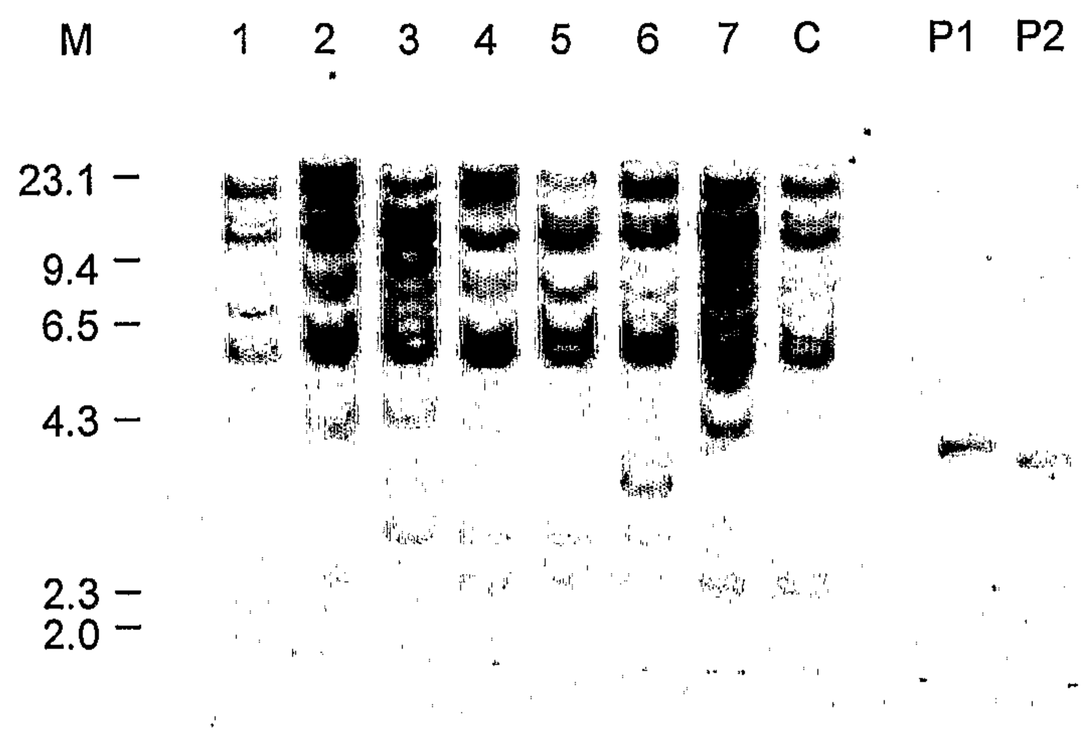

1 2 3 4 5 C P
699 bp
npt2

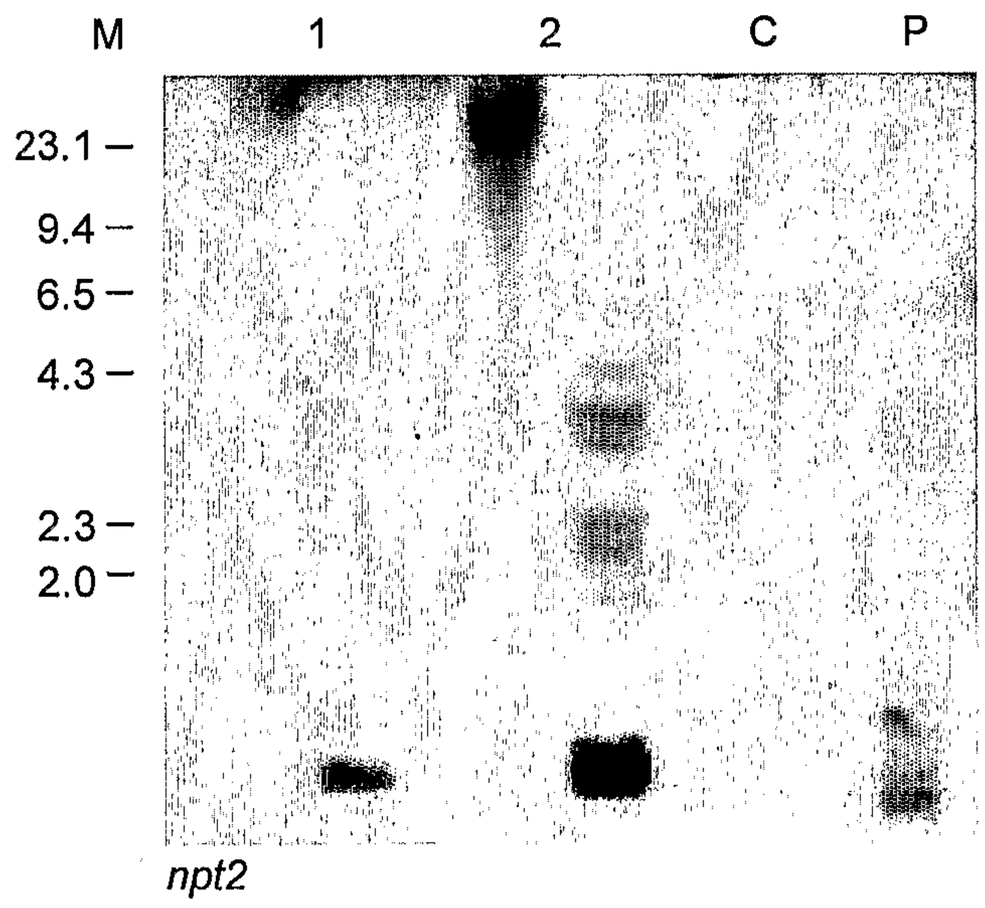
FIGURE 35

FIGURE 35 CONTINUED

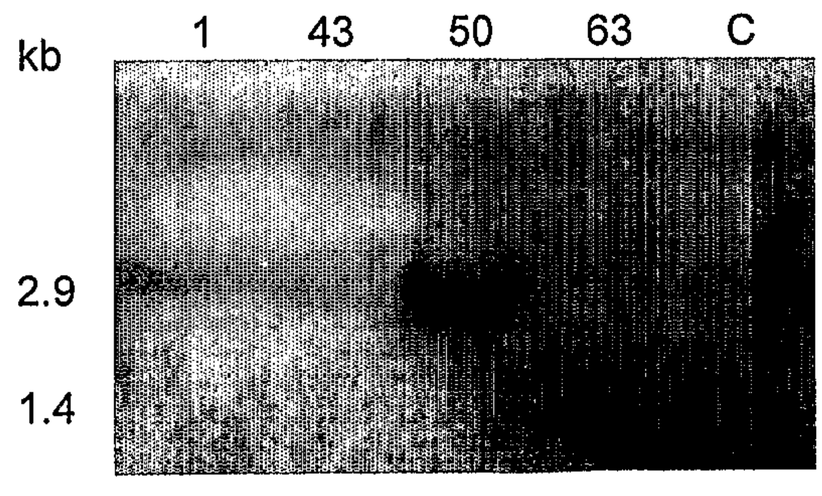
C



D



E



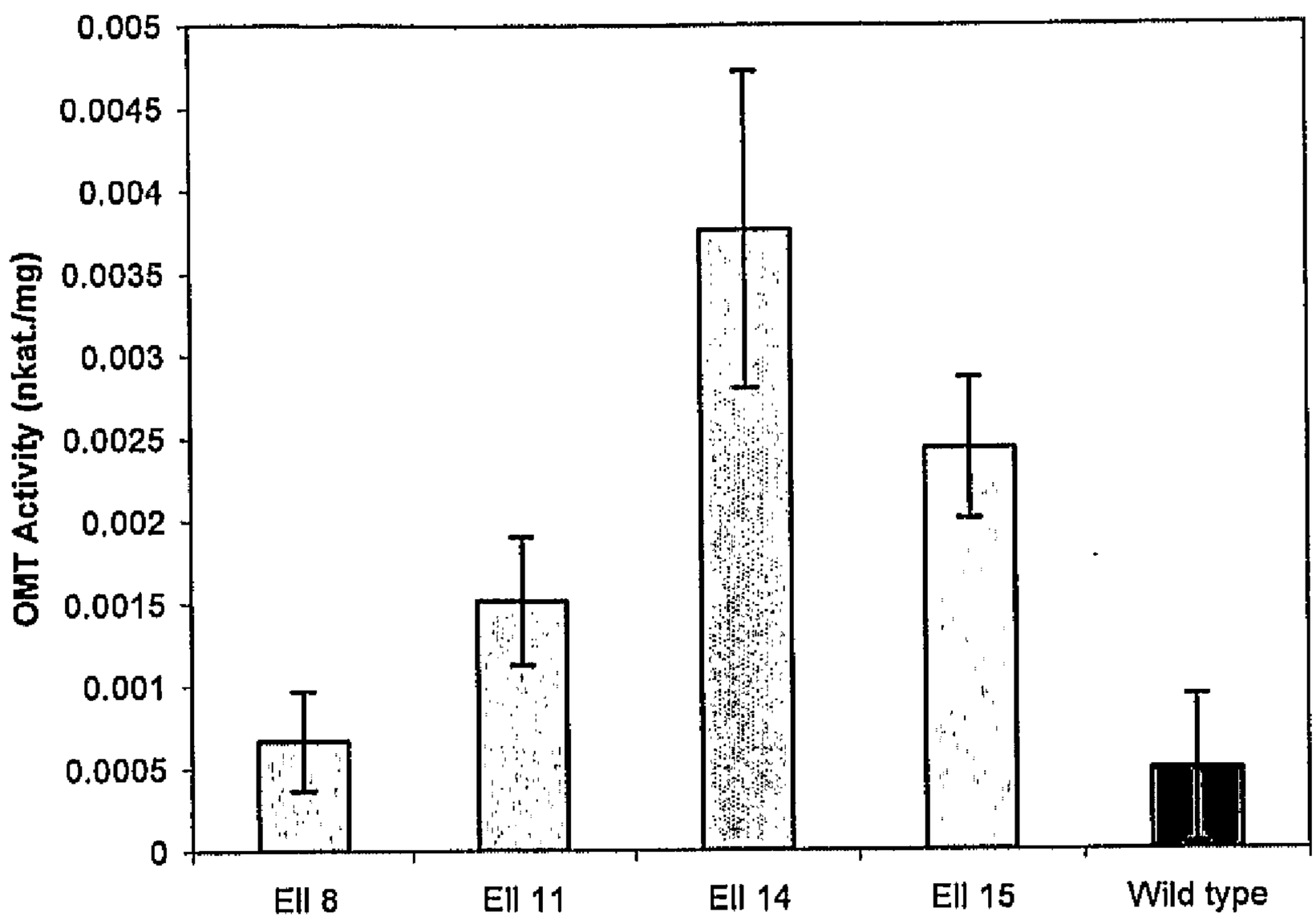

0.005
0.0045
0.004
0.0035
0.003
0.0025
0.002
0.0015
0.001
0.0005
0
OMT Activity (nkat./mg)
Ell 8
Ell 11
Ell 14
Ell 15
Wild type

FIGURE 36

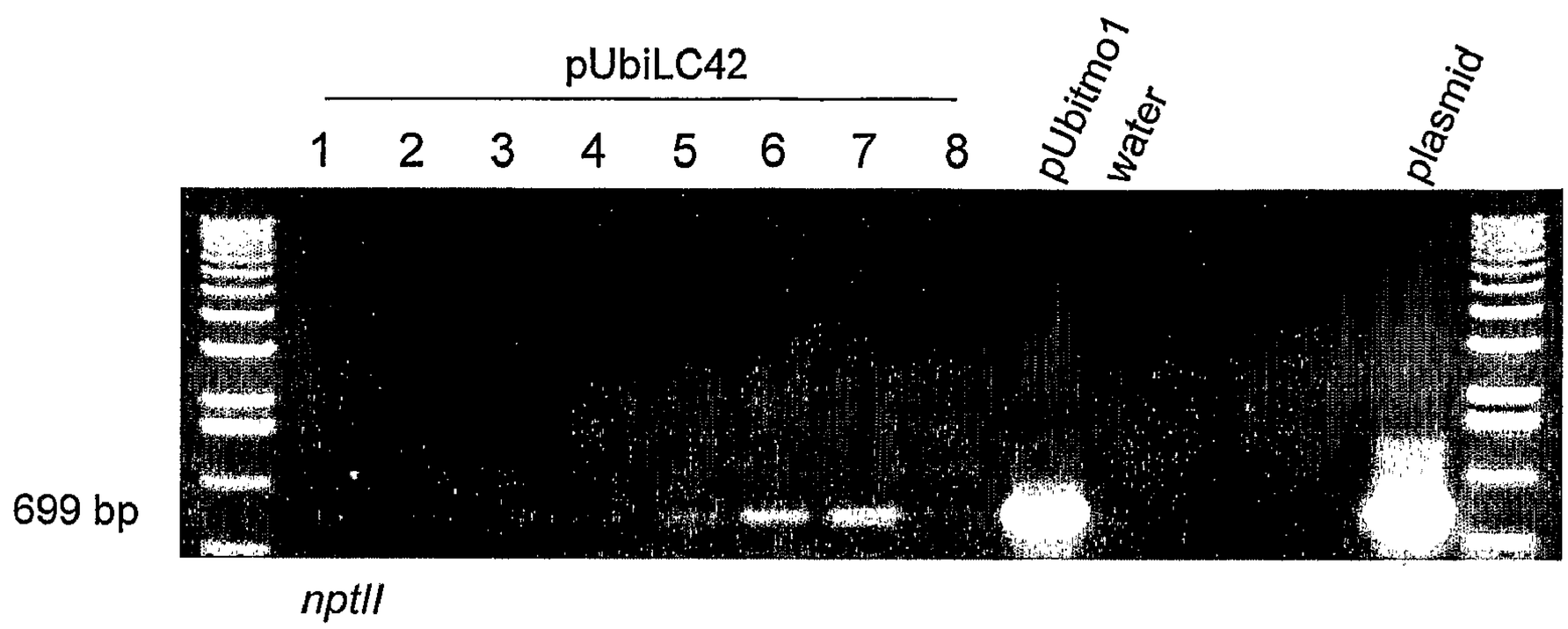


FIGURE 37

```
       CGGGATCAACTTGGATGTCCTTTGCGGGCACGGTTTCAGGAACAACGACACATGCAGCAG
-2206  ------+---------+---------+---------+---------+---------+---   -2147

       GGATCTCCTCCAAAGACTCACACAAAGGTGACATGAGCGCCCGCTTTTTTGAAGCCAAGT
-2146  ------+---------+---------+---------+---------+---------+---   -2087

        TGGCTAAGAAATCGCAAAGCTTGGTGGAGTCGGCCACCTCAGGATCTGCAACAAAAGGCA
-2086  ------+---------+---------+---------+---------+---------+---   -2027

       CCAAGGGAGCTGCCAACACATCAACCACAACATCATGTTCAAACGCAGTCTCCTCAAGCC
-2026  ------+---------+---------+---------+---------+---------+---   -1967

       TCGAATGCTCAACCGAAAGAGAGGCAGAAGCTTCAACAAAAAACTCAGCCAACCCAAAGC
-1966  ------+---------+---------+---------+---------+---------+---   -1907

       CCTCGACGTCATCAGAGATTAGGCTCTGAGGACCCGCAGGGAAGCAACCTTGTCAACAAC
-1906  ------+---------+---------+---------+---------+---------+---   -1847

       CGCATCCGGCAGAAAAGGAGCAAGACCGGAGCAACCCTCAAGAGGCACACGAAAGACGTC
-1846  ------+---------+---------+---------+---------+---------+---   -1787

       GAAGCCAAGAGGAGACGAGTCGCAGGGACGGCGGACAGGCGAGAAGGGGCCGTAGAACTC
-1786  ------+---------+---------+---------+---------+---------+---   -1727

       CAAGAGCTCGGCGTCCCTCGACCTAGCATCCGAAGCACTGACCGGGGCACTCAATGCATA
-1726  ------+---------+---------+---------+---------+---------+---   -1667

       ACTTTATCTTGATGGCATATGTACTCAAACCCATACAATGTTCACCATGCATTATCTATG
-1666  ------+---------+---------+---------+---------+---------+---   -1607

       GAACATTCCTTCATATACAACTtCTGAGTGGTCAGTGCATAGGAATTTTCATTAACAACC
-1606  ------+---------+---------+---------+---------+---------+---   -1547

       AAAAACATACTTGGGGCCTACACACACTTTCACAGCATGGAAAACTTGTTAGCTTTTTAA
-1546  ------+---------+---------+---------+---------+---------+---   -1487

       AGAGTTGCAAAATCTGTCAAGCGAATGTTCTTGTGATAATTGGAACGAAGCATGTTTCCC
-1486  ------+---------+---------+---------+---------+---------+---   -1427

       CATTTTCAATGTGTGTCTCTTACCCTAACTAGCACCCGACCAACAAAATCTGACCATCCT
-1426  ------+---------+---------+---------+---------+---------+---   -1367
```

FIGURE 38

```
       AGTTATATCATCATAGAGACCCACATGTAGGTTGACCCCCATAACACTTGTGTGGATATC
-1366  ------+---------+---------+---------+---------+---------+---   -1307

       ATGGAAAATGGCCTTGATCAACACTTTCTTTCCTACTTGGTACAAATGGTTATGGACTTA
-1306  ------+---------+---------+---------+---------+---------+---   -1247

       CTCAATTAGTGCTTTAGAGAGCTTTGGCTGCAGACTTTGTAGCTTCCCAATATTCATAGG
-1246  ------+---------+---------+---------+---------+---------+---   -1187

       TCCCTCCGGAGTGGGCAGCCCCATCTACATAGGCTCAAAACCAGATTTTTGTAACATGTT
-1186  ------+---------+---------+---------+---------+---------+---   -1127

       AGACACTTTCAACTTCATCATAGACCATCAAGGAGCTGGCATGTGACAGTGATATATGTA
-1126  ------+---------+---------+---------+---------+---------+---   -1067

       TCAATTACCCATTCAACACGAATAGCTTGCTCATGCATGGTTAGTCTTGCGGCGGCGGGG
-1066  ------+---------+---------+---------+---------+---------+---   -1007

       CGGGACCATCGAACACACCGCCGGGCGGTCAGTAGGCTAGGGTTAGATAAAATCTAGCCG
-1006  ------+---------+---------+---------+---------+---------+---   -947

       TTTTCATTCAAACTTGTGATATATAATCAAATTTAAATAAAAACCTTTATTTTCGTGCAT
-946   ------+---------+---------+---------+---------+---------+---   -887

       TTTTATTTATTTGAGGGCGTGTTTGGGGGACACGGCTGGAAAGTGACATCCCCAAACACT
-886   ------+---------+---------+---------+---------+---------+---   -827

       GCACGAAGAAAACGCGTCGCCAAAAAATTCGATCCGGCGTCAGTCCTTTGGGAGACGATT
-826   ------+---------+---------+---------+---------+---------+---   -767

       TGGATGACGCGGCTAGAGATGCTCTAAGTTCTCCACGCCATGTTTCTTTCTATATATACA
-766   ------+---------+---------+---------+---------+---------+---   -707

       CACAGCCCAAGGTCCATGAAAAGTAAAACGGCACGACGACACGCACCGGCGACAACTTCA
-706   ------+---------+---------+---------+---------+---------+---   -647

       CATTACGGCACATCGCTATTACGGACCACATACAACTCCACCGCTATTCTCAGCCAAGTC
-646   ------+---------+---------+---------+---------+---------+---   -587

       ATACATGACATGATCCAATGGACGACTTTGTGAGCGAAACTAGAACCTTGCGGGGTTTAG
-586   ------+---------+---------+---------+---------+---------+---   -527

       ATTTTCCAATGTGGATAAGTTGTACGCGCCGACTAGCTTTACACTTGGTTGAAAAAAGCT
-526   ------+---------+---------+---------+---------+---------+---   -467
```

FIGURE 38 CONTINUED

```
     TATTGTAGCACGACTTCTCACTGACATAGGAATGTAAACAGTCTCTCCACGCCATGTTTC
-466 ------+---------+---------+---------+---------+---------+---  -407

     TTTCTAGTAGTAGCATACTAGTAGTAACTTCTCTTTGTCCTACACACACCCAGGGTCCAA
-406 ------+---------+---------+---------+---------+---------+---  -347

     GAAAGGAAAACGGCACGACGGCACCCACCGACGACGACGACTCCACATCACGGTTCGGTA
-346 ------+---------+---------+---------+---------+---------+---  -287

     AAAAAAGTCAAAACTCGCTGACGTGGCACCACCGGTCGCAGTCAACTGACGCGCTCCTCT
-286 ------+---------+---------+---------+---------+---------+---  -227

     GCGCAGGTyTCACTtCAAGTTTCACCTACCACTGTGGGCCCACCGCCAaTGTGGGCCCCG
-226 ------+---------+---------+---------+---------+---------+---  -167

     CGAGCTtCTtACTCACTGACCTGTCTCCCACCAGCCTCCTCGCCGGTATATTACCCCGGC
-166 ------+---------+---------+---------+---------+---------+---  -107

     CCCCAATTTCCTCTGCCTTCCCACGAGCAGCAGCCGGAGCACGGAATCCCGGCCGCCATT
-106 ------+---------+---------+---------+---------+---------+---  -47

     CCTCCACCTTCAGCTCCGCCCAAAGATTTCCATCCGGCGAGATCCATGGGCTCCATCGCG
 -46 ------+---------+---------+---------+---------+---------+---  13
                                                M  G  S  I  A

     GCGGACGCGCCTCCCGCGGAGCTGGTGTTCCGGTCCAAGCTCCCGGACATCGAGATCCCG
  14 ------+---------+---------+---------+---------+---------+---  73
     A  D  A  P  P  A  E  L  V  F  R  S  K  L  P  D  I  E  I  P

     ACCCACCTGACGCTGCAGGACTACTGCTTCCAGCGCCTGCCGGAGCTCTCCGCGCGCGCC
  74 ------+---------+---------+---------+---------+---------+---  133
     T  H  L  T  L  Q  D  Y  C  F  Q  R  L  P  E  L  S  A  R  A

     TGCCTCATCGACGGCGCCACGGGCGCCGCGCTCACCTACGCCGACGTGGACGCCCTCACG
 134 ------+---------+---------+---------+---------+---------+---  193
     C  L  I  D  G  A  T  G  A  A  L  T  Y  A  D  V  D  A  L  T

     CGCCGCTGCGCCGCGGGCCTCCGCCGCCTGGGGGTCCGCAAGGGCGACGTCGTCATGGCG
 194 ------+---------+---------+---------+---------+---------+---  253
     R  R  C  A  A  G  L  R  R  L  G  V  R  K  G  D  V  V  M  A

     CTGCTCCGCAACTGCCCCGAGTTCGCCTTCGTGTTCCTCGGCGCCGCCCGGCTCGGCGCC
 254 ------+---------+---------+---------+---------+---------+---  313
     L  L  R  N  C  P  E  F  A  F  V  F  L  G  A  A  R  L  G  A
```

FIGURE 38 CONTINUED

```
     GCCACCACCACCGCCAACCCGTTCTACACGCCCCACGAGATCCACCGCCAGGCGACCGCC
314  ------+---------+---------+---------+---------+---------+---   373
     A  T  T  T  A  N  P  F  Y  T  P  H  E  I  H  R  Q  A  T  A

     GCCGGGGCCAGGGTCATCGTCACCGAGGCCTGCGCCGTCGAGAAGGTGCGCGCCTTCGCC
374  ------+---------+---------+---------+---------+---------+---   433
     A  G  A  R  V  I  V  T  E  A  C  A  V  E  K  V  R  A  F  A

     GCCGAGAGAG
434  ------+---  443
     A  E  R
```

FIGURE 38 CONTINUED

```
       TCGACGCGGCCGCGTAATACGACTCACTATAGGGCGAAGAATTCGGATCATATGGATTCG
-6735  -----+---------+---------+---------+---------+---------+----  -6676

       ACACTGGAATTTACTCCCATCGGGAGCGTGCAAACAAAAAGGTGTTATAGCAAGAAGACA
-6675  -----+---------+---------+---------+---------+---------+----  -6616

       CTGGCAACATTGCCAGCACAGAATTTGTTACAATCATAGAAAGTTTTATGACAGGACATT
-6615  -----+---------+---------+---------+---------+---------+----  -6556

       GTTTCAACCGAAAGCAAGATTACAACAATATAATCAAGGGCTTGGGTCTGGTTGGACATG
-6555  -----+---------+---------+---------+---------+---------+----  -6496

       CTCGGTCCAATGGACGATTTATTTGCCGAGACCAGCTCAAGGAGTTGACGAGCACACTTA
-6495  -----+---------+---------+---------+---------+---------+----  -6436

       AGCGCCGAGATCTTAAAGGCACCCAAGTCAACAAGTCGCCCATCTTGCTCTTTTGGCAGC
-6435  -----+---------+---------+---------+---------+---------+----  -6376

       TCCTTGGACATCTCTTCGATATTGGCTTTGAAGCCATGACCCATCATAAGCTGAAAGGCT
-6375  -----+---------+---------+---------+---------+---------+----  -6316

       AGGAGGGCACCATAGGTACGCGAAGTACGTTTGAATACCTCGAGGACCTCCCTCGTGTTG
-6315  -----+---------+---------+---------+---------+---------+----  -6256

       ATGGCGAAAGCATCGATCAGCTGCCCCAAGGTCTTGTTTTGATCGATCTTGGGGAAGATC
-6255  -----+---------+---------+---------+---------+---------+----  -6196

       ATCGAGTGCATCCGCGTCATGGATCCTTTACCCTTCTGAAGGAGGTCCTGAAAAAGCTGG
-6195  -----+---------+---------+---------+---------+---------+----  -6136

       TGAGACCCGAGGGTCATTGACAAAGCATTCGCCGGAGAATTATTCGGCAATTTATCTAGA
-6135  -----+---------+---------+---------+---------+---------+----  -6076

       GCCTCAGCAGGGATGTAGGCAGCTTCTGGAGAAAGTGAAAGAGGAGGAGCTCACTAACCA
-6075  -----+---------+---------+---------+---------+---------+----  -6016

       AAATCAAATCGATAAAGCAAAAATCGGAAAGGAGGCCAAAAGGGGATTACTGAGCAAGGC
-6015  -----+---------+---------+---------+---------+---------+----  -5956

       CAAGGAAGATTGGCGAAGGAGCTCATCTTTTTCAATCGCCCGAGCTTCGGCAGCAAGCCT
-5955  -----+---------+---------+---------+---------+---------+----  -5896

       GGATGCCTCTTCATCCTTCAGCCTCTTTCTTAGCCCCTCGAGCTCATCCTTAAAGGAATC
-5895  -----+---------+---------+---------+---------+---------+----  -5836

       AACCTCCTGGCGGGCCTCGGCAGCTATCTTTATCGCACCCTCCAGCTTCGAGGAAGAAGA
-5835  -----+---------+---------+---------+---------+---------+----  -5776

       CTCGACCTCCTTTTGCAGCCGAGTCTTGTCAACTTCCAGAGAAGTGTATTGGGAGGCGAA
-5775  -----+---------+---------+---------+---------+---------+----  -5716

       GGCCTCCAGAGAAGAGATAACAGCTCACAAATCCTTAAGAGATAAGGAAAAATAATTAGA
-5715  -----+---------+---------+---------+---------+---------+----  -5656

       CGAAGAACTGGTTGTCAACAAACTTATAATTTGATCAGGGAAATCGTCCCACATGGATAT
-5655  -----+---------+---------+---------+---------+---------+----  -5596

       ATCGTTAAAACAGGAAAAGCTTACAGGTTTCCCTGGAGGAGAAGCTGTAACCACGGCAGT
-5595  -----+---------+---------+---------+---------+---------+----  -5536
```

FIGURE 39

```
       CAAAGAAATCTCCTTCCCTTTGGAAAGGGAAGAAGTTGTCGATATTTGAGCCATGGGGGC
-5535  -----+---------+---------+---------+---------+---------+----   -5476

       TGCGGCAGGAGTCGAAGCCTCGGAAGCGGCTGGATTCGGCACGATGGCACCAGATTTGGC
-5475  -----+---------+---------+---------+---------+---------+----   -5416

       CTTCTTGGCCGGAGGCTCGATGAAGCCATCTTCACTGCAAGAACAAAAAACTAGCGAAGT
-5415  -----+---------+---------+---------+---------+---------+----   -5356

       CAGAATTCAATGCATATGGCGAAGTTAGAACACAATCCTGGAAAAGGAAGCAAGGACTTA
-5355  -----+---------+---------+---------+---------+---------+----   -5296

       CAATTCATAGAGACCATCTTCATCGGCAAAGCCGCCGGATGATCTCTTTGGAGGTAGTGC
-5295  -----+---------+---------+---------+---------+---------+----   -5236

       CTCGGCCTTTTCCGTAGCTGCATCAACAAAGGCAGCACGATCAGCATCGTCATCATGCAT
-5235  -----+---------+---------+---------+---------+---------+----   -5176

       TGACCCCGCTGTATCGCTCATATCATCGGCAGAGAATCGAGGATTGATGGAAAAAGCCTC
-5175  -----+---------+---------+---------+---------+---------+----   -5116

       AGGATTCATCGGATCATCATGTTGATCTATCGGGCTTGCATTCCCTAGAGTATGGGACCC
-5115  -----+---------+---------+---------+---------+---------+----   -5056

       TACAAGGACTAAGGAATCCCTTTTCTTGGAAAAATTGTTCGACAGGTCTTGCAAACGTTC
-5055  -----+---------+---------+---------+---------+---------+----   -4996

       AAGAGCCGTAAGGATCTGTCGTAGTTGACGAGTGAGAATAATGGCAGTTAAAATAATCAA
-4995  -----+---------+---------+---------+---------+---------+----   -4936

       AGGAACATGACAATAAGAGCATAAAGGGGAAATTTACCTCGGTTGGCAGATGACCAGCGT
-4935  -----+---------+---------+---------+---------+---------+----   -4876

       CAAATGGCGGTTGAGGAGATATCAGTGGAATTGAATCTTCCTGGCTAAAGAGGGTGAGAC
-4875  -----+---------+---------+---------+---------+---------+----   -4816

       ACCGGACTTCGTCAAGCAGTTCTTTTTCGGATAATTCAGCAATATTTACTCTAGTCTCGT
-4815  -----+---------+---------+---------+---------+---------+----   -4756

       CCCTGGGACCCGAATACAACCACATCGGATGGGTCCTAGACATGATCGGCTGAACTCGAT
-4755  -----+---------+---------+---------+---------+---------+----   -4696

       GTTTTAAGAACACAGCGGCTACCTCAGTACCTATCATGGTTTGACCATCGGATTCTTTGA
-4695  -----+---------+---------+---------+---------+---------+----   -4636

       TCCGAAGGAATCTATCAAATAACTTGTCTACTGTTGGTTTTTCATCGGGTGAGAGGATAT
-4635  -----+---------+---------+---------+---------+---------+----   -4576

       TTTTCCAAGACTTCTTGGGCTTGCTTCTAGAACATCGGAGAATTGGGGGGAGCTGGGAG
-4575  -----+---------+---------+---------+---------+---------+----   -4516

       TCGGCTGCTGATGAGTCCTTAATATAAAACCACTTCAGCCTCCAGCCTTGCACGGATTCT
-4515  -----+---------+---------+---------+---------+---------+----   -4456

       TTCATCGGGAAGTTGAAGTAGTTGACTTCCTTACGAGCAACAAAACCAACCCCACCAATG
-4455  -----+---------+---------+---------+---------+---------+----   -4396
```

FIGURE 39 CONTINUED

```
        ACGAAGGACCCACCACTGCTGTTATATCTTTTCACGAAGAAAATCTTCTTCCACAAACCA
-4395   -----+---------+---------+---------+---------+---------+----   -4336

        AAGTGGGGCTCAATGCCCAAAAACGCTTCGCAGAGGGTGATAAAGATGGCAAGGTGAAGG
-4335   -----+---------+---------+---------+---------+---------+----   -4276

        ATTGAGTTGGGGGTTAACTTCCATAATTGAATCTCATACACTCGAAGGAGGTGGTGAAGA
-4275   -----+---------+---------+---------+---------+---------+----   -4216

        AATTTGTGAGCGGGAAGCGAAAGACCTCGGTACAAGAAGGATAAGAACATCACAGTAAAA
-4215   -----+---------+---------+---------+---------+---------+----   -4156

        CCGGCAGGAGGATTGGGCCGTGAAATTGCACCTGGAAGAATAACATTCCCCTCGTCAGAA
-4155   -----+---------+---------+---------+---------+---------+----   -4096

        GAAATTATTACGAGGCTCCGGGCCCTCTTTTCATCTCGCTTCGTGGTGGTAGAAGCTGGC
-4095   -----+---------+---------+---------+---------+---------+----   -4036

        CAATCGCCAGGGATAGGCCCGGCCGTGGAGCTTGACGGCGCTGGCGGTGCCGGAGCTGAG
-4035   -----+---------+---------+---------+---------+---------+----   -3976

        GGAGGAGCATCTGGCGCGCTTCTCCGCGGCGGATTCGAAGGAGCCCTGACGGTGGTGCCA
-3975   -----+---------+---------+---------+---------+---------+----   -3916

        CTGCTCACGGCGCTGGTGGCGAGAGTGGGATTCTTCTTCTTCACCATTGTGAGATTTGAG
-3915   -----+---------+---------+---------+---------+---------+----   -3856

        GGAGATCTGGGAGTTGCGACGGTGGCGTGGTAGTTGCAAACGAAAAGGATGAATGAGGAA
-3855   -----+---------+---------+---------+---------+---------+----   -3796

        GAAGGGACGCAAGGATGAAGTGTGGAAAGGGGAGTTTACCCCAAGAGATTATAAAGTGAA
-3795   -----+---------+---------+---------+---------+---------+----   -3736

        AGGAAAACCTGAGAATTGAGCGGGCACGTGTCGTTGCTCTCAATTTATTGAGGGGATTTT
-3735   -----+---------+---------+---------+---------+---------+----   -3676

        TTCTCATCATAGATCGCGGAAATCGAGGAGTCACCTTGGTAACTGCACGCAAGTAGTGGT
-3675   -----+---------+---------+---------+---------+---------+----   -3616

        CATTTCTTAAACAGAACCGCATAGAAGTAGGATGGGACCGTCAGGTCACGTCCTATCAGT
-3615   -----+---------+---------+---------+---------+---------+----   -3556

        CAGATTTACAACAGTAATTACATCATCACTGACGTCAAAGTATGCTTGAAGTATCCGAAG
-3555   -----+---------+---------+---------+---------+---------+----   -3496

        AAAAGTCGAAATTTGGGCTCGAAGACTTTCTTGCAGAGAAGCGCGTGAAAGGAATATCTA
-3495   -----+---------+---------+---------+---------+---------+----   -3436

        AGGAAAGGGTCAAAACATTCGGCTCGAGTCTACGCACGGATTGCAAGCATCCGTACCTAG
-3435   -----+---------+---------+---------+---------+---------+----   -3376

        ACTCGGGGGCTACTCCCATCGGGAGCGCTGGACGTGCACCCGATAAATTTAGACGAGGAT
-3375   -----+---------+---------+---------+---------+---------+----   -3316

        GAAAACCGGAAACCCAAGTGCTACTCCCATCGGGAGCGCCGATTACGCACCCGACAAACT
-3315   -----+---------+---------+---------+---------+---------+----   -3256

        TTTTTGCACTCCAGGATCATGCCCGGGGACTTAATTCTGTGTAGAGTAGCGTTGTTTTGT
-3255   -----+---------+---------+---------+---------+---------+----   -3196
```

FIGURE 39 CONTINUED

```
       CTTCGGCAGTTAACCAGCAAAGCTGGACACGTTACTCAATATCCTTTACGCATTAAACCC
-3195  -----+---------+---------+---------+---------+---------+----   -3136

       TTACTTGAAGAATTGAAGCCCCGATGCAAATATATCGGATGACCTATGAAGGCCTGCGGA
-3135  -----+---------+---------+---------+---------+---------+----   -3076

       AAGCTTCGGGAGAAGAAGACATTCGAGTGGCACAACTTGAGTCTACGAACGGATTGCAAG
-3075  -----+---------+---------+---------+---------+---------+----   -3016

       CATCCGTACCTAGACTCGGGGGCTACTCCCATCGGGAGCGCTGGACTCGCACCCGATAGA
-3015  -----+---------+---------+---------+---------+---------+----   -2956

       AGGAGATGATGATATTACAAGAAGGACAAGAAGTATCAAGGGAGAAGAACATTCGGTGGA
-2955  -----+---------+---------+---------+---------+---------+----   -2896

       GGCATGCTTTAGTCTCTACCCGAAAAAACTTCGGCTAGACACTCGGGGGGCTACTGACGT
-2895  -----+---------+---------+---------+---------+---------+----   -2836

       GGGCATTACCCTTCGGGTAACTGATATTGCCCTATCCTGTACGACCCAACTGGAGGCCCA
-2835  -----+---------+---------+---------+---------+---------+----   -2776

       TGAAGACACTCGAAGGCAAGGTGGACCACTACGTCGGTGCCGAAGGGGGTTCCTTGAAGA
-2775  -----+---------+---------+---------+---------+---------+----   -2716

       ACAAGACGAAGAAAAGAAGAATACAAGAAAAGTATAGAACTAGGATCTTTTGTAACCTGG
-2715  -----+---------+---------+---------+---------+---------+----   -2656

       TCGTACCCGGACAGATCTCTCGAGACCTGGCCCCCTACATATGGGCTAGGAGAGGGGCTG
-2655  -----+---------+---------+---------+---------+---------+----   -2596

       CCGAGAGGGACACACACAATCTTAGCAATTTTAGCCACCATAAGTCCAGAGCAAGGTCCC
-2595  -----+---------+---------+---------+---------+---------+----   -2536

       CGTAGAACTTAGCCTCTCGACGAGATCACAGCCGAAACCTTCGGCACCCCATTGTAACCC
-2535  -----+---------+---------+---------+---------+---------+----   -2476

       GATATTTTCATAGTCAAGATCAGACAGGTAGGACGTAAGGGTTTTACCTCATCGAGGGCC
-2475  -----+---------+---------+---------+---------+---------+----   -2416

       CCGAACCTGGGTAAATCGCTCTCCCCGCTTGTTTGATAACCGATGGCTTGTGTCAGCTTA
-2415  -----+---------+---------+---------+---------+---------+----   -2356

       CATGATTCCATCTACCCTAAACCTCAAACGGAGGGCATTGCCGAGGAGTACCCTCGACAT
-2355  -----+---------+---------+---------+---------+---------+----   -2296

       TCCCCTCCACCAATGGTCTCACATAAATTCAACAAAGCAAACTCATAAAAAGTTTAATGA
-2295  -----+---------+---------+---------+---------+---------+----   -2236

       GTTTCAGAAAGAAATAAAACTAGGCCCCTCCTTTGAGAATCTACGAATGATTCACCATAT
-2235  -----+---------+---------+---------+---------+---------+----   -2176

       CATCTCGCAGTTAGTGATGAGTAACTAAGTCTCAAATTTCCCGACGCATGGCGAAAAAGG
-2175  -----+---------+---------+---------+---------+---------+----   -2116

       TAGCGAACTTAAAATGTGAGGAATGAATGCCACATATGCATGGTGCATCGAGTATTCTCA
-2115  -----+---------+---------+---------+---------+---------+----   -2056

       TTTTAGTCTTGGATTACTCCCTTTAGATGTTGACACCATCCCAAAAATACAACTTGGACA
-2055  -----+---------+---------+---------+---------+---------+----   -1996
```

FIGURE 39 CONTINUED

```
        AGTTGTTCATTTCACTAGTATGAATTTCAGTAAATCGGGCAATACTCCAACACTCATTCA
-1995   -----+---------+---------+---------+---------+---------+----   -1936

        CCCCCTAGGCGAGGTTAGCTCAGATCAACGTCGGGTGTCTTCATCGAGTTAATGTCGTCA
-1935   -----+---------+---------+---------+---------+---------+----   -1876

        CACGCACACACACGTACGCGCACACACACGTGCGCAAACAAAAAGAAAACTAGGAACCTT
-1875   -----+---------+---------+---------+---------+---------+----   -1816

        CTCACGTAGCCTAGGTCTTGTCCTGTAAGAAAAAACCCAGGTCCACCCTAGTTTCGAACC
-1815   -----+---------+---------+---------+---------+---------+----   -1756

        AAAATATTTTTGAAGATACATTAGTAAGATATTTTTGAAAATAAAACCGCAAAAAGGGAA
-1755   -----+---------+---------+---------+---------+---------+----   -1696

        TTGAAAAATATGGACTGGCTGTTTTGTCCAAAACCACATCTTTCGGAGAACCACGAGGGT
-1695   -----+---------+---------+---------+---------+---------+----   -1636

        ATCTATTGATGGGCTCATACTATACCTGGGCATGTGTTGGGCCAGGCCTCATGTCGGGCC
-1635   -----+---------+---------+---------+---------+---------+----   -1576

        GAGGAAAGCCCGACGCTGAAAAATCAGGCCCAAGCTTAACCCGGCCCGACCAAATACCCA
-1575   -----+---------+---------+---------+---------+---------+----   -1516

        CCAAACCCGTTGGGCCATCAGGTTGCGGGCCGGGCAGTAGTGTAAAACACCGATTTCGGG
-1515   -----+---------+---------+---------+---------+---------+----   -1456

        CTACATAGGCCCGGCTCGTTTGTCGGGCAAACATTTCTAGACCTAAGCCCGAGTTTTTCG
-1455   -----+---------+---------+---------+---------+---------+----   -1396

        GGCCGGGCTGCCCATGGCCAGGTATAGCTCATAACGACGTATGACATTTCGAGCAATTGA
-1395   -----+---------+---------+---------+---------+---------+----   -1336

        TGCAAAGCACGTGTAGGGTTTTATCCCATCCGTGTGGCGTGTGTAGGGTGTAAATGAATA
-1335   -----+---------+---------+---------+---------+---------+----   -1276

        GGATAATTTCCTCGCCGAAACTGGTCCCAAATTCGCTTTGAAGTGTCCATATATGATTTT
-1275   -----+---------+---------+---------+---------+---------+----   -1216

        AAAGAATGTGACAAATAAAGATATCCAATTTCGAAATAGTGCTCCGGATACGGTATAGGA
-1215   -----+---------+---------+---------+---------+---------+----   -1156

        TATGGTATAGCAAATAACATGCTGATATGGATTGTCCGATATTAAATTAAGATAATCCAA
-1155   -----+---------+---------+---------+---------+---------+----   -1096

        ATGTTTTAAACCGCATAATTCGATTTTTGAGTCAAAAGCGAATGCCAATTCAGAAGGTTA
-1095   -----+---------+---------+---------+---------+---------+----   -1036

        GCAGTTATTGAGTTTCAAAATTTATTTGGCGAGCATATCTAGTTCTAAATTCTATCACGT
-1035   -----+---------+---------+---------+---------+---------+----   -976

        AAATTGTGTCTTTTTTTAATAACTACACAAGACTAAAAGTTTAAATCTCTCTCAAGATTT
-975    -----+---------+---------+---------+---------+---------+----   -916

        GCGAAAACTATAGCTATCTACTGATATATATATCCGACTATATTTGTTTTCGGACCGCAT
-915    -----+---------+---------+---------+---------+---------+----   -856

        GCGTCCTATTTCCGATTCGAATCTGCACTCCGATATATCCACATTGAATCTAAAACCGAT
-855    -----+---------+---------+---------+---------+---------+----   -796
```

FIGURE 39 CONTINUED

```
      CAATATTTGCTCCGATCTAAATCCGGAAAAATATGTGGTGAAGGATATGGTATAAGCAAA
-795  -----+---------+---------+---------+---------+---------+----  -736

      ATCCGATTTGATCCATTTGTACCTCTAGGCGTGTGCAAGACCTGGAGGAAAGAATGGCGC
-735  -----+---------+---------+---------+---------+---------+----  -676

      ATCTGTAGGGTGCAGTCCCACCGGTGGAAAATGTGAGCTCACCGTATTGTCCCCCGATGG
-675  -----+---------+---------+---------+---------+---------+----  -616

      AGCATCGAAACGGAGTCGGAACACGATTTGCGCCACGTACAGAGCATGCATGATTTCCCT
-615  -----+---------+---------+---------+---------+---------+----  -556

      TGTATGCGGTCCAGGATCTTAAACTGCCTTCCATTTCCAGGAACCTACCGATTGGCTGCA
-555  -----+---------+---------+---------+---------+---------+----  -496

      AGCCGTAGCTAGCGGTTTGAAGTCACGGCATTGCCGCCCCCGATTAACCCACCCGTCGCG
-495  -----+---------+---------+---------+---------+---------+----  -436

      CGCGCGGTCGGTCGTTTCACCGTCCTGCCTAGGCTACGCACGCGCGCGCGCAGTTGGGCC
-435  -----+---------+---------+---------+---------+---------+----  -376

      AGTTGTAGGTAAGCCGACTCGAGATCACACACCCGGCCTCACCTACTACCTCTCGCCGTC
-375  -----+---------+---------+---------+---------+---------+----  -316

      GCGGTCACCGTGTCACACTCACGCCCAGGGGAGCCACCCGCCCACACGGCGCCTAGCTCA
-315  -----+---------+---------+---------+---------+---------+----  -256

      TCCCCTCTCACTACTCTTCTTCTCCTCCCTCTCACCTCGCCGTCGACCCAGCTCCCGGCT
-255  -----+---------+---------+---------+---------+---------+----  -196

      CTATAAATTCCGCACTACTCGAACCAACATCGCCCAGGCCTTTGCCTTTTACGACGAATC
-195  -----+---------+---------+---------+---------+---------+----  -136

      CTACCAAACCGAGCTACCAGATCCTTCTCTACTAATCGAGCTCCCTACGCTGCTCCGCCT
-135  -----+---------+---------+---------+---------+---------+----  -76

      GTCTTCGTTTCCGCCTCACCGCCGGCCGGTTCTCCGCTCCAAGCTACGTCCGTCCGTCCA
 -75  -----+---------+---------+---------+---------+---------+----  -16

      CATATATAGCATCGACATGACCATCGCCGAGGTCGTGGCTGCCGGAGACACCGCCGCCGC
 -15  -----+---------+---------+---------+---------+---------+----  44
                      M  T  I  A  E  V  V  A  A  G  D  T  A  A  A

      GGTGGTGCAGCCCGCCGGGAACGGGCAGACCGTGTGCGTGACCGGCGCCGCCGGGTACAT
 45   -----+---------+---------+---------+---------+---------+----  104
       V  V  Q  P  A  G  N  G  Q  T  V  C  V  T  G  A  A  G  Y  I

      CGCGTCGTGGCTCGTCAAGCTGCTGCTGGAGAAGGGGTACACCGTCAAGGGCACCGTCAG
105   -----+---------+---------+---------+---------+---------+----  164
       A  S  W  L  V  K  L  L  L  E  K  G  Y  T  V  K  G  T  V  R

      GAACCCAGGCATGTCACCCATGCATTCATCATTTTCTTACTAGTCGTATGCGTTATGCGA
165   -----+---------+---------+---------+---------+---------+----  224
       N  P  G

      CTTGTGTATTAACTATTGTGGACTGCATGCAGACGACCCGAAGAACGCGCACCTGAGGGC
225   -----+---------+---------+---------+---------+---------+----  284
                                          D  P  K  N  A  H  L  R  A
```

FIGURE 39 CONTINUED

```
     GCTCGACGGCGCCGCCGACCGGCTGGTCCTCTGCAAGGCCGACCTCCTCGACTACGACGC
285  -----+---------+---------+---------+---------+---------+----  344
      L  D  G  A  A  D  R  L  V  L  C  K  A  D  L  L  D  Y  D  A

     CATCCGCCGCGCCATCGACGGCTGCCACGGCGTCTTCCACACCGCGTCCCCCGTCACCGA
345  -----+---------+---------+---------+---------+---------+----  404
      I  R  R  A  I  D  G  C  H  G  V  F  H  T  A  S  P  V  T  D

     CGACCCCGTACGTACTCCATAGAACTCGGCACCCCTAGCTTCTCTCCGTTCTCTCTGTAT
405  -----+---------+---------+---------+---------+---------+----  464
      D  P

     GTCTGTCACCGTCGATCGCCATGGCAGCACGCATGCATGCGCGCGCAACGCTAGCTAGAC
465  -----+---------+---------+---------+---------+---------+----  524

     GCTGACCGACTCATTGTGCAGGAGCAAATGGTGGAGCCGGCGGTGAGGGGCACGCAGTAC
525  -----+---------+---------+---------+---------+---------+----  584
                          E  Q  M  V  E  P  A  V  R  G  T  Q  Y

     GTCATAGACGCGGCGGCGGAGGCCGGCACGGTGCGGCGGATGGTGCTCACCTCCTCCATC
585  -----+---------+---------+---------+---------+---------+----  644
     V  I  D  A  A  A  E  A  G  T  V  R  R  M  V  L  T  S  S  I

     GGCGCCGTCACCATGGACCCCAACCGCGGGCCGGACGTGGTCGTCGACGAGTCGTGCTGG
645  -----+---------+---------+---------+---------+---------+----  704
     G  A  V  T  M  D  P  N  R  G  P  D  V  V  V  D  E  S  C  W

     AGCGACCTCGACTTCTGCAAGAAAACCAGGGTGGGTGCTGCATGCTCAATTTTTATTATC
705  -----+---------+---------+---------+---------+---------+----  764
     S  D  L  D  F  C  K  K  T  R

     ATAGCTACCCTTTTTCTGCACCATGCTGCATTTCTTTTCCAAAAACAACTCTCAAAAGAT
765  -----+---------+---------+---------+---------+---------+----  824

     ATGCTACGTGGTGAGTTCCTATAGCTGAATTATTACAACTACCACCCTATCGATCACTAC
825  -----+---------+---------+---------+---------+---------+----  884

     CGCCCTAAAAGTGTTCAACTTTTGAAGGCAACCAAAACCAATACATGAACGACGATCGTG
885  -----+---------+---------+---------+---------+---------+----  944

     TGCGCTTGTCGTCGTTATCATTAGCCTCTGTAGCTCTAATTTTCACCTATGTACGCATGG
945  -----+---------+---------+---------+---------+---------+----  1004

     ATAGACGATTCGGAAATACAGTTCAGTTTACCTACCATATACTATGCCGAAATCGAACGC
1005 -----+---------+---------+---------+---------+---------+----  1064

     ACACAGGTGTGAGGCAGCAGCCGCTCACGAGTTATGCGCCGAAACCGACATCTCGGAATC
1065 -----+---------+---------+---------+---------+---------+----  1124

     TTCAGTCCACAATCAAAAAATAGACACCTGGTACCACTACAAAATTATACTCCTACTGTA
1125 -----+---------+---------+---------+---------+---------+----  1184

     TATTGGTAAAACAAAACATTTTCTTTTTTATTTGATAGGAGTGCTGCAAATTAAAGTTCT
1185 -----+---------+---------+---------+---------+---------+----  1244

     TTGTGTCATTTTTCAAAGGAAAAAAAAACACCTTTACCACTCTTCTTCCTTGCCATCAT
1245 -----+---------+---------+---------+---------+---------+----  1304
```

FIGURE 39 CONTINUED

```
      TTTTTTTTTACCAAAGTTTGTTCTGTCAAATGAACATATATATAGTTCGGTGCTATGTCA
1305  -----+---------+---------+---------+---------+---------+----   1364

      GTGCCATTTACCGGCCACTAGCTAGTAGGACTGCCATGTTCCAGCAAATTGTCTAGTGGA
1365  -----+---------+---------+---------+---------+---------+----   1424

      CCGGAGTGGCCAAAAGGAGCCAATTATGTAGGGTTGCAAGCGGGATCACACAAAAGCCTC
1425  -----+---------+---------+---------+---------+---------+----   1484

      GCCTCTAGTTCATTTTATCAATTAAGTGGTACTTTCTCAGGGACCCCCCTTGCAACTCTA
1485  -----+---------+---------+---------+---------+---------+----   1544

      CCATTACATCCGTGCAAAATAAAAGCTAGCATCACGCACCAGATTTAGTACTCCCTCCGT
1545  -----+---------+---------+---------+---------+---------+----   1604

      TTTTATTTAGTTCGCATTCTAGGTTCAGCCAAAGTCATACTTTGCAAAGTTTAACCAAAA
1605  -----+---------+---------+---------+---------+---------+----   1664

      TTATAAGAAAAAAATATCAATAATCATCATACAAAATACATATAATATAAGAGTAAACCT
1665  -----+---------+---------+---------+---------+---------+----   1724

      TATAACGATTCTACAATAGATTTTTTTATTGCATATGTCAATATTTTTTCATAAATATTT
1725  -----+---------+---------+---------+---------+---------+----   1784

      ACTCAAAATTATAAGGTTTGACTTTGACTAAACCCAGAACCTTCTTAGAGAGGAAGAAAT
1785  -----+---------+---------+---------+---------+---------+----   1844

      GCATGGGCAAAAGCAAATCATGCATATGGGCAGGAGTAACATTTTTTTGACTTTCATAGA
1845  -----+---------+---------+---------+---------+---------+----   1904

      AAGTACTGTATGGCACTAAACGGTCTAAACCGGACACTGGAAGCAAATCGTGCACGTGGG
1905  -----+---------+---------+---------+---------+---------+----   1964

      CAATATTATCTACCGTCGCGTCGCCAGTCTCCCCATGCCCATGACCATGCTTGGAATTTT
1965  -----+---------+---------+---------+---------+---------+----   2024

      AGTCTCGCCGGAGCTGCCGAGTGCATGCATAGTGACGAGTTTCAATAGGCCACTATATAT
2025  -----+---------+---------+---------+---------+---------+----   2084

      GTGATCATGGCTCTTGATTTGTCACTTTCTTTTTTTGCCGAAGGATATAGTAGTATTACT
2085  -----+---------+---------+---------+---------+---------+----   2144

      TTCTCTGCTATCACAAAGAAAGAACTGATTGTGTCTAGTCTAGGTGGTCTCAGAATTCTG
2145  -----+---------+---------+---------+---------+---------+----   2204

      CATGACTCCAGAGTATTCTTGATGCCACTTGTTTGTTATTGCAAGAAACTTAATTCGGAG
2205  -----+---------+---------+---------+---------+---------+----   2264

      ACAACCAAAAGCTCATCCCATGTCTCTGGAACTAGTAGACATAAGAAAATCTCATGGTAT
2265  -----+---------+---------+---------+---------+---------+----   2324

      CAGTTTGCTATTTATCTACAACTGAAACGGCATGTTTGGTTTTATTAAATTCAGAACTGG
2325  -----+---------+---------+---------+---------+---------+----   2384
                                                          N  W

      TACTGCTACGGGAAGGCGGTTGCGGAGCAGGCGGCATCGGAGTTGGCGCGGCAGCGCGGC
2385  -----+---------+---------+---------+---------+---------+----   2444
      Y  C  Y  G  K  A  V  A  E  Q  A  A  S  E  L  A  R  Q  R  G
```

FIGURE 39 CONTINUED

```
     GTGGACCTTGTGGTGGTGAACCCGGTGCTGGTGATCGGCCCCCTGCTGCAGCCGACGGTG
2445 -----+---------+---------+---------+---------+---------+----  2504
     V  D  L  V  V  V  N  P  V  L  V  I  G  P  L  L  Q  P  T  V

     AACGCCAGCATCGGCCACATCCTCAAGTACCTGGACGGGTCGGCCAGCAAGTTCGCCAAC
2505 -----+---------+---------+---------+---------+---------+----  2564
     N  A  S  I  G  H  I  L  K  Y  L  D  G  S  A  S  K  F  A  N

     GCCGTGCAGGCGTACGTGGACGTCCGCGACGTGGCCGACGCCCACCTCCGCGTCTTCGAG
2565 -----+---------+---------+---------+---------+---------+----  2624
     A  V  Q  A  Y  V  D  V  R  D  V  A  D  A  H  L  R  V  F  E

     TGCGCCGCCGCGTCCGGCCGCCACCTCTGCGCCGAGCGCGTCCTCCACCGCGAGGACGTC
2625 -----+---------+---------+---------+---------+---------+----  2684
     C  A  A  A  S  G  R  H  L  C  A  E  R  V  L  H  R  E  D  V

     GTGCGCATCCTCGCCAAGCTCTTCCCCGAGTACCCCGTCCCCACCAGGTACGCGTACGAC
2685 -----+---------+---------+---------+---------+---------+----  2744
     V  R  I  L  A  K  L  F  P  E  Y  P  V  P  T  R

     CTGCTTGCTAGCCGCTTCCGTTAATTCCATTGCCTTAATTGATTGCATGATGCCGCTCCT
2745 -----+---------+---------+---------+---------+---------+----  2804

     AATTTACTCACTTGCGTAACTAATTGCATTCATATATGATCTACCAACCGTGGAGAAAAT
2805 -----+---------+---------+---------+---------+---------+----  2864

     TAGCAAGAGTCTGTCGGGGCGTCCCGGTCCAGTGCAGTTAACCTGCATGTCGATGGTCTG
2865 -----+---------+---------+---------+---------+---------+----  2924

     CAGGTTGCAGCTTACTTGTGGTTCTTTAGTTCAGAGACACAGAGCAATTGGGCACTAAGC
2925 -----+---------+---------+---------+---------+---------+----  2984

     AAAACTGACATCACTGGTAATTAGGTAGCTCCCACACACTGAAGTGGGTGGATCCCATCG
2985 -----+---------+---------+---------+---------+---------+----  3044

     GTAGTAGGTAAGGGTGGATAGTACTGGACGAGAGCTCGATCGTTGTTGTAAAAAAGCGAG
3045 -----+---------+---------+---------+---------+---------+----  3104

     TGACCACCACTTCACCATCCACTGCAAGTAGCTGCTAGTGAACCATCCAACCAGCTCCCT
3105 -----+---------+---------+---------+---------+---------+----  3164

     GGATCACTCTGCTCCGTCCGTACCTTCAGCTACCTACAGAAGCGACATGAACACACAGAC
3165 -----+---------+---------+---------+---------+---------+----  3224

     ACACAAGGCCGGCTCACCATTCGCATAGGTCAAACCAAATGTTGGTGAACGGCAACATCG
3225 -----+---------+---------+---------+---------+---------+----  3284

     CCACAAGTCGCGTGCTAGTTCGAGGTTGTGTCCGGTGTACCGAGGCCACACTATTCGTGC
3285 -----+---------+---------+---------+---------+---------+----  3344

     TGCCCGTCGCTGATATTTGCACGCGTAGCTGTCGACGAAAGTAGGTGGACTGACAGATAC
3345 -----+---------+---------+---------+---------+---------+----  3404

     ACATATCCTCATTGCCTTCTCTGCTCGGTTTCTGCTAGGATTGCCATCTTCAGGAGTGCC
3405 -----+---------+---------+---------+---------+---------+----  3464

     TATCCGCACGGCAGAAACGCGTAGCATCAGGCCAGAAAGCAGCGTGCGTGATATCGTAAC
3465 -----+---------+---------+---------+---------+---------+----  3524
```

FIGURE 39 CONTINUED

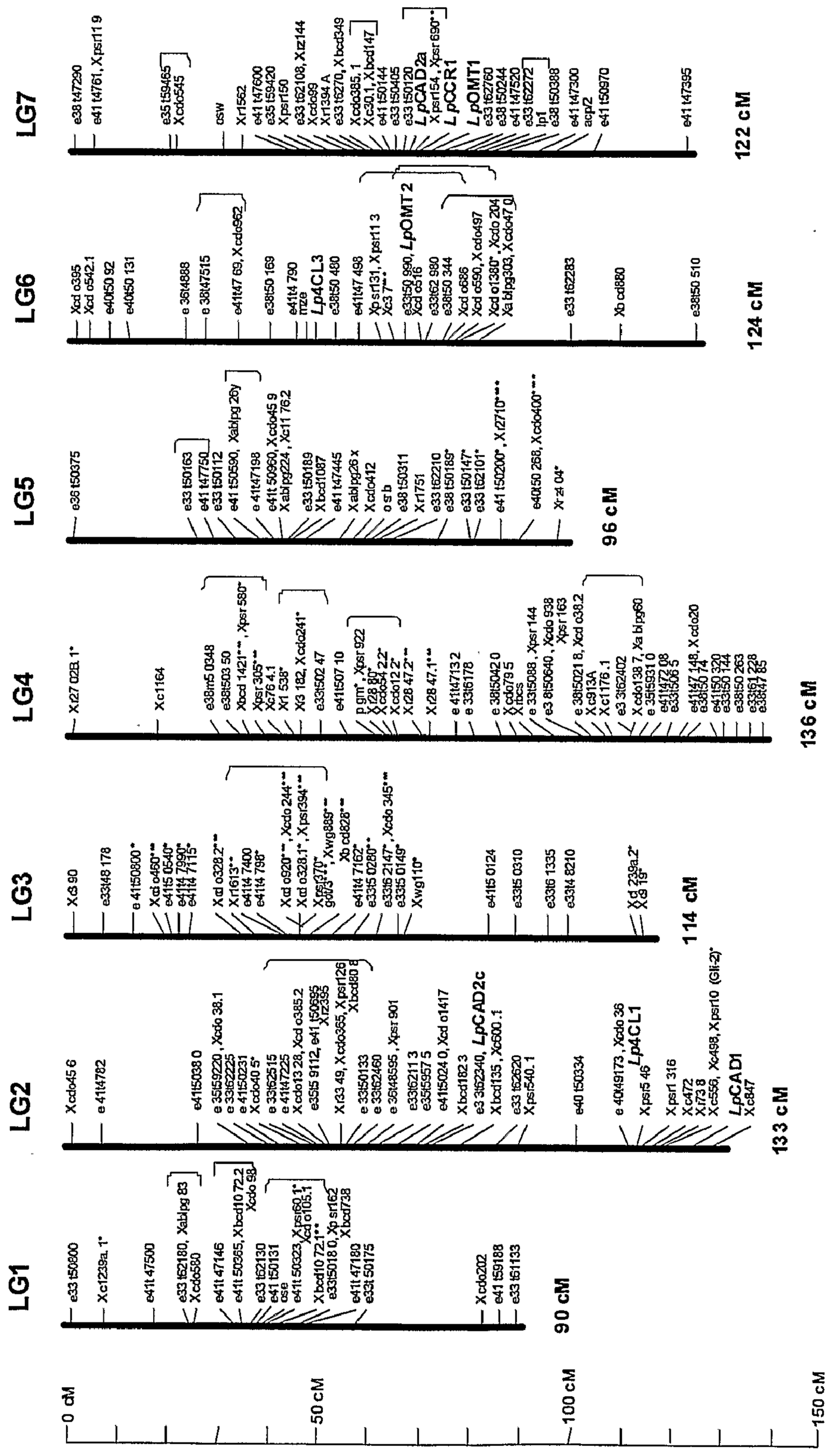
```
     CCAGACGGTCTTCACCTGTCCATTCTGGGCTACCTGGCATACTACCTCGGTGCCGCTGTG
3525 -----+---------+---------+---------+---------+---------+----  3584

     CCGCTGACCAATTCGTGCACGACCACTATAGCAAAACCCTATGCATGTAACTGCTTCAAG
3585 -----+---------+---------+---------+---------+---------+----  3644

     ATCAGCAGTGACATGTGCAATATAAACCTCAAGTGTGCACTCTAGTGCGTACTGATAAAA
3645 -----+---------+---------+---------+---------+---------+----  3704

     CCGTATAACTGGTGACCCAGTCATTCTTCTCTTTTTTATTTGTTTGGACCAAACGAACAC
3705 -----+---------+---------+---------+---------+---------+----  3764

     AGCATGTTATCCATCACCAACAAGTGGCGCTGATTTTTCAAACTACACTGGGATCATACT
3765 -----+---------+---------+---------+---------+---------+----  3824

     GGAAACCAAAGCAGGAGAACATCTTCGAACCAAGAGATGTTTACTAAATTTGAAAGAAAA
3825 -----+---------+---------+---------+---------+---------+----  3884

     TGTACTGACAAGTAATCTGTCTGAAGCAAGACACATACTACCTCGGTTCGAACGTGGGAC
3885 -----+---------+---------+---------+---------+---------+----  3944

     ACCATGCCCGTGCCATATTTGCTAGGCACCACTCTGCCGTCGATTGTATCCCAACGGAGG
3945 -----+---------+---------+---------+---------+---------+----  4004

     GAGTATCGATTTGCGCAAAGTTCCTACATACATAGCCGCTCAAGATATAATCTTACGACC
4005 -----+---------+---------+---------+---------+---------+----  4064

     TTCCGTCGAAATCGGTGATACGTCGCAACCTATAGCTAACTTGGCAGAGCATAAAATAAC
4065 -----+---------+---------+---------+---------+---------+----  4124

     TATCTAAGGTTGGGGTCTCCCTCTTTTCAATCAACCTTTCATACCGAATGATGGGAGTGT
4125 -----+---------+---------+---------+---------+---------+----  4184

     TTGTGAAAACATCTCTTGGTCGACTCAGCATTAGCGCCCTACCAATTTCTCTGTGGACAA
4185 -----+---------+---------+---------+---------+---------+----  4244

     TGCCACCTTAAATCGTTTTTTAGTCTTCATGATTTACTCCCCCTTATATCTGGCCGTAGT
4245 -----+---------+---------+---------+---------+---------+----  4304

     CCCTCTTTTCCATTTTTCTTGTCTGGTTTTAAGTCAAATTTAGACTACTAAAACAACAGC
4305 -----+---------+---------+---------+---------+---------+----  4364

     AAGATTTTATGGAAGGGAGGTAGTGCAAAACAGAAAGTCCGATCGAAATGCGTGCCAATT
4365 -----+---------+---------+---------+---------+---------+----  4424

     TGTCGTCGCGGCGGCCGGACTAAAATGGATCTGCATGTGCATACCGTTCGTCGGAGTATC
4425 -----+---------+---------+---------+---------+---------+----  4484

     CTGCGAACGGTCGTGTGTTTAGTCAACATTAATGTGAGGTTCATGTGATACTCTTGCTTG
4485 -----+---------+---------+---------+---------+---------+----  4544

     AAAGATACTACTACTGCTACCTCGTAGAACTGAATGAAAGTATGTGGGACTGTTCAGCTC
4545 -----+---------+---------+---------+---------+---------+----  4604

     TCTGCACATGTCAAATGTCGTTACTCATACCTTTCGTCAGAGCATCCTGCGACGCGCGCC
4605 -----+---------+---------+---------+---------+---------+----  4664

     GGTGCCGAAATTTCGCCGTGTGTTTAGTCAAGATCAACGTGAGGTTCATGCGGTACCCTA
4665 -----+---------+---------+---------+---------+---------+----  4724
```

FIGURE 39 CONTINUED

```
      TCTGGCTTCGAAGATACCAAGCAGACTGCGGCTAGATTGTCATTTTGATGTCGCAATCTT
4725  -----+---------+---------+---------+---------+---------+----   4784

      CACCAAACCTGCCCTTCCGGACCACAGCAGCAGTACGTAACAATGGTGTCATCGCCATGC
4785  -----+---------+---------+---------+---------+---------+----   4844

      GTTGCTCGTGTCCAAGGAAACGGAGGAATCTCGGCTTCCCACAAGTCACGCATCGATGTT
4845  -----+---------+---------+---------+---------+---------+----   4904

      CACACCTGAATTGGTCGACGTTTCTTCTTCTAGACTAGAAAAAGATTACAGAACAACGCA
4905  -----+---------+---------+---------+---------+---------+----   4964

      AGCTTCGTTCAAGTCCATACTTCTGTTCAGTATACTCCTGATGATTGCAGTTATATCAGC
4965  -----+---------+---------+---------+---------+---------+----   5024

      ATGTCTATTCTGAATTTTTGCACTTCTATTCAAAGGATGGGCTGGAATTGCTACTGACTT
5025  -----+---------+---------+---------+---------+---------+----   5084

      TGGTGTGATGTGTGTGGCACAGGTGCTCTGATGAGACGAACCCGAGGAAGCAGCCATACA
5085  -----+---------+---------+---------+---------+---------+----   5144
                                C  S  D  E  T  N  P  R  K  Q  P  Y  K

      AGATGTCGAACCAGAAGCTCCAGGACCTCGGACTCGAGTTCAGGCCGGTGAGCCAGTCCC
5145  -----+---------+---------+---------+---------+---------+----   5204
        M  S  N  Q  K  L  Q  D  L  G  L  E  F  R  P  V  S  Q  S  L

      TGTACGAGACGGTGAAGAGCCTCCAGGAGAAGGGCCACCTTCCGGTGCTCAGCGAGCAGG
5205  -----+---------+---------+---------+---------+---------+----   5264
        Y  E  T  V  K  S  L  Q  E  K  G  H  L  P  V  L  S  E  Q  A

      CAGAGGCGGACAAGGAAACCCTAGCTGCCGAGCTGCAGGCAGGGGTTACCATCCGAGCAT
5265  -----+---------+---------+---------+---------+---------+----   5324
        E  A  D  K  E  T  L  A  A  E  L  Q  A  G  V  T  I  R  A  *

      GAGGAACAAGAAATCAACCATGTCCATACTGCTACTGTCATGTAAACCAGCTGTTGAATG
5325  -----+---------+---------+---------+---------+---------+----   5384

      CCTAAAATCTAAGTTCTTGTAATACTGTGTTGTTTCATGTGGACTAGATTGATCG
5385  -----+---------+---------+---------+---------+---------   5439
```

FIGURE 39 CONTINUED

FIGURE 40


LG1
LG2
LG3
LG4
LG5
LG6
LG7
90 cM
133 cM
114 cM
136 cM
96 cM
124 cM
122 cM
0 cM
50 cM
100 cM
150 cM
LpCAD1
Lp4CL1
LpCAD2c
Lp4CL3
LpOMT2
LpCAD2a
LpCCR1
LpOMT1

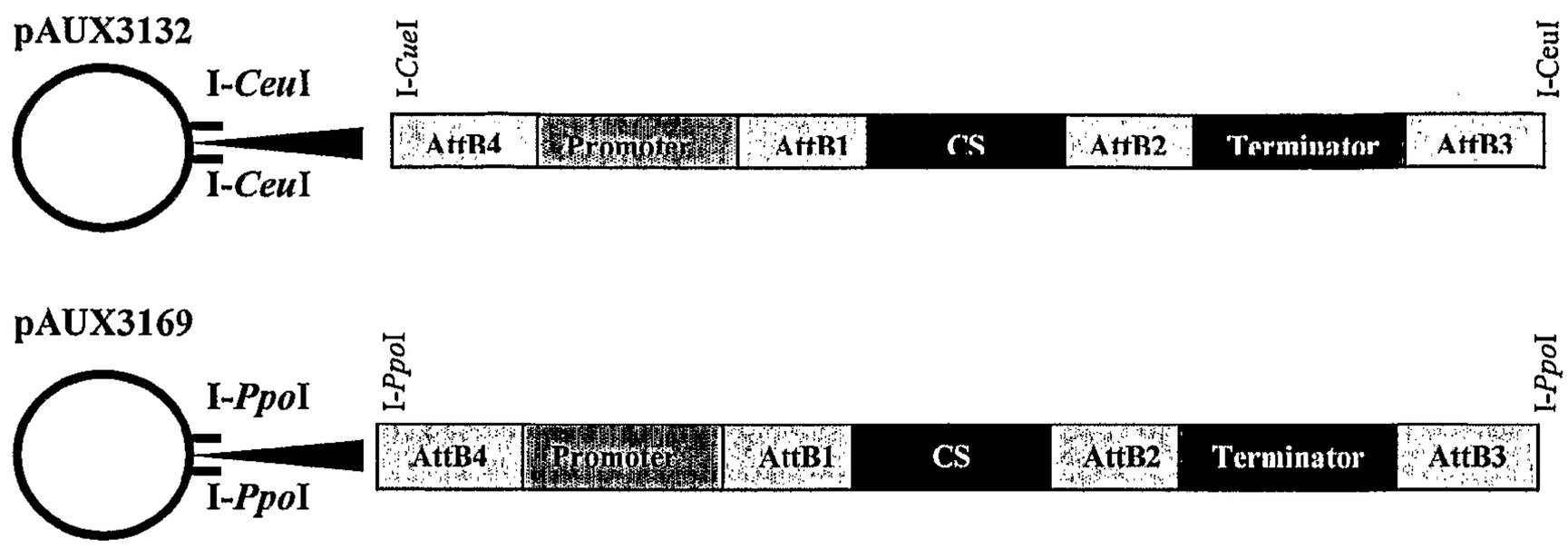

pAUX3132
I-CeuI
I-CeuI
I-CueI
I-CeuI
AttB4
Promoter
AttB1
CS
AttB2
Terminator
AttB3
pAUX3169
I-PpoI
I-PpoI
I-PpoI
I-PpoI
AttB4
Promoter
AttB1
CS
AttB2
Terminator
AttB3

FIGURE 41

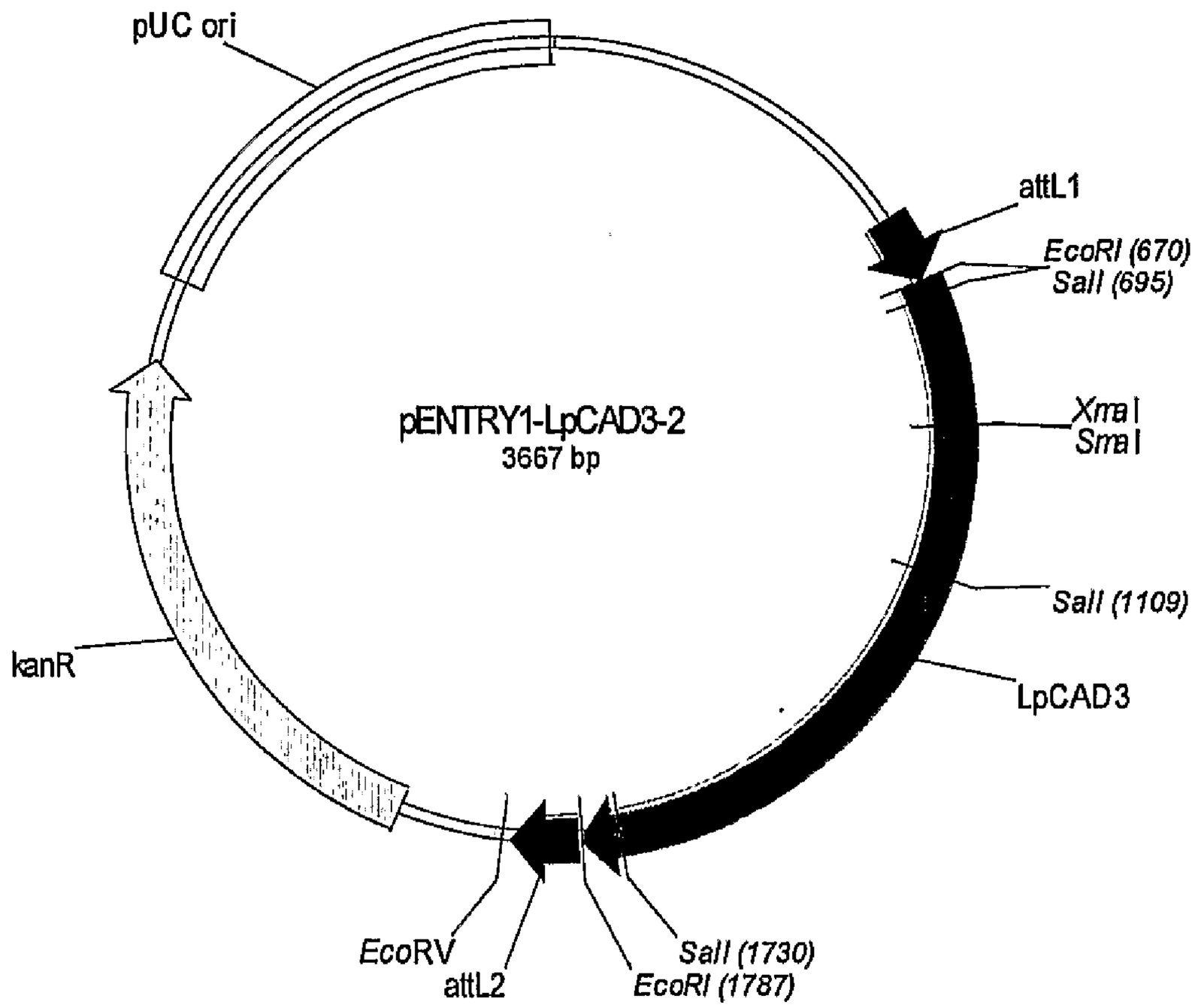

pUC ori
attL1
EcoRI (670)
SalI (695)
pENTRY1-LpCAD3-2
3667 bp
XmaI
SmaI
SalI (1109)
kanR
LpCAD3
EcoRV
attL2
SalI (1730)
EcoRI (1787)

FIGURE 42

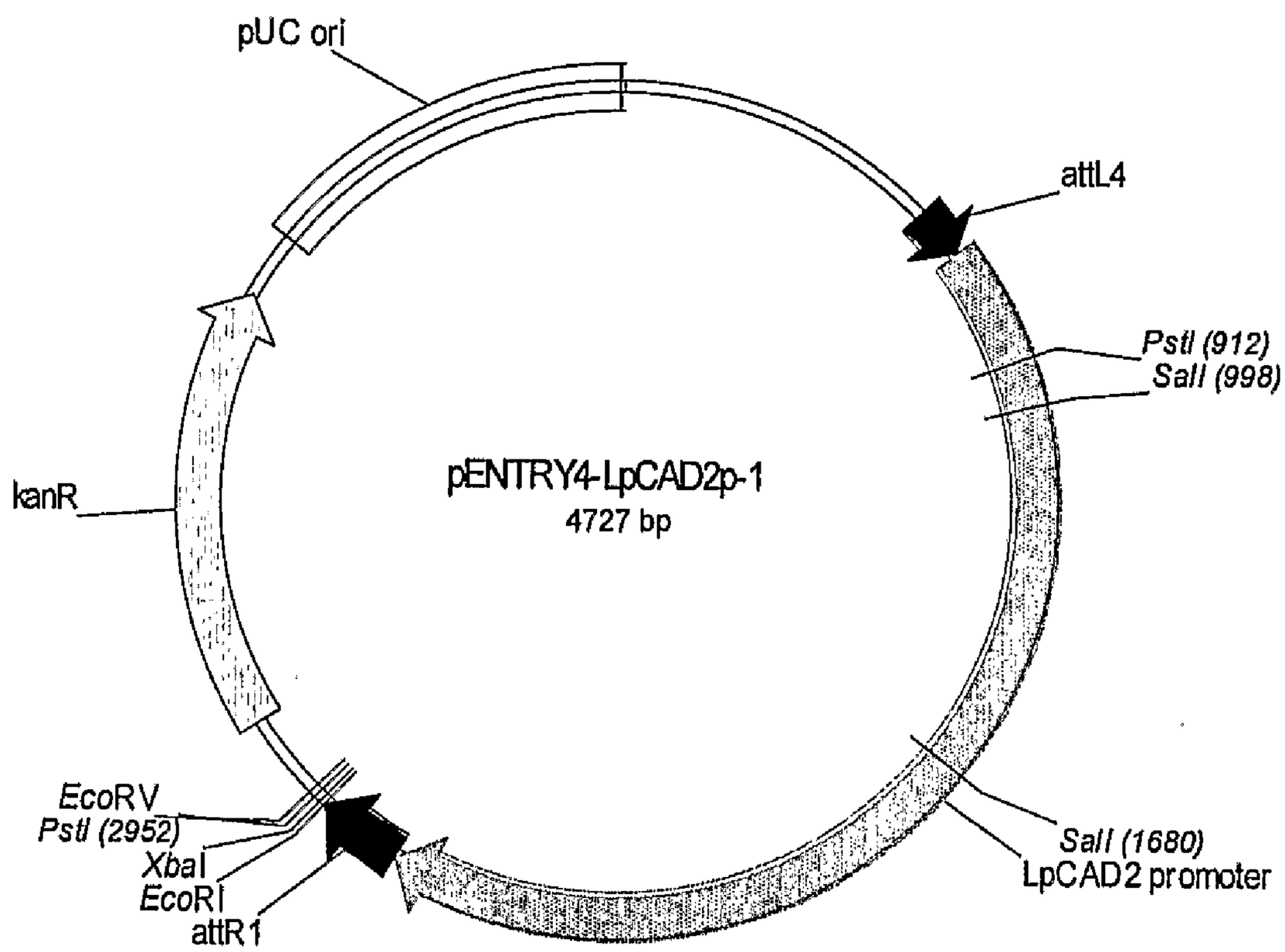

pUC ori
attL4
PstI (912)
SalI (998)
pENTRY4-LpCAD2p-1
4727 bp
kanR
EcoRV
PstI (2952)
XbaI
EcoRI
attR1
SalI (1680)
LpCAD2 promoter

FIGURE 43

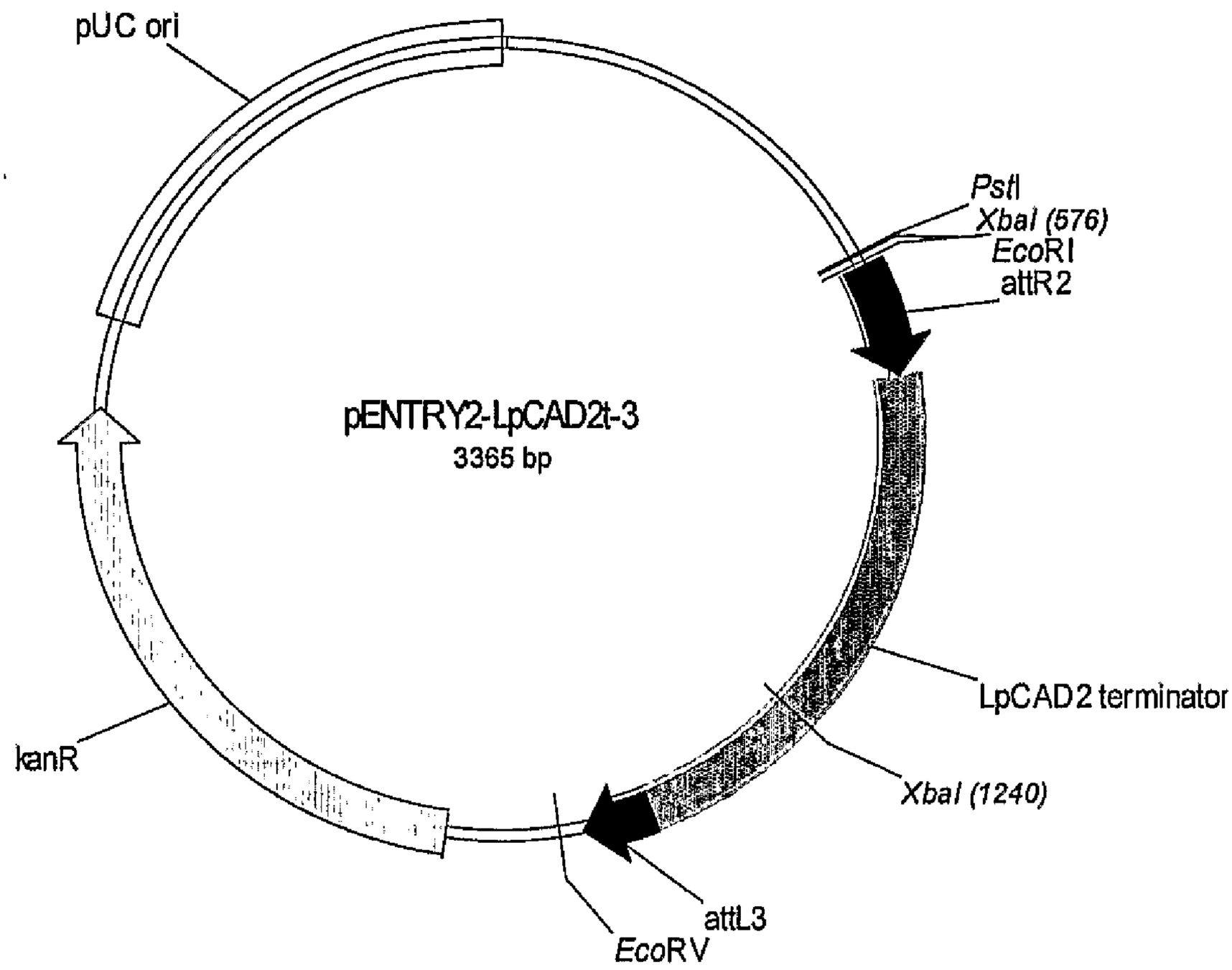

pUC ori
PstI
XbaI (576)
EcoRI
attR2
pENTRY2-LpCAD2t-3
3365 bp
LpCAD2 terminator
kanR
XbaI (1240)
attL3
EcoRV

FIGURE 44

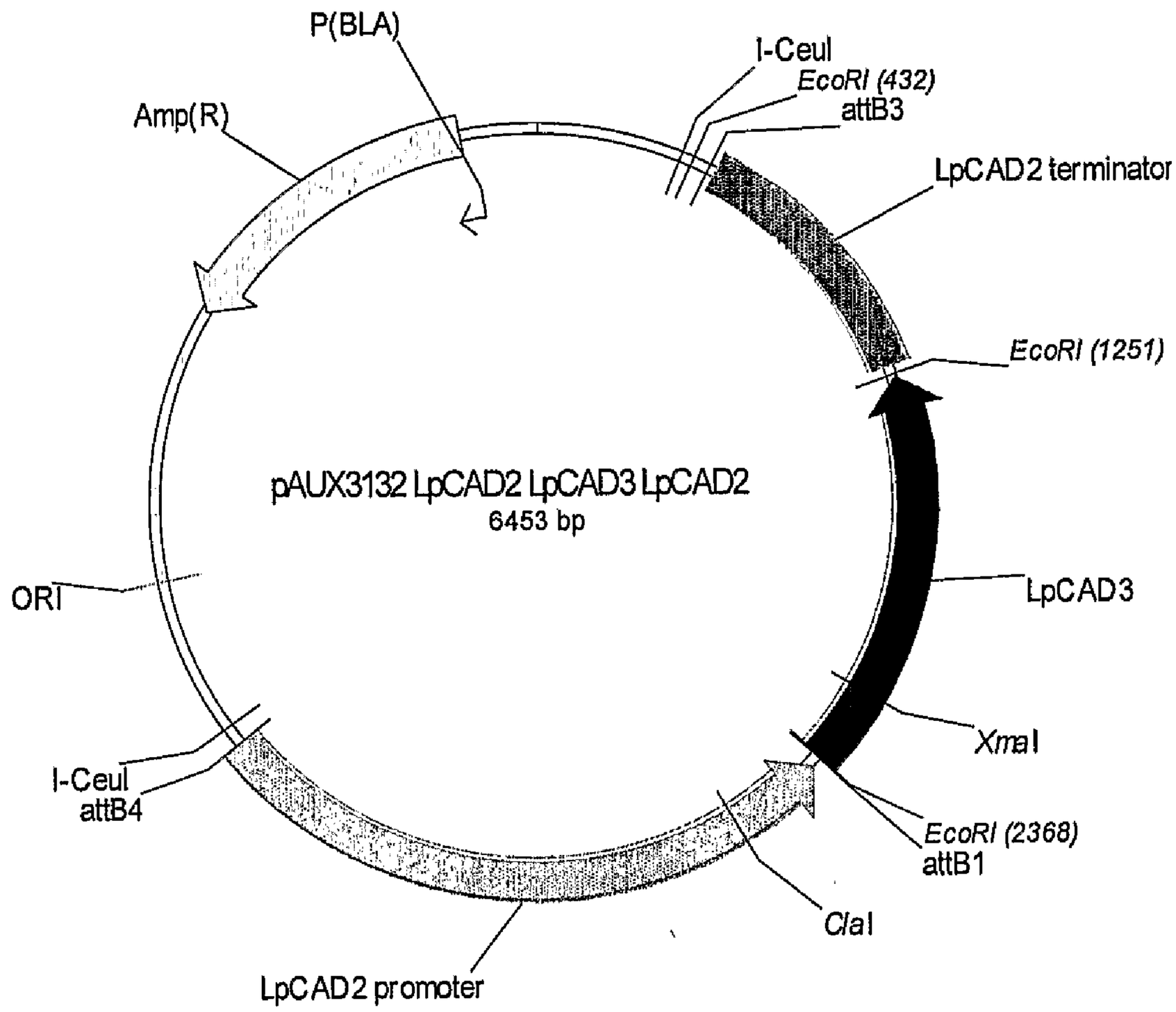

P(BLA)
I-CeuI
EcoRI (432)
attB3
Amp(R)
LpCAD2 terminator
EcoRI (1251)
pAUX3132 LpCAD2 LpCAD3 LpCAD2
6453 bp
ORI
LpCAD3
XmaI
I-CeuI
attB4
EcoRI (2368)
attB1
ClaI
LpCAD2 promoter

FIGURE 45

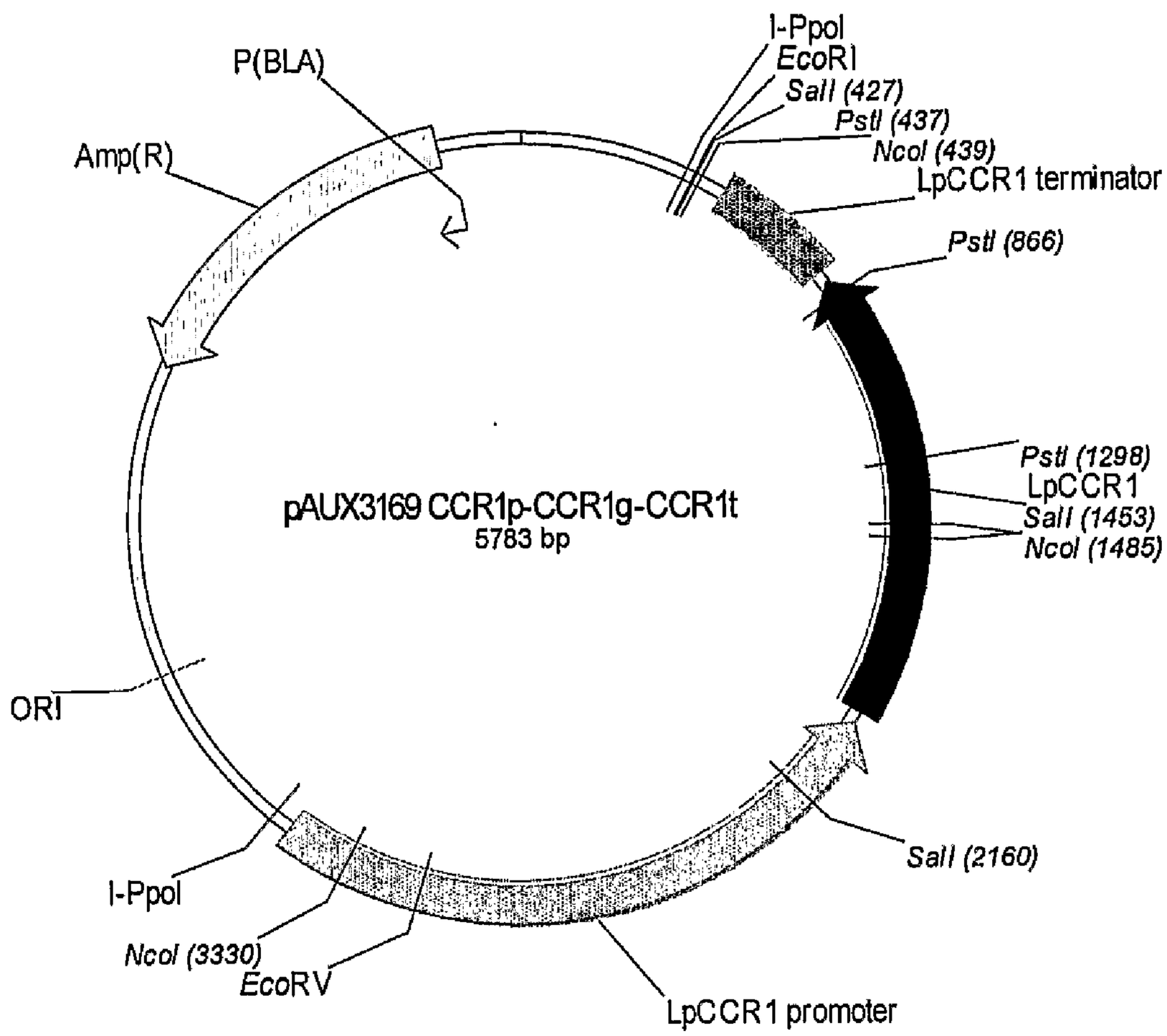

I-PpoI
EcoRI
SalI (427)
PstI (437)
NcoI (439)
LpCCR1 terminator
PstI (866)
P(BLA)
Amp(R)
PstI (1298)
LpCCR1
SalI (1453)
NcoI (1485)
pAUX3169 CCR1p-CCR1g-CCR1t
5783 bp
ORI
SalI (2160)
I-PpoI
NcoI (3330)
EcoRV
LpCCR1 promoter

FIGURE 46

LpCCR1 cDNA sequence

Deleted base

ATGACCATCGCCGAG

Modified forward primer sequence

CAAGTTTGTACAAAAAAGCAGGCTATGACCATCGCGAG

*Att*B1 site

Start site

FIGURE 47

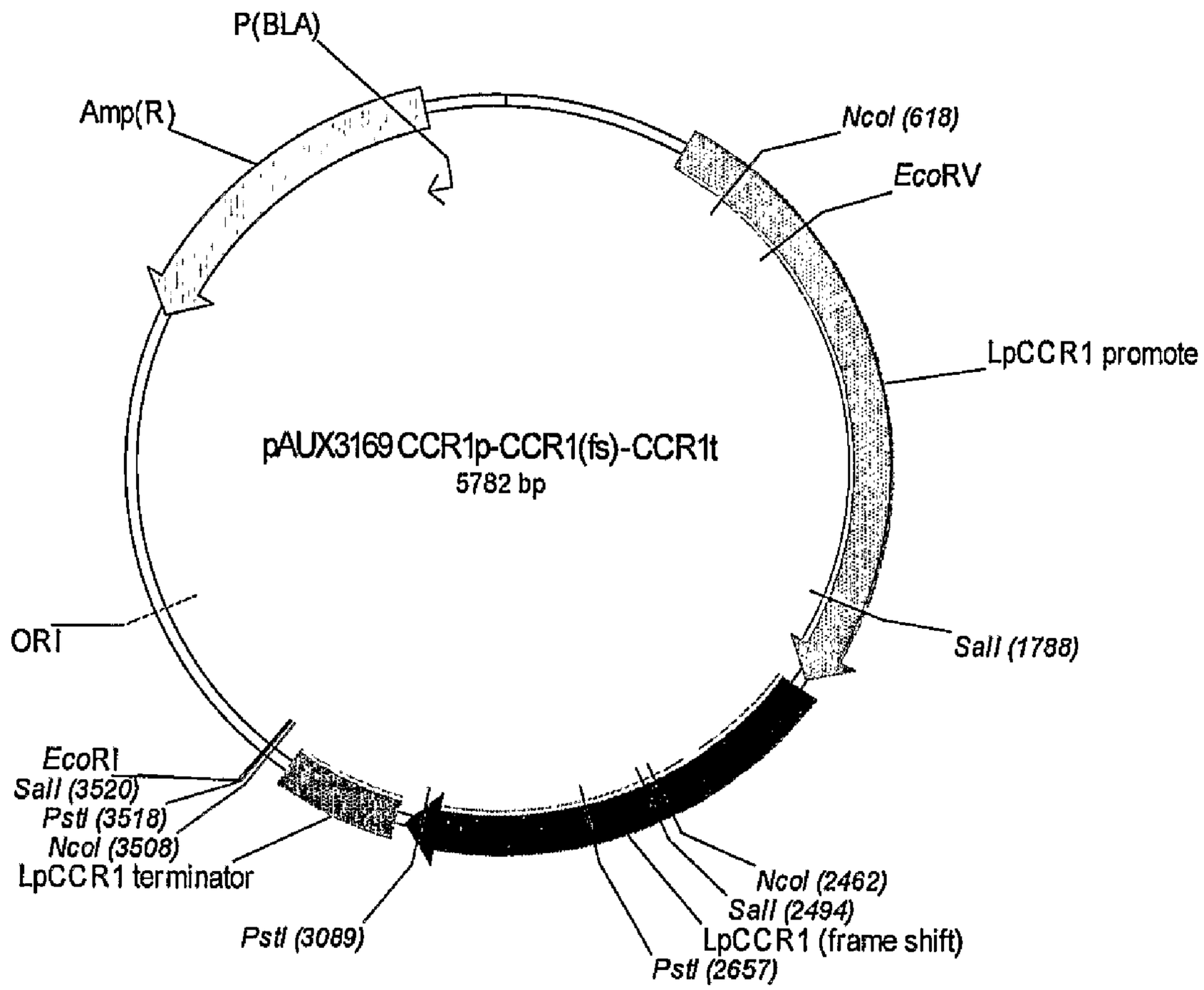

P(BLA)
Amp(R)
NcoI (618)
EcoRV
LpCCR1 promote
pAUX3169 CCR1p-CCR1(fs)-CCR1t
5782 bp
ORI
SalI (1788)
EcoRI
SalI (3520)
PstI (3518)
NcoI (3508)
LpCCR1 terminator
PstI (3089)
NcoI (2462)
SalI (2494)
LpCCR1 (frame shift)
PstI (2657)

FIGURE 48

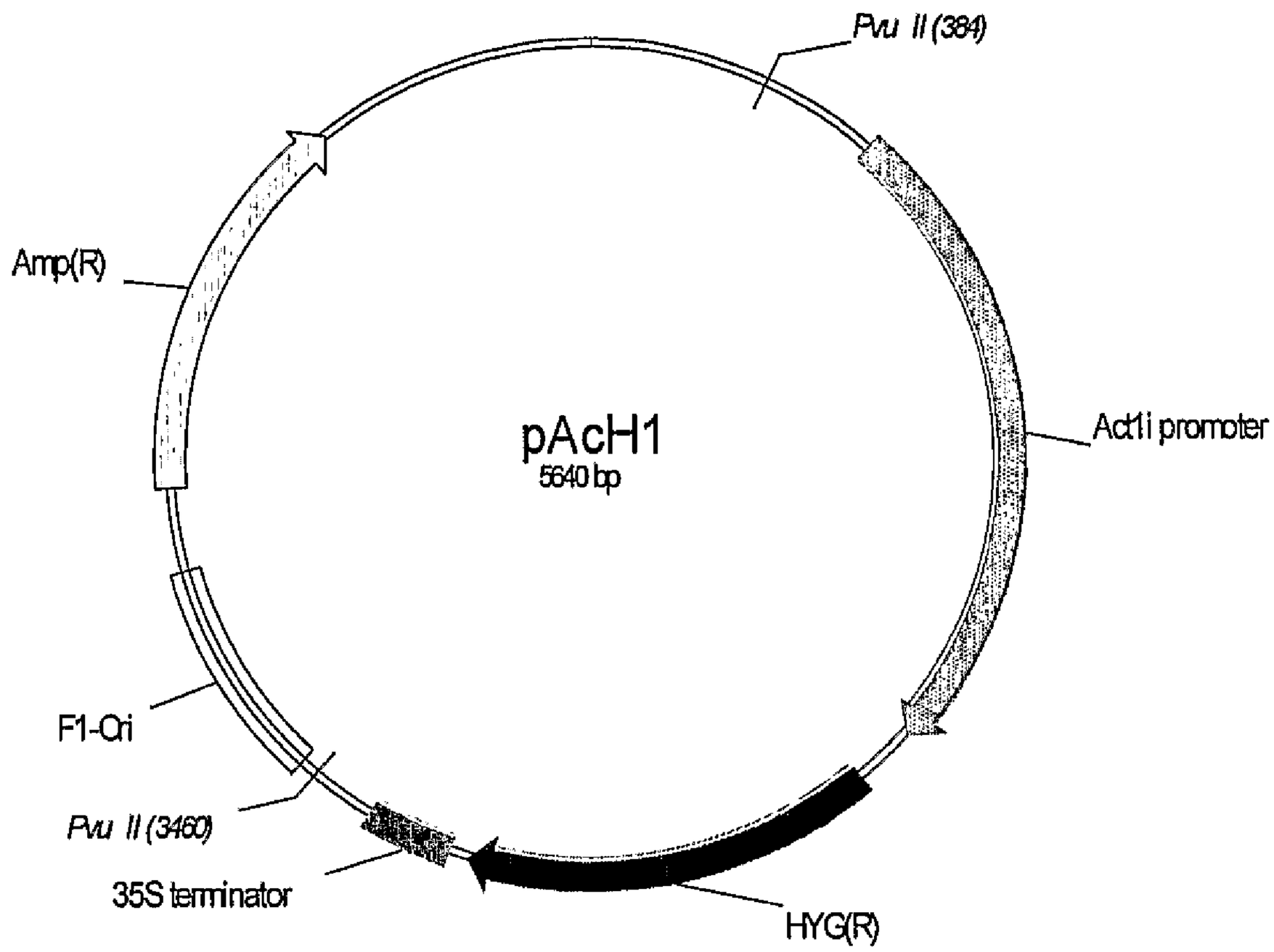

Pvu II (384)
Amp(R)
pAcH1
5640 bp
Act1i promoter
F1-Ori
Pvu II (3460)
35S terminator
HYG(R)

FIGURE 49

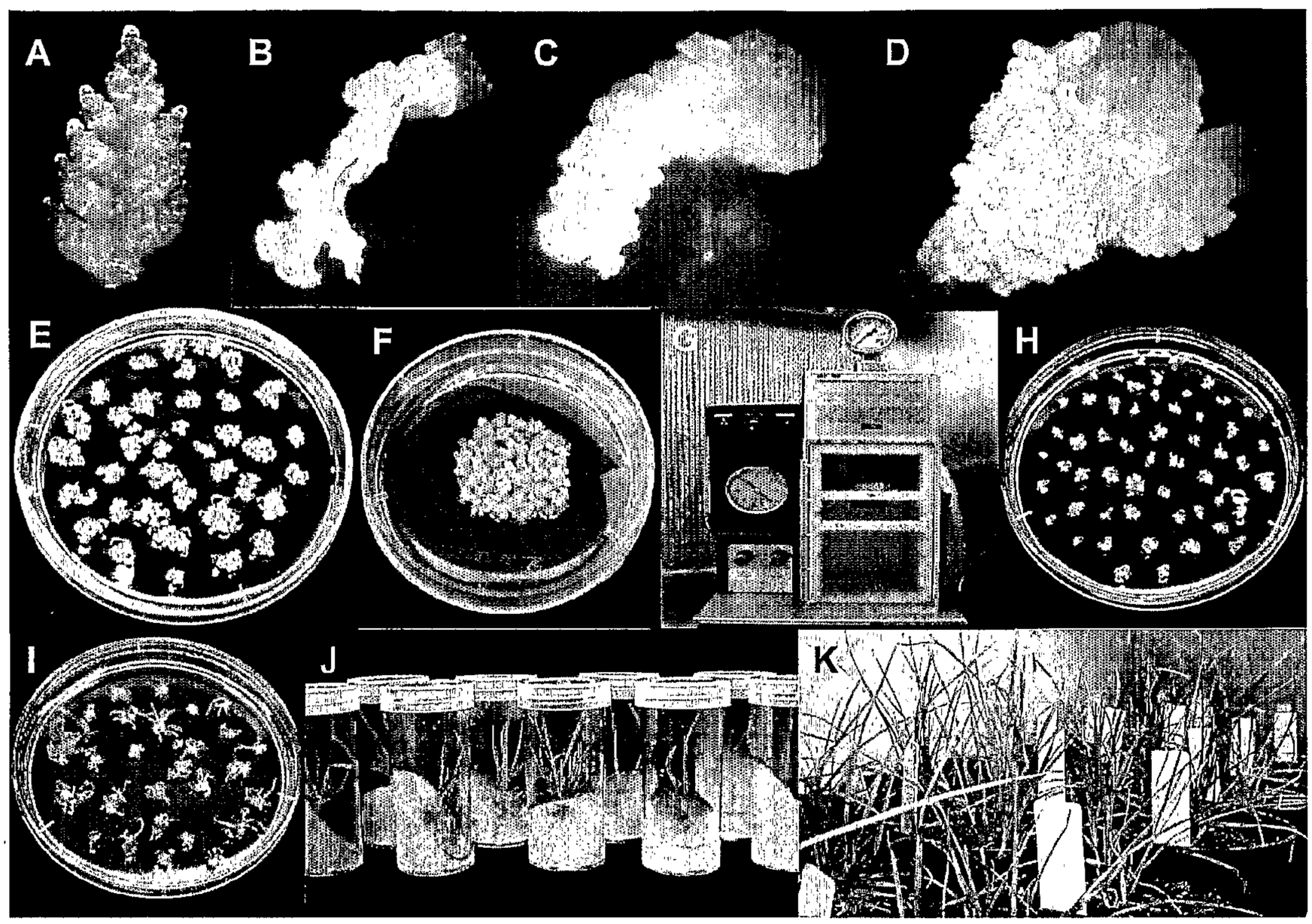


FIGURE 50

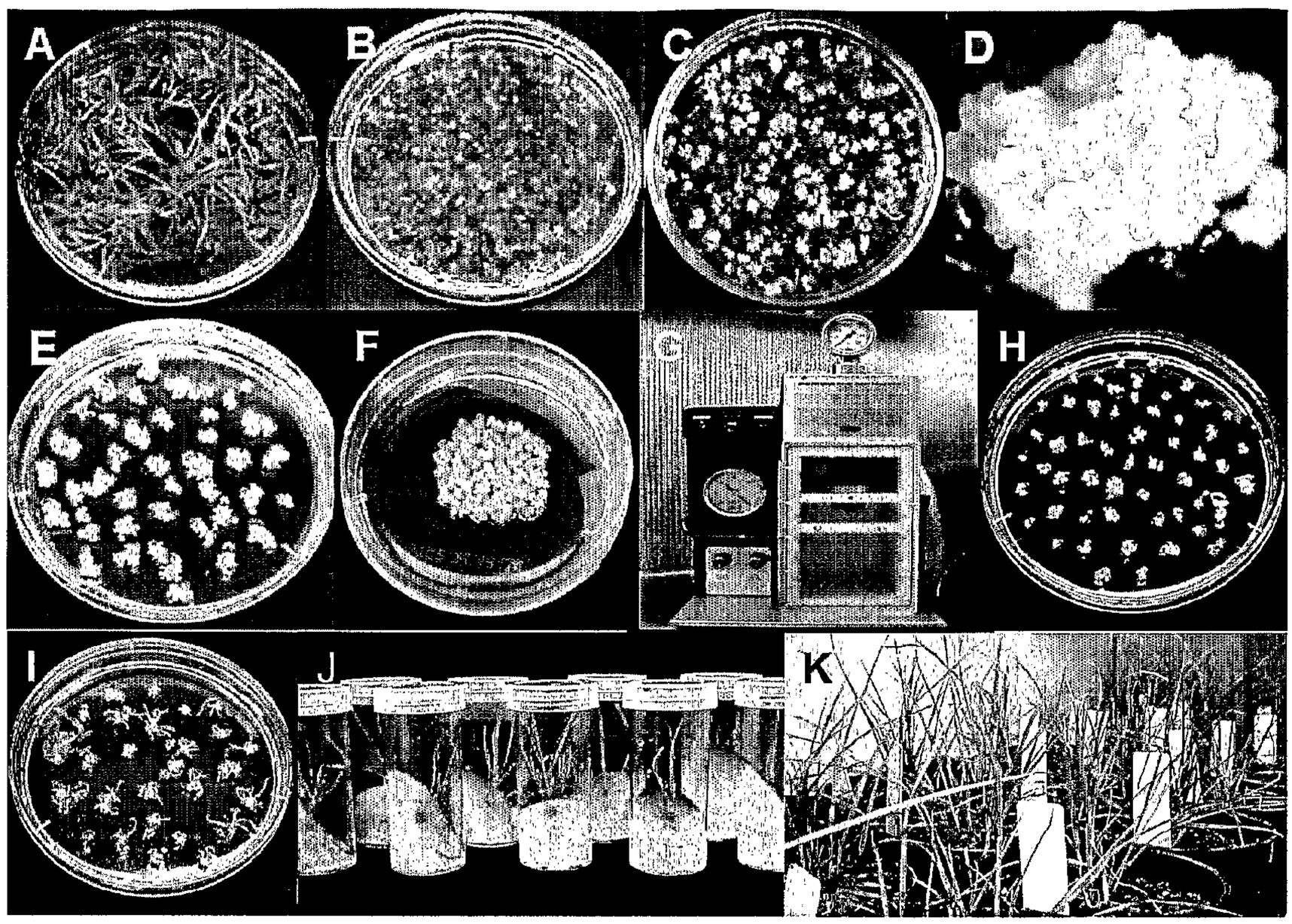


FIGURE 51

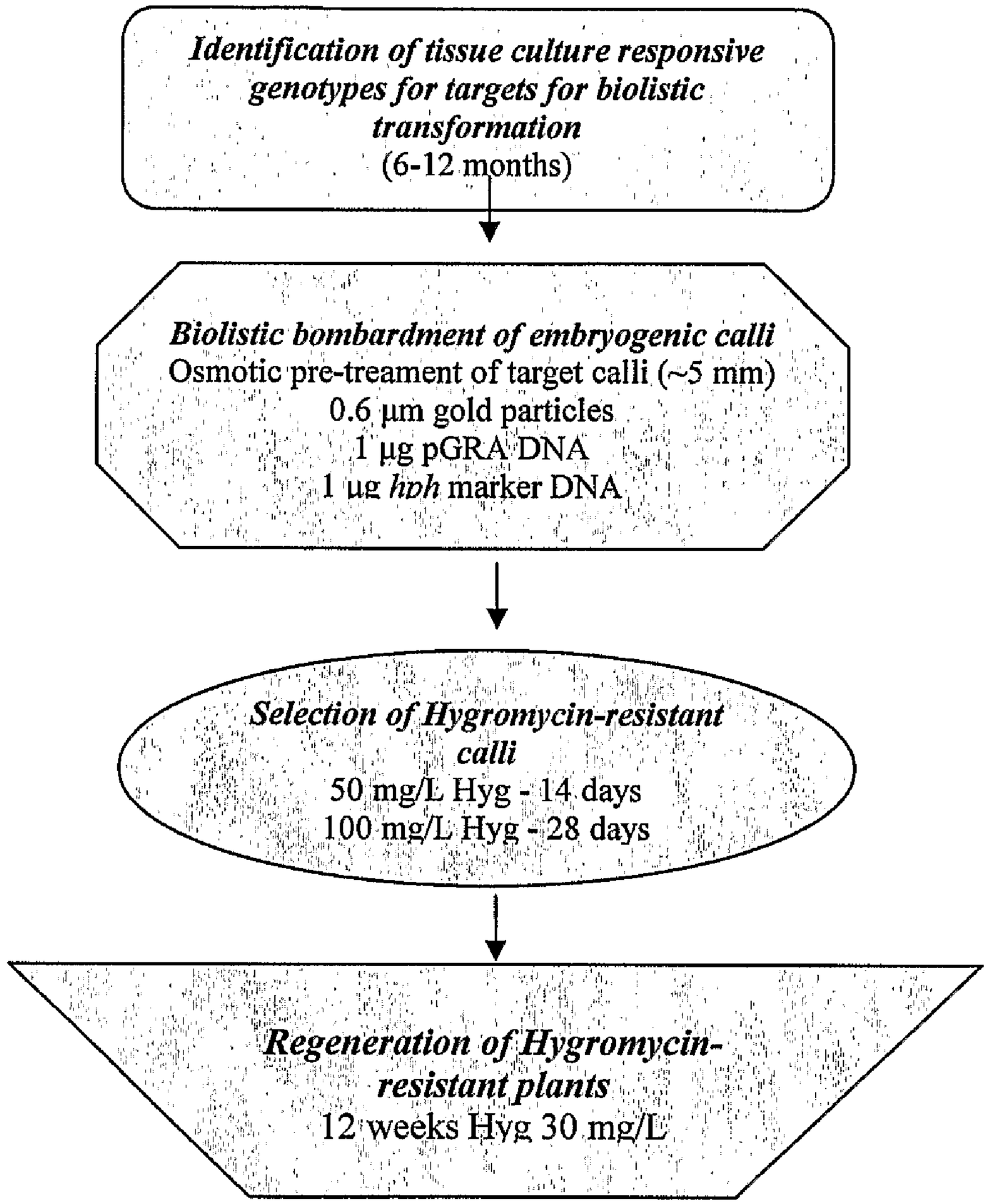

Identification of tissue culture responsive genotypes for targets for biolistic transformation
(6-12 months)
Biolistic bombardment of embryogenic calli
Osmotic pre-treament of target calli (~5 mm)
0.6 µm gold particles
1 µg pGRA DNA
1 µg hph marker DNA
Selection of Hygromycin-resistant calli
50 mg/L Hyg - 14 days
100 mg/L Hyg - 28 days
Regeneration of Hygromycin-resistant plants
12 weeks Hyg 30 mg/L

FIGURE 52

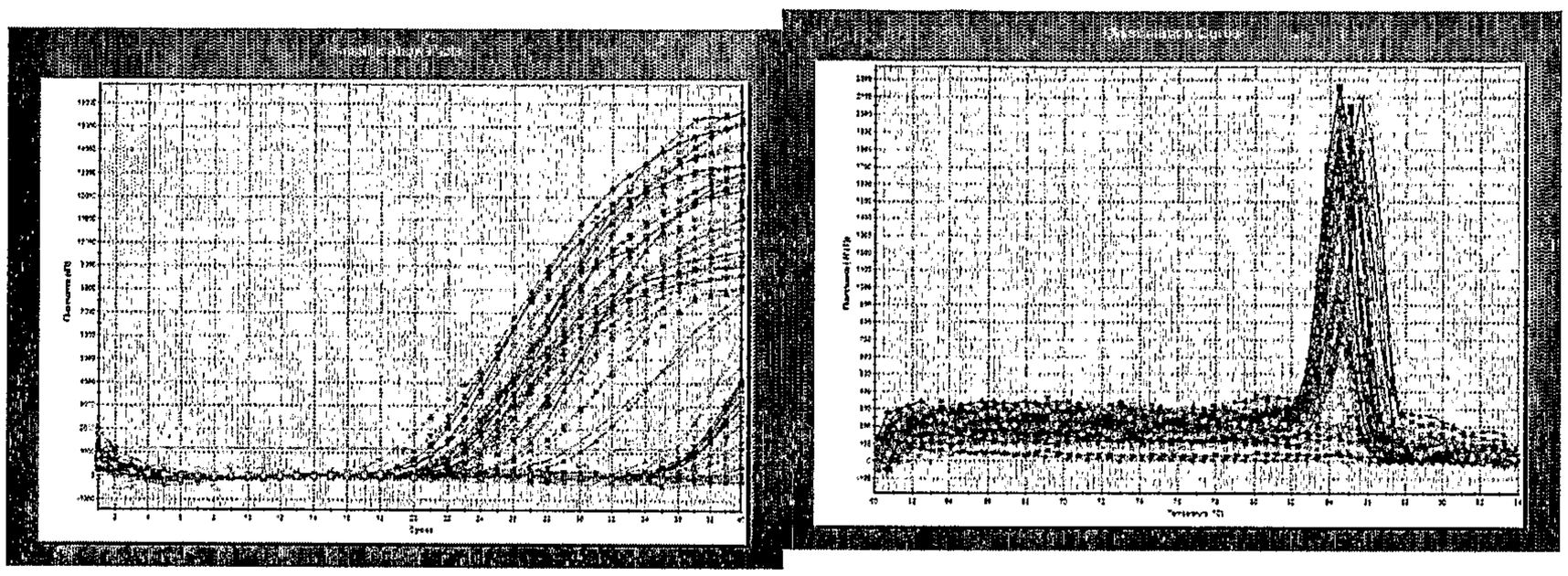


FIGURE 53

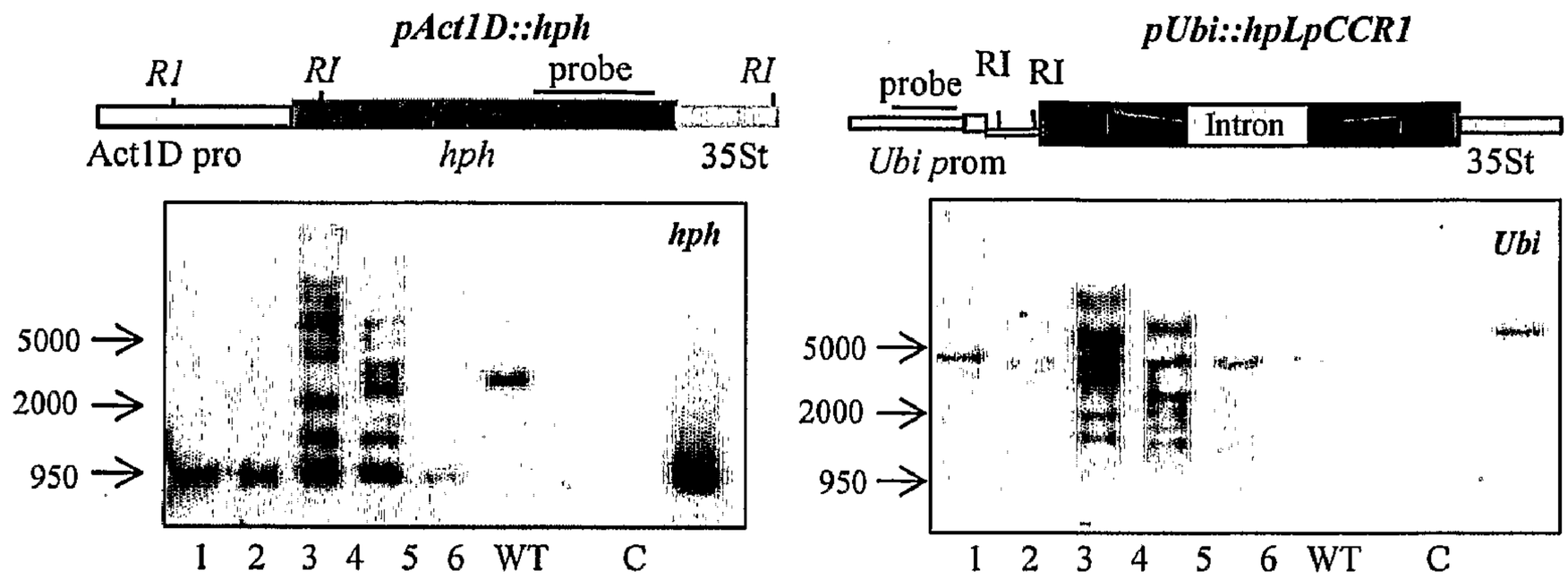

pAct1D::hph
RI
RI
probe
RI
Act1D pro
hph
35St
hph
5000
2000
950
1 2 3 4 5 6 WT C
pUbi::hpLpCCR1
probe
RI
RI
Intron
Ubi prom
35St
Ubi
5000
2000
950
1 2 3 4 5 6 WT C

FIGURE 54

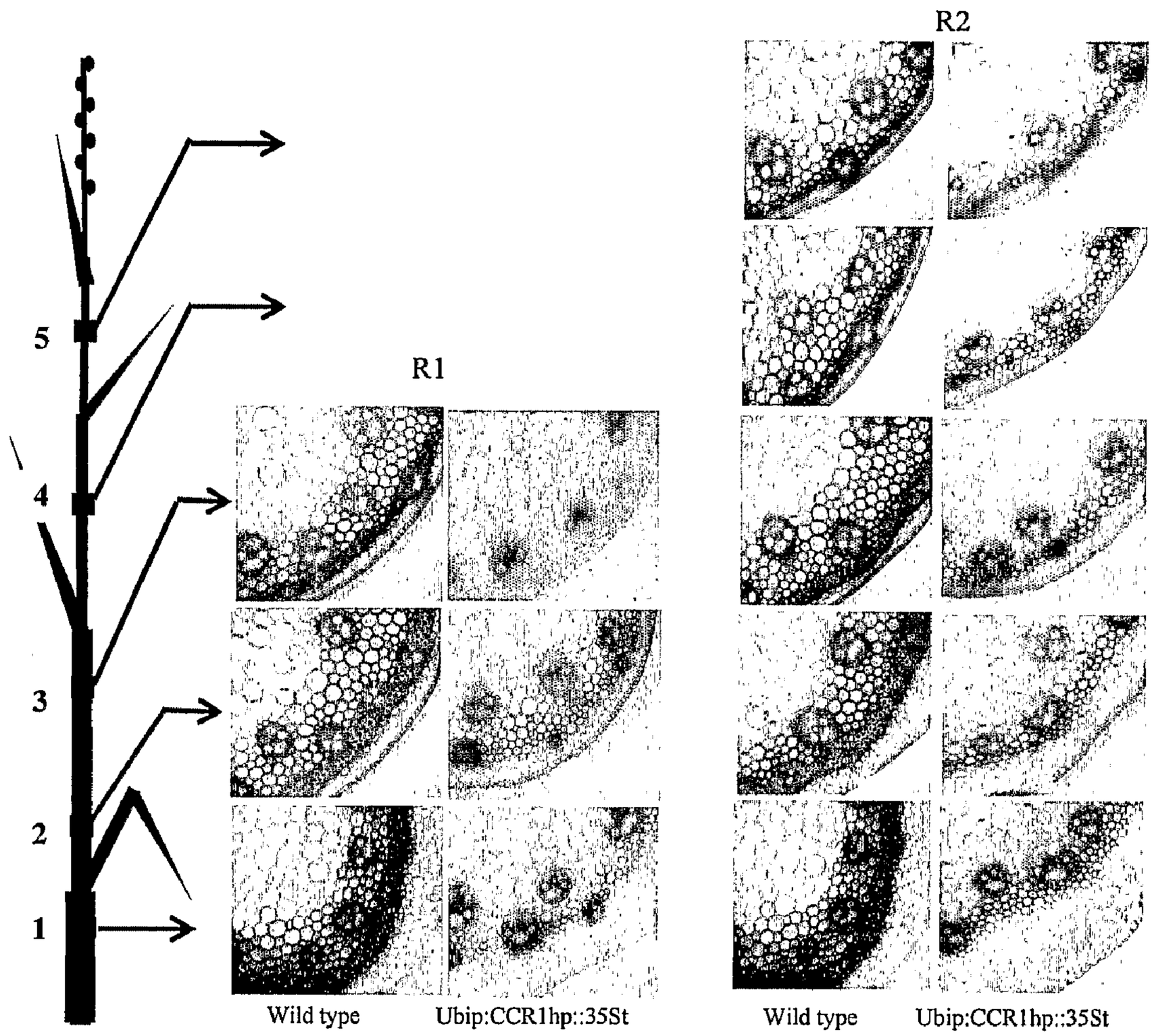

R2
R1
5
4
3
2
1
Wild type
Ubip:CCR1hp::35St
Wild type
Ubip:CCR1hp::35St

FIGURE 55

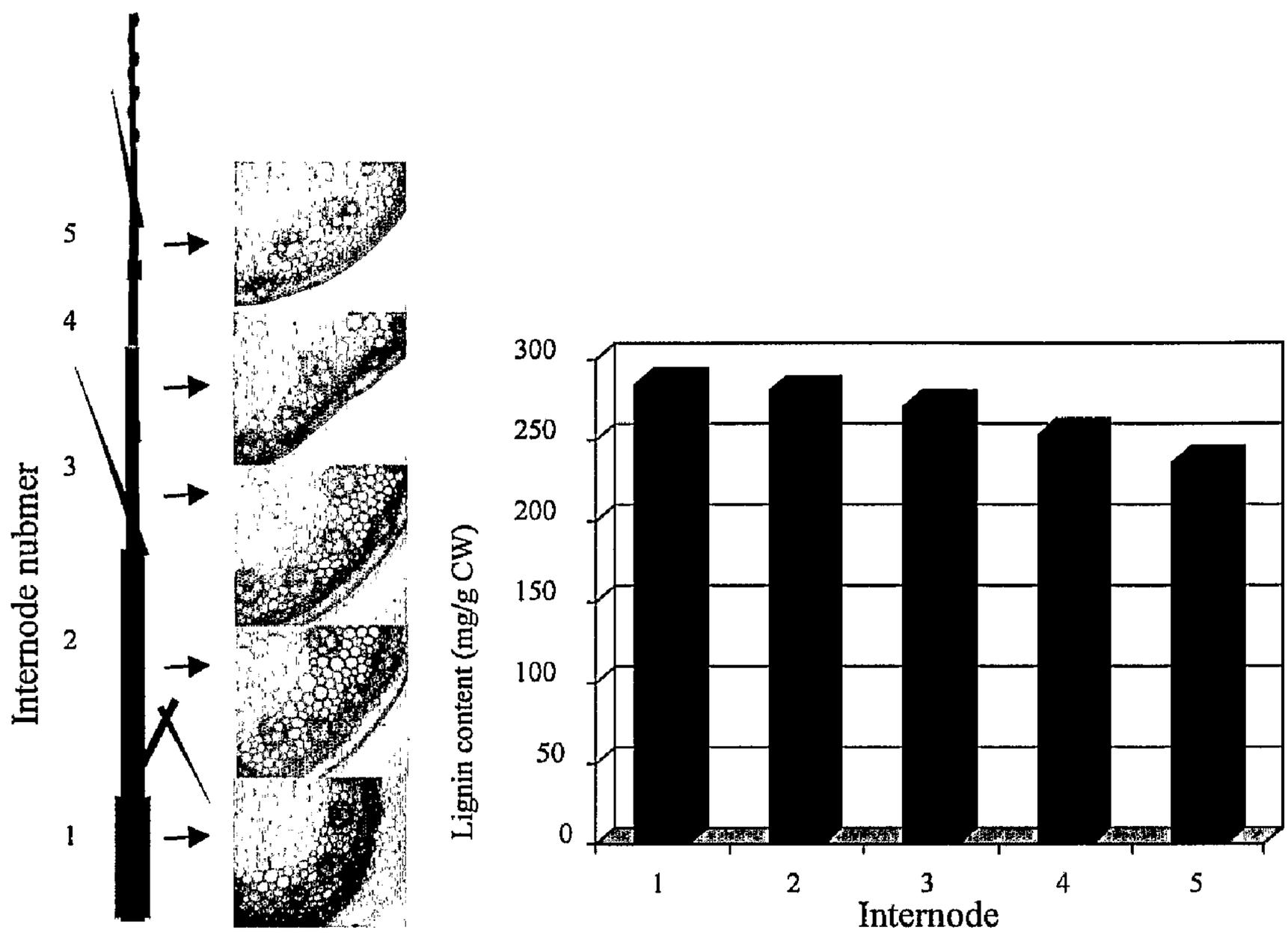

Internode nubmer
5
4
3
2
1
Lignin content (mg/g CW)
300
250
200
150
100
50
0
1
2
3
4
5
Internode

FIGURE 56

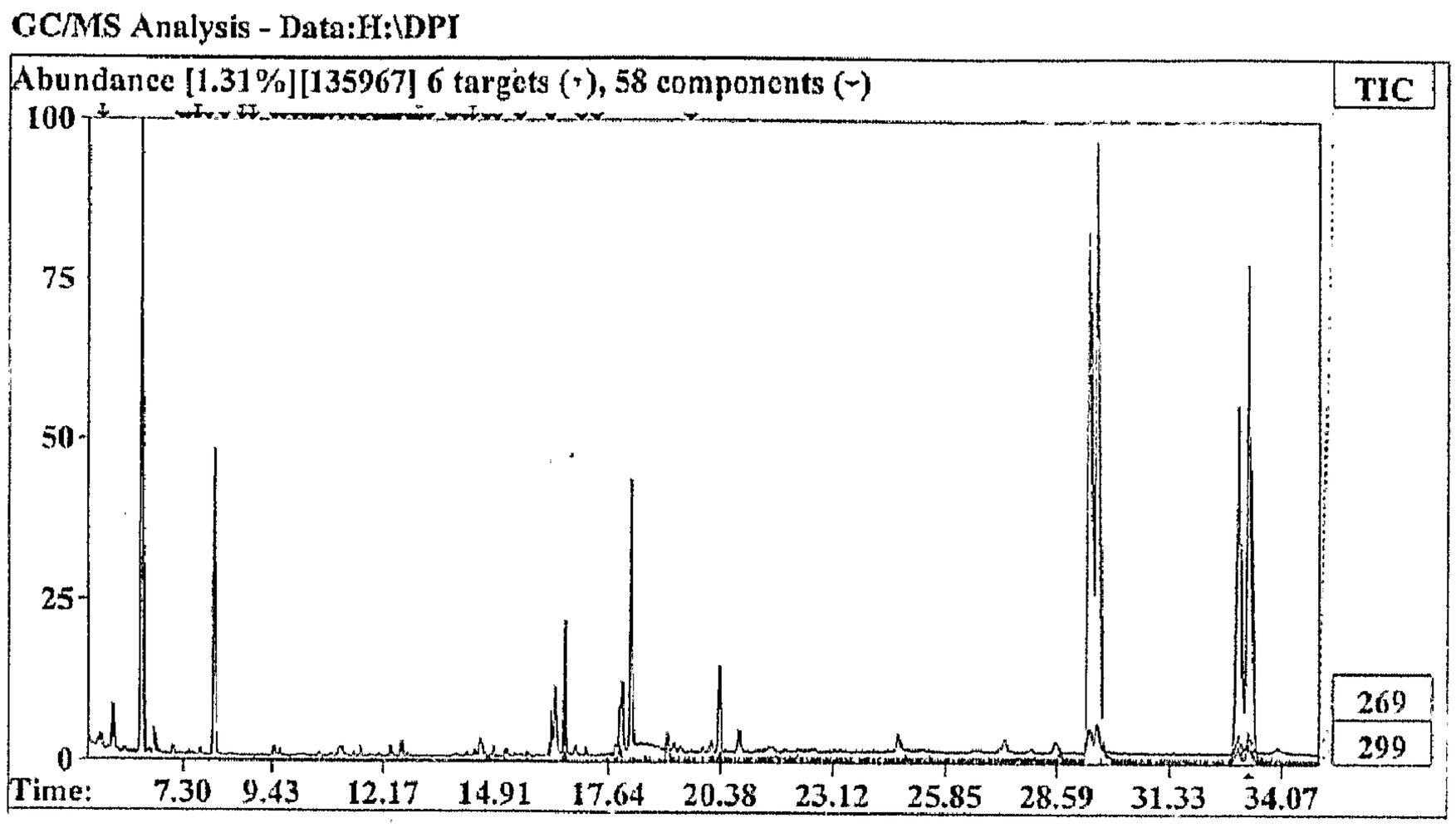

GC/MS Analysis - Data:H:\DPI
Abundance [1.31%][135967] 6 targets (·), 58 components (~)
TIC
100
75
50
25
0
269
299
Time:
7.30
9.43
12.17
14.91
17.64
20.38
23.12
25.85
28.59
31.33
34.07

FIGURE 57

MODIFICATION OF LIGNIN BIOSYNTHESIS VIA SENSE SUPPRESSION

[0001] The present invention relates to the modification of lignin biosynthesis in plants, to enzymes involved in the lignin biosynthetic pathway and nucleic acids encoding such enzymes and, more particularly, to methods of modifying lignin biosynthesis via sense suppression and to related nucleic acids and constructs.

[0002] The present invention also relates to a regulatory element and, more particularly, to a promoter capable of causing expression of an exogenous gene in plant cells, such as a gene encoding an enzyme involved in the lignin biosynthetic pathway in plants.

[0003] The invention also relates to vectors including the nucleic acids and regulatory elements of the invention, plant cells, plants, seeds and other plant parts transformed with the regulatory elements, nucleic acids and vectors, and methods of using the nucleic acids, regulatory elements and vectors.

[0004] Lignins are complex phenolic polymers that strengthen plant cell walls against mechanical and chemical degradation. The process of lignification typically occurs during secondary thickening of the walls of cells with structural, conductive or defensive roles. Three monolignol precursors, sinapyl, coniferyl and p-coumaryl alcohol combine by dehydrogenative polymerisation to produce respectively the syringyl(S), guaiacyl(G) and hydroxyl(H) subunits of the lignin polymer, which can also become linked to cell-wall polysaccharides through the action of peroxidases and other oxidative isozymes. In grasses, biosynthesis of the monolignol precursors is a multistep process beginning with the aromatic amino-acids phenylalanine and tyrosine. It is the final two reduction/dehydrogenation steps of the pathway, catalysed by Cinnamoyl CoA Reductase (CCR) and Cinnamyl Alcohol Dehydrogenase (CAD) that are considered to be specific to lignin biosynthesis. The proportions of monolignols incorporated into the lignin polymer vary depending on plant species, tissue, developmental stage and sub-cellular location.

[0005] Caffeic acid O-methyl transferase (OMT), 4 coumarate CoA-ligase (4CL), cinnamoyl-CoA reductase (CCR) and cinnamyl alcohol dehydrogenase (CAD) are key enzymes involved in lignin biosynthesis.

[0006] Worldwide permanent pasture is estimated to cover 70% of agriculturally cultivated area. Ryegrasses (*Lolium* spp.) together with the closely related fescues (*Festuca* spp.) are of significant value in temperate grasslands. The commercially most important ryegrasses are Italian or annual ryegrass (*L. multiflorum* Lam.) and perennial ryegrass (*L. perenne* L.). They are the key forage species in countries where livestock production is an intensive enterprise, such as the Netherlands, United Kingdom and New Zealand. The commercially most important fescues are tall fescue (*F. anundinacea* Schreb.), meadow fescue (*F. pratensis*) and red fescue (*F. rubra*).

[0007] Perennial ryegrass (*Lolium perenne* L.) is the major grass species sown in temperate dairy pastures in Australia, and the key pasture grass in temperate climates throughout the world. A marked decline of the feeding value of grasses is observed in temperate pastures of Australia during late spring and early summer, where the nutritive value of perennial ryegrass based pasture is often insufficient to meet the metabolic demands of lactating dairy cattle. Perennial ryegrass is also an important turf grass.

[0008] Grass and legume in vitro dry matter digestibility has been negatively correlated with lignin content. In addition, natural mutants of lignin biosynthetic enzymes in maize, sorghum and pearl millet that have higher rumen digestibility have been characterised as having lower lignin content and altered S/G subunit ratio. Thus, lignification of plant cell walls is the major factor identified as responsible for lowering digestibility of forage tissues as they mature.

[0009] It would be desirable to have methods of altering lignin biosynthesis in plants, including grass species such as ryegrasses and fescues, by reducing the activity of key biosynthetic enzymes in order to reduce lignin content and/or alter lignin composition for enhancing dry matter digesitibility and improving herbage quality. However, for some applications it may be desirable to enhance lignin biosynthesis to increase lignin content and/or alter lignin composition, for example to increase mechanical strength of wood, to increase mechanical strength of turf grasses, to reduce plant height and reduce lodging or improve disease resistance.

[0010] While nucleic acid sequences encoding some of the enzymes involved in the lignin biosynthetic pathway have been isolated for certain species of plants, there remains a need for materials useful in the modification of lignin biosynthesis in plants, particularly grass species such as ryegrasses and fescues.

[0011] Other phenotypic traits which may be improved by transgenic manipulation of plants include disease resistance, mineral content, nutrient quality and drought tolerance.

[0012] However, transgenic manipulation of phenotypic traits in plants requires the availability of regulatory elements capable of causing the expression of exogenous genes in plant cells.

[0013] It is an object of the present invention to overcome, or at least alleviate, one or more of the difficulties or deficiencies associated with the prior art.

[0014] In one aspect, the present invention provides substantially purified or isolated nucleic acids or nucleic acid fragments encoding the following enzymes from a ryegrass (*Lolium*) or fescue (*Festuca*) species: 4 coumarate CoA-ligase (4CL), cinnamoyl-CoA reductase (CCR) and cinnamyl alcohol dehydrogenase (CAD).

[0015] The ryegrass (*Lolium*) or fescue (*Festuca*) species may be of any suitable type, including Italian or annual ryegrass, perennial ryegrass, tall fescue, meadow fescue and red fescue. Preferably the ryegrass or fescue species is a *Lolium* species such as *Lolium perenne* or *Lolium arundinaceum* which is otherwise known as *Festuca arundinacea*.

[0016] By 'nucleic acid' is meant a chain of nucleotides capable of carrying genetic information. The term generally refers to genes or functionally active fragments or variants thereof and or other sequences in the genome of the organism that influence its phenotype. The term 'nucleic acid' includes DNA (such as cDNA or genomic DNA) and RNA (such as mRNA or microRNA) that is single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases, synthetic nucleic acids and combinations thereof.

[0017] The nucleic acid or nucleic acid fragment may be of any suitable type and includes DNA (such as cDNA or genomic DNA) and RNA (such as mRNA) that is single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases, and combinations thereof.

[0018] By 'substantially purified' is meant that the nucleic acid or promoter is free of the genes, which, in the naturally-occurring genome of the organism from which the nucleic

acid or promoter of the invention is derived, flank the nucleic acid or promoter. The term therefore includes, for example, a nucleic acid or promoter which is incorporated into a vector; into an autonomously replicating plasmid or virus; or into the genomic DNA of a prokaryote or eukaryote; or which exists as a separate molecule (eg. a cDNA or a genomic or cDNA fragment produced by PCR or restriction endonuclease digestion) independent of other sequences. It also includes a nucleic acid or promoter which is part of a hybrid gene. Preferably, the substantially purified nucleic acid or promoter is at least approximately 90% pure, more preferably at least approximately 95% pure, even more preferably at least approximately 98% pure.

[0019] The term "isolated" means that the material is removed from its original environment (eg. the natural environment if it is naturally occurring). For example, a naturally occurring nucleic acid present in a living plant is not isolated, but the same nucleic acid separated from some or all of the coexisting materials in the natural system, is isolated. Such nucleic acids could be part of a vector and/or such nucleic acids could be part of a composition, and still be isolated in that such a vector or composition is not part of its natural environment.

[0020] In a preferred embodiment of this aspect of the invention, the substantially purified or isolated nucleic acid or nucleic acid fragment encoding 4CL includes a nucleotide sequence selected from the group consisting of (a) sequences shown in FIGS. **2**, **3** and **4** hereto (Sequence ID Nos: 1, 3 and 5; respectively) (b) complements of the sequences shown in FIGS. **2**, **3** and **4** hereto (Sequence ID Nos: 1, 3 and 5, respectively); (c) sequences antisense to the sequences recited in (a) and (b); and (d) functionally active fragments and variants of the sequences recited in (a), (b) and (c).

[0021] In a further preferred embodiment of this aspect of the invention, the substantially purified or isolated nucleic acid or nucleic acid fragment encoding CCR includes a nucleotide sequence selected from the group consisting of (a) the sequence shown in FIG. **10** hereto (Sequence ID No: 7); (b) the complement of the sequence shown in FIG. **10** hereto (Sequence ID No: 7); (c) sequences antisense to the sequences recited in (a) and (b); and (d) functionally active fragments and variants of the sequences recited in (a), (b) and (c).

[0022] In a still further preferred embodiment of this aspect of the invention, the substantially purified or isolated nucleic acid or nucleic acid fragment encoding CAD includes a nucleotide sequence selected from the group consisting of (a) the sequences shown in FIGS. **13**, **14**, **26** and **27** hereto (Sequence ID Nos: 9, 11, 14 and 16, respectively); (c) sequences antisense to the sequences recited in (a) and (b); and (d) functionally active fragments and variants of the sequences recited in (a), (b) and (c).

[0023] By "functionally active" is meant that the fragment or variant (such as an analogue, derivative or mutant) is capable of participating in or modifying lignin biosynthesis in a plant. Such variants include naturally occurring allelic variants and non-naturally occurring variants. Additions, deletions, substitutions and derivatizations of one or more of the nucleotides are contemplated so long as the modifications do not result in loss of functional activity of the fragment or variant. Preferably the functionally active fragment or variant has at least approximately 80% identity to the relevant part of the above mentioned sequence to which the fragment or variant corresponds, more preferably at least approximately 90% identity, even more preferably at least approximately 95% identity, most preferably at least approximately 98% identity. Such functionally active variants and fragments include, for example, those having conservative nucleic acid changes. By 'conservative nucleic acid changes' is meant nucleic acid substitutions that result in conservation of the amino acid in the encoded protein, due to the degeneracy of the genetic code. Such functionally active variants and fragments also include, for example, those having nucleic acid changes which result in conservative amino acid substitutions of one or more residues in the corresponding amino acid sequence. By 'conservative amino acid substitutions' is meant the substitution of an amino acid by another one of the same class, the classes being as follows:

[0024] Nonpolar: Ala, Val, Leu, Ile, Pro, Met Phe, Trp

[0025] Uncharged polar: Gly, Ser, Thr, Cys, Tyr, Asn, Gln

[0026] Acidic: Asp, Glu

[0027] Basic: Lys, Arg, His

[0028] Other conservative amino acid substitutions may also be made as follows:

[0029] Aromatic: Phe, Tyr, His

[0030] Proton Donor: Asn, Gln, Lys, Arg, His, Trp

[0031] Proton Acceptor: Glu, Asp, Thr, Ser, Tyr, Asn, Gln

[0032] Preferably the fragment has a size of at least 20 nucleotides, more preferably at least 50 nucleotides, more preferably at least 100 nucleotides, more preferably at least 200 nucleotides, more preferably at least 500 nucleotides.

[0033] In a still further preferred embodiment of this aspect of the invention the functionally active fragment or variant may be capable of modifying lignin biosynthesis in a plant via sense suppression.

[0034] Accordingly, the present invention provides a substantially purified or isolated nucleic acid including a fragment or variant of a gene encoding a lignin biosynthetic enzyme, said nucleic acid being capable of modifying lignin biosynthesis in a plant via sense suppression.

[0035] By "sense suppression" is meant that when the functionally active fragment or variant is introduced into the plant in sense orientation, it causes an identifiable decrease in expression of the corresponding gene in the transformed plant relative to an untransformed control plant.

[0036] By "sense" orientation is meant that the nucleic acid is in the same orientation or has the same polarity as a messenger RNA copy that is translated or translatable into protein.

[0037] Fragments and variants for sense suppression include those with additions, deletions, substitutions or derivatizations of one or more nucleotides in the nucleic acid or nucleic acid fragment according to the present invention.

[0038] Fragments and variants for sense suppression preferably include those with short deletions of, for example 1 to approximately 500, 1 to approximately 300 or 1 to approximately 100 nucleotides, preferably consecutive nucleotides. In a preferred embodiment, the short deletion may be located at or near, for example within approximately 200, 100, 50 or 20 bases of, the 3' or 5' end of the gene upon which the fragment or variant is based.

[0039] In a preferred embodiment of this aspect of the invention, the functionally active fragment or variant capable of modifying lignin biosynthesis via sense suppression may be a functionally active fragment or variant of a nucleic acid or nucleic acid fragment encoding 4CL, CCR or CAD, for example as herein before described, or as described in International patent applications WO 02/26994 or WO 03/40306;

or a functionally active fragment or variant of a nucleic acid or nucleic acid fragment encoding cinnamate-4-hydroxase (C4H), caffeoyl-CoA3-O-methyltransferase (CCoAOMT or CCoAMT), caffeic acid O-methyltransferase (OMT or COMT), ferulate-5-hydroxylase (F5H) or phenylalanine ammonia lyase (PAL), for example as described in International patent application WO 02/26994 or WO 03/40306; or a functionally active fragment or variant of cinnamate-3-hydroxylase (C3H), for example as described in International patent application WO 2008/064289.

[0040] Preferably the functionally active fragment or variant encodes a 4CL, CCR CAD, C3H, C4H, CCoAOMT, COMT, F5H or PAL polypeptide without enzymatic activity or with substantially reduced enzymatic activity.

[0041] By "substantially reduced enzymatic activity" is meant enzymatic activity which is significantly lower, for example at least approximately 25%, 50% or 75% lower, than the enzymatic activity in a wild type plant.

[0042] Preferably the functionally active fragment or variant includes a frame-shift mutation relative to the corresponding gene upon which the fragment or variant is based. This may result in a loss of or substantial reduction in enzymatic activity in the encoded polypeptide.

[0043] By a "frame-shift mutation" is meant a mutation that inserts or deletes a number of nucleotides that is not evenly divisible by three from a nucleic acid sequence. Due to the triplet nature of gene expression by codons, the insertion or deletion may disrupt the reading frame, or the grouping of the nucleotides into codons, resulting in a different translation from the original. The earlier in the sequence the deletion or insertion occurs, the greater is the proportion of the protein that is altered.

[0044] A frame-shift mutation may cause the reading of codons to be different, so most codons after the mutation (with a few exceptions due to redundancy or coincidental similarity) will code for different amino acids than the corresponding codon in the wild type sequence, leading to a substantially altered polypeptide sequence. Furthermore, the stop codon "UAA, UGA, or UAG" may not be read, or a stop codon may be created at an earlier site. The protein being created may be abnormally short, abnormally long, and/or contain the wrong amino acids. It is unlikely to be functional.

[0045] Deletions or additions occurring at or near the 5' end may preferably be within a short distance, for example within approximately 20, 50, 100 or 200 bases of the ATG start codon, preferably within a short distance downstream of the ATG start codon, for example within approximately 20, 50, 100 or 200 bases downstream of the ATG start codon.

[0046] By "downstream" is meant in the 5'→3' direction along the nucleic acid. Preferably, such deletions or additions occurring at or near the 5' end may result in a frame-shift mutation, so that the resulting polypeptide has little or no enzymatic activity.

[0047] In a particularly preferred embodiment, the deletion or addition at or near the 5' end may be a deletion or addition of one, two, four, five, seven or eight bases, prefer-ably consecutive bases, within a short distance downstream of the ATG start codon, so as to result in a frame-shift mutation, and a resulting polypeptide with little or no enzymatic activity. More preferably the frame-shift mutation is a deletion of one base.

[0048] Deletions occurring at or near the 3' end may preferably start at the 3' end or within a short distance, for example approximately 20, 50, 100 or 200 bases, of the 3' end, and extend in a 5' direction. Preferably, such deletions have a size of between approximately 50 to 500 nucleotides, more preferably approximately 100 to 300 nucleotides.

[0049] In a particularly preferred embodiment of this aspect of the invention, the functionally active fragment or variant capable of modifying lignin biosynthesis via sense suppression may be a functionally active fragment or variant of a nucleic acid or nucleic acid fragment encoding CCR, 4CL or CAD, C3H, C4H, CCoAOMT, COMT, F5H or PAL.

[0050] Accordingly, in a preferred embodiment the present invention provides a fragment or variant of a nucleic acid selected from the group consisting of nucleic acids with the sequences shown in FIG. **10** hereto (Sequence ID No: 7), and in FIGS. 38, 40, 41, 43 and 44 of WO 02/26994 (Sequence ID Nos: 244 to 251, respectively) and in SEQ ID Nos: 147 and 148 of WO 03/40306 (Seq ID Nos. 117 and 121 of this application); wherein said fragment or variant is capable of modifying lignin biosynthesis in a plant via sense suppression of a gene encoding CCR in said plant.

[0051] In a further preferred embodiment the present invention provides a fragment or variant of a nucleic acid selected from the group consisting of the nucleic acids with sequences shown in FIGS. **13**, **14**, **26** and **27** hereto (Sequence ID Nos: 9, 11, 14 and 16, respectively), and in FIGS. 9, 11, 13, 15, 16, 18, 19, 21 and 22 of WO 02/26994 (Sequence ID Nos: 252 to 269, respectively) and in SEQ ID No: 7 of WO 2008/064289 (Seq ID No. 361 of this application) and in SEQ ID Nos: 35 and 145 of WO 03/40306 (Seq ID Nos 53 and 57 of this application; wherein said fragment or variant is capable of modifying lignin biosynthesis in a plant via sense suppression of a gene encoding CAD in said plant.

[0052] In a further preferred embodiment, the present invention provides a fragment or variant of a nucleic acid selected from the group consisting of the nucleic acids with sequences shown in FIGS. **2**, **3** and **4** hereto (Sequence ID Nos: 1, 3 and 5, respectively) and in FIGS. 68, 70, 71 and 73 of WO 02/26994 (Sequence ID Nos: 235-243, respectively) and in SEQ ID Nos: 29, 31, 27, 142 and 143 of WO 03/40306 (Seq ID Nos 21, 25, 33 and 37 of this application); wherein said fragment or variant is capable of modifying lignin biosynthesis in a plant via sense suppression of a gene encoding 4CL in said plant.

[0053] In a further preferred embodiment, the present invention provides a fragment or variant of a nucleic acid selected from the group consisting of the nucleic acids with sequences shown in FIGS. 32, 34, 36 and 76 of WO 02/26994 (Sequence ID Nos: 270-273, respectively) and in SEQ ID No: 6 of WO 2008/064289 (Seq ID No. 49 of this application) and in SEQ ID Nos: 33 and 144 of WO 03/40306 (Seq ID Nos. 41 and 45 of this application); wherein said fragment or variant is capable of modifying lignin biosynthesis in a plant via sense suppression of a gene encoding C4H in said plant.

[0054] In a further preferred embodiment, the present invention provides a fragment or variant of a nucleic acid selected from the group consisting of the nucleic acids with sequences shown in FIGS. 1, 3, 4, 6, 7, 82 and 87 of WO 02/26994 (Sequence ID Nos: 274 to 294, respectively) and in SEQ ID No: 8 of WO 2008/064289 (Seq ID No. 362 of this application) and in SEQ ID Nos: 37 and 146 of WO 03/40306 (Seq ID Nos. 89 and 93 of this application); wherein said fragment or variant is capable of modifying lignin biosynthesis in a plant via sense suppression of a gene encoding CCoAOMT in said plant.

[0055] In a further preferred embodiment, the present invention provides a fragment or variant of a nucleic acid selected from the group consisting of the nucleic acids with sequences shown in FIGS. 24, 26, 27, 29, 30, 93 and 99 of WO 02/26994 (Sequence ID Nos: 295 to 342, respectively) and SEQ ID Nos: 2, 8 and 9 of WO 2008/064289 (Seq ID Nos. 360, 362 and 363 of this application) and in SEQ ID Nos: 149, 42, 150 and 43 of WO 03/40306 (Seq ID Nos. 133, 137, 141 and 145 of this application); wherein said fragment or variant is capable of modifying lignin biosynthesis in a plant via sense suppression of a gene encoding COMT in said plant.

[0056] In a further preferred embodiment, the present invention provides a fragment or variant of a nucleic acid selected from the group consisting of the nucleic acids with sequences shown in FIGS. 59 and 61 of WO 02/26994 (Sequence ID Nos: 343 to 346, respectively) and in SEQ ID Nos: 45 and 151 of WO 03/40306 (Seq ID Nos. 173 and 177 of this application); wherein said fragment or variant is capable of modifying lignin biosynthesis in a plant via sense suppression of a gene encoding F5H in said plant.

[0057] In a further preferred embodiment, the present invention provides a fragment or variant of a nucleic acid selected from the group consisting of the nucleic acids with sequences shown in FIGS. 62, 64, 65 and 67 of WO 02/26994 (Sequence ID Nos: 347 to 358, respectively) and in SEQ ID Nos: 152, 153, 50, 54, 48, 53, 156, 49, 51, 154, 52 and 155 of WO 03/40306 (Seq ID Nos. 181, 185, 189, 193, 197, 201, 205, 209, 213, 217, 220 and 224 of this application); wherein said fragment or variant is capable of modifying lignin biosynthesis in a plant via sense suppression of a gene encoding PAL in said plant.

[0058] In a further preferred embodiment, the present invention provides a fragment or variant of a nucleic acid selected from the group consisting of the nucleic acids with sequences shown in SEQ ID No: 1 of WO 2008/064289 (Seq ID No. 359 of this application) wherein said fragment or variant is capable of modifying lignin biosynthesis in a plant via sense suppression of a gene encoding C3H in said plant.

[0059] Preferably, the fragment or variant includes a short deletion at or near the 3' or 5' end of a sequence as hereinbefore described.

[0060] Preferably, the fragment or variant includes a frame-shift mutation relative to a sequence, as hereinbefore described.

[0061] In a particularly preferred embodiment, the fragment or variant comprises sequence selected from the group of frame shift DNA sequences shown in Tables 1 and 2 or encodes a polypeptide comprising a sequence selected from the group of frame shift protein sequences shown in Tables 1 and 2.

TABLE 1

SEQ ID NO. DNA	SEQ ID NO. PROT	Species	Gene name	Abbrev	DNA Seq ID NO.	No of NT
4CL						
29	91	Fescue	4 Coumarate CoA ligase 2	4CL-2	21	1934
31	93	Fescue	4 Coumarate CoA ligase 3	4CL-3	25	2073
27	89	Lolium	4 Coumarate CoA ligase 1	4CL-1	29	1855
142	90	Lolium	4 Coumarate CoA ligase 2	4CL-2	33	2039
143	178	Lolium	4 Coumarate CoA ligase 3	4CL-3	37	2006
C4H						
33	95	Fescue	cinnamate 4-hydroxylase	C4H	41	1775
144	179	Lolium	cinnamate 4-hydroxylase	C4H	45	1789
		Fescue	cinnamate 4-hydroxylase	LaC4H	49	1518
CAD3						
35	97	Fescue	cinnamyl alcohol dehydrogenase	CAD	53	1313
145	180	Lolium	cinnamyl alcohol dehydrogenase	CAD	57	1358
CAD						
		Fescue	cinnamyl alcohol dehydrogenase	LaCAD1a	61	1501
		Fescue	cinnamyl alcohol dehydrogenase	LaCAD1b	65	1339
		Fescue	cinnamyl alcohol dehydrogenase	LaCAD2a	69	1322
		Fescue	cinnamyl alcohol dehydrogenase	LaCAD2b	73	1526
		Lolium	cinnamyl alcohol dehydrogenase	LpCAD1	77	1325
		Lolium	cinnamyl alcohol dehydrogenase	LpCAD2	81	1378
		Lolium	cinnamyl alcohol dehydrogenase	LpCAD3	85	1382

TABLE 1-continued

CCoAOMT						
37	99	Fescue	Caffeoyl CoA O-methyltransferase	CCoAOMT	89	1063
146	98	Lolium	Caffeoyl CoA O-methyltransferase	CCoAOMT	93	1051
		Lolium	Caffeoyl CoA O-methyltransferase	LpCCoAOMT1	97	1126
		Lolium	Caffeoyl CoA O-methyltransferase	LpCCoAOMT2	101	1164
		Lolium	Caffeoyl CoA O-methyltransferase	LpCCoAOMT3	105	1088
		Lolium	Caffeoyl CoA O-methyltransferase	LpCCoAOMT4	109	1241
		Lolium	Caffeoyl CoA O-methyltransferase	LpCCoAOMT5	113	1151
CCR						
148	101	Fescue	cinnamoyl CoA reductase	CCR	117	1236
147	181	Lolium	cinnamoyl CoA reductase	CCR	121	1332
		Lolium	cinnamoyl CoA reductase	LpCCR1	125	1395
		Lolium	cinnamoyl CoA reductase	LpCCR2	129	1207
COMT						
149	182	Fescue	caffeic acid O-methyltransferase	COMT	133	1428
42	104	Fescue	caffeic acid O-methyltransferase 1	COMT-1	137	1452
150	103	Lolium	caffeic acid O-methyltransferase 1	COMT-1	141	1455
43	105	Lolium	caffeic acid O-methyltransferase 3	COMT-3	145	1440
		Fescue	caffeic acid O-methyltransferase 1	LaCOMT1c	149	1438
		Fescue	caffeic acid O-methyltransferase 1	LaCOMT3	153	1430
		Lolium	caffeic acid O-methyltransferase 3	LpOMT1	157	1542
		Lolium	caffeic acid O-methyltransferase 3	LpOMT2	161	1496
		Lolium	caffeic acid O-methyltransferase 3	LpOMT3	165	1505
		Lolium	caffeic acid O-methyltransferase 3	LpOMT4	169	1366
F5H						
45	107	Fescue	Ferulate 5-hydroxylase	F5H	173	2051
151	183	Lolium	Ferulate 5-hydroxylase	F5H	177	2101
PAL						
152	108	Lolium	Phenylalanine ammonia lyase	PAL	181	2460
153	184	Fescue	Phenylalanine ammonia lyase	PAL	185	2595
50	112	Fescue	Peroxidase	PER	189	1205
54	116	Fescue	Peroxidase	PER	193	1266
48	110	Fescue	Peroxidase	PER	197	1301
53	115	Lolium	Peroxidase	PER	201	1059
156	185		Peroxidase	PER	205	1204
49	111	Lolium	Peroxidase	PER	209	1236
51	113	Lolium	Peroxidase	PER	213	1382
154			Peroxidase	PER	217	1382
52	114	Lolium	Peroxidase	PER	220	1261
155			Peroxidase	PER	224	1260

TABLE 1-continued

SEQ ID NO. DNA	PROT SEQ ID NO.	No of AA	ORF start	ORF end	Frame shift DNA SEQ ID NO.	Frame shift protein SEQ ID NO.	Important Info
4CL							
29	22	559	72	1751	23	24	From WO 03/40306
31	26	557	137	1810	27	28	From WO 03/40306
27	30	539	3	1622	31	32	From WO 03/40306
142	34	559	85	1764	35	36	From WO 03/40306 Corrected SEQ 28, no change in protein seq
143	38	557	126	1799	39	40	From WO 03/40306 Corrected SEQ 30 and 92
C4H							
33	42	505	80	1597	43	44	From WO 03/40306
144	46	501	61	1566	47	47	From WO 03/40306 Corrected SEQ 32 and 94
	50	506	1	1518	51	52	From US patent WO 2008/064289 also known as PCT/US2007/085344
CAD3							
35	54	361	86	1171	55	56	From WO 03/40306
145	58	361	67	1152	59	60	From WO 03/40306 FL of SEQ 34 and 96
CAD							
	62	361	40	1125	63	64	AF188292
	66	361	40	1125	67	68	AF188293; Also in US patent WO 2008/064289 also known as PCT/US2007/085344
	70	361	91	1176	71	72	AF188294
	74	361	91	1176	75	76	AF188295
	78	407	22	1245	79	80	
	82	370	102	1214	83	84	
	86	361	81	1166	87	88	
CCoAOMT							
37	90	265	75	872	91	92	From WO 03/40306
146	94	265	55	852	95	96	From WO 03/40306 Corrected SEQ 36, no change in protein seq
	98	261	132	917	99	100	
	102	261	135	920	103	104	
	106	243	171	902	107	108	
	110	243	170	901	111	112	
	114	265	137	934	115	116	
CCR							
148	118	342	90	1118	119	120	From WO 03/40306 Corrected SEQ 39, no change in protein seq
147	122	363	148	1239	123	124	From WO 03/40306 Corrected SEQ 38 and 100
	126	362	150	1238	127	128	
	130	344	1	1035	131	132	From NCBI Accession # AF278698
COMT							
149	134	360	27	1109	135	136	From WO 03/40306 Corrected SEQ 40 and 102. Also in US patent WO 2008/064289 also known as PCT/US2007/085344
42	138	360	64	1146	139	140	From WO 03/40306
150	142	360	66	1148	143	144	From WO 03/40306 Corrected SEQ 41, no change in protein seq

TABLE 1-continued

43	146	361	85	1170	147	148	From WO 03/40306
	150	360	62	1144	151	152	NCBI accession no AF153825
	154	360	78	1160	155	156	NCBI accession no AF153826
	158	360	139	1221	159	160	
	162	351	135	1187	163	164	
	166	361	156	1238	167	168	
	170	367	107	1209	171	172	
F5H							
45	174	542	93	1721	175	176	From WO 03/40306
151	178	543	87	1718	179	180	From WO 03/40306 Corrected SEQ 44 and 106
PAL							
152	182	711	111	2246	183	184	From WO 03/40306 Corrected SEQ 46, no change in protein seq
153	186	713	143	2284	187	188	From WO 03/40306 Corrected SEQ 47 and 109
50	190	326	87	1067	191	192	From WO 03/40306
54	194	311	80	1015	195	196	From WO 03/40306
48	198	323	22	993	199	200	From WO 03/40306
53	202	293	1	882	203	204	From WO 03/40306
156	206	324	46	1017	207	208	From WO 03/40306 FL of SEQ 53 and 115, now two sequences with difference in 5' region with SEQ ID NO 162
49	210	344	4	1038	211	212	From WO 03/40306
51	214	358	59	1135	215	216	From WO 03/40306
154					218	219	From WO 03/40306 Corrected SEQ 51, no change in protein seq
52	221	344	18	1052	222	223	From WO 03/40306
155					225	226	From WO 03/40306 Corrected SEQ 52, no change in protein seq

TABLE 2

Grass	Species	Gene name	Abbrev	DNA SEQ ID NO.	No of NT	PROT SEQ ID NO.	No of AA	ORF start	ORF end	Frame shift DNA SEQ ID NO.	Frame shift protein SEQ ID NO.	Important Info
Bermuda grass	*Cynodon dactylon*	cinnamate 3-hydroxylase	C3H	227	1539	228	512	1	1539	229	230	From US patent WO 2008/064289 also known as PCT/US2007/085344
	Cynodon dactylon	caffeic acid O-methyltransferase	COMT	231	789	232	262	1	789	233	234	From US patent WO 2008/064289 also known as PCT/US2007/085344

[0062] In a second aspect of the present invention there is provided a genetic construct or a vector including a nucleic acid or nucleic acid fragment according to the present invention.

[0063] In a preferred embodiment of this aspect of the invention, the vector may include a regulatory element such as a promoter, a nucleic acid or nucleic acid fragment, according to the present invention and a terminator; said regulatory element, nucleic acid or nucleic acid fragment and terminator being operatively linked.

[0064] By 'genetic construct' is meant a recombinant nucleic acid molecule.

[0065] By a 'vector' is meant a genetic construct used to transfer genetic material to a target cell.

[0066] By 'operatively linked' is meant that the nucleic acid(s) and a regulatory sequence, such as a promoter, are linked in such a way as to permit expression of said nucleic acid under appropriate conditions, for example when appropriate molecules such as transcriptional activator proteins are bound to the regulatory sequence. Preferably an operatively linked promoter is upstream of the associated nucleic acid.

[0067] The vector may be of any suitable type and may be viral or non-viral. The vector may be an expression vector. Such vectors include chromosomal, non-chromosomal and synthetic nucleic acid sequences, eg. derivatives of plant viruses; bacterial plasmids; derivatives of the Ti plasmid from *Agrobacterium tumefaciens*; derivatives of the Ri plasmid from *Agrobacterium rhizogenes*; phage DNA; yeast artificial

chromosomes; bacterial artificial chromosomes; binary bacterial artificial chromosomes; vectors derived from combinations of plasmids and phage DNA. However, any other vector may be used as long as it is replicable or integrative or viable in the plant cell.

[0068] The regulatory element and terminator may be of any suitable type and may be endogenous to the target plant cell or may be exogenous, provided that they are functional in the target plant cell.

[0069] Preferably the regulatory element is a promoter. A variety of promoters which may be employed in the vectors of the present invention are well known to those skilled in the art. Factors influencing the choice of promoter include the desired tissue specificity of the vector, and whether constitutive or inducible expression is desired and the nature of the plant cell to be transformed (eg. monocotyledon or dicotyledon). Particularly suitable promoters include the Cauliflower Mosaic Virus 35S (CaMV 35S) promoter, the maize Ubiquitin promoter, the rice Actin promoter, and ryegrass endogenous OMT, 4CL, CCR or CAD promoters.

[0070] A variety of terminators which may be employed in the vectors of the present invention are also well known to those skilled in the art. The terminator may be from the same gene as the promoter sequence or a different gene. Particularly suitable terminators are polyadenylation signals, such as the CaMV 35S polyA and other terminators from the nopaline synthase (nos) and the octopine synthase (ocs) genes.

[0071] The vector, in addition to the regulatory element, the nucleic acid or nucleic acid fragment of the present invention and the terminator, may include further elements necessary for expression of the nucleic acid or nucleic acid fragment, in different combinations, for example vector backbone, origin of replication (ori), multiple cloning sites, spacer sequences, enhancers, introns (such as the maize Ubiquitin Ubi intron), antibiotic resistance genes and other selectable marker genes [such as the neomycin phosphotransferase (npt2) gene, the hygromycin phosphotransferase (hph) gene, the phosphinothricin acetyltransferase (bar or pat) gene], and reporter genes (such as beta-glucuronidase (GUS) gene (gusA)]. The vector may also contain a ribosome binding site for translation initiation. The vector may also include appropriate sequences for amplifying expression.

[0072] As an alternative to use of a selectable marker gene to provide a phenotypic trait for selection of transformed host cells, the presence of the vector in transformed cells may be determined by other techniques well known in the art, such as PCR (polymerase chain reaction), Southern blot hybridisation analysis, histochemical GUS assays, northern and Western blot hybridisation analyses.

[0073] Those skilled in the art will appreciate that the various components of the vector are operatively linked, so as to result in expression of said nucleic acid or nucleic acid fragment. Techniques for operatively linking the components of the vector of the present invention are well known to those skilled in the art. Such techniques include the use of linkers, such as synthetic linkers, for example including one or more restriction enzyme sites.

[0074] The vectors of the present invention may be incorporated into a variety of plants, including monocotyledons (such as grasses from the genera *Lolium, Festuca, Cynodon, Bracharia, Paspalum, Panicum, Miscanthus, Pennisetum, Phalaris*, and other forage, turf and bioenergy grasses, corn, oat, sugarcane, wheat and barley), dicotyledons (such as *Arabidopsis*, tobacco, legumes, Alfalfa, oak, *Eucalyptus*, maple, *Populus*, canola, soybean and chickpea) and gymnosperms (such as *Pinus*). In a preferred embodiment, the vectors are used to transform monocotyledons, preferably grass species such as *Lolium, Festuca, Cynodon, Bracharia, Paspalum, Panicum, Miscanthus, Pennisetum, Phalaris*, and other forage, turf and bioenergy grasses, more preferably a *Lolium* species such as *Lolium perenne* or *Lolium arundinaceum*, including cultivars for forage and turf applications.

[0075] Techniques for incorporating the vectors of the present invention into plant cells (for example by transduction, transfection or transformation) are well known to those skilled in the art. Such techniques include *Agrobacterium* mediated introduction, electroporation to tissues, cells and protoplasts, protoplast fusion, injection into reproductive organs, injection into immature embryos and high velocity projectile introduction to cells, tissues, calli, immature and mature embryos. The choice of technique will depend largely on the type of plant to be transformed.

[0076] Cells incorporating the vector of the present invention may be selected, as described above, and then cultured in an appropriate medium to regenerate transformed plants, using techniques well known in the art. The culture conditions, such as temperature, pH and the like, will be apparent to the person skilled in the art. The resulting plants may be reproduced, either sexually or asexually, using methods well known in the art, to produce successive generations of transformed plants.

[0077] In a further aspect of the present invention there is provided a transformed plant cell, plant, plant seed or other plant part, or plant biomass, including digestible biomass such as hay, including, eg transformed with, a nucleic acid, genetic construct or vector of the present invention. Preferably the transgenic plant cell, plant, plant seed or other plant part is produced by a method according to the present invention.

[0078] The present invention also provides a transgenic plant, plant seed or other plant part, or plant biomass, derived from a plant cell of the present invention and including a nucleic acid, genetic construct or vector of the present invention.

[0079] The present invention also provides a transgenic plant, plant seed or other plant, part or plant biomass, derived from a plant of the present invention and including a nucleic acid, genetic construct or vector of the present invention.

[0080] The nucleic acid, genetic construct or vector of the present invention may be stably integrated into the genome of the plant, plant seed, other plant part or plant biomass.

[0081] The plant cell, plant, plant seed or other plant part may be from any suitable species, including monocotyledons, dicotyledons and gymnosperms. In a preferred embodiment the plant cell, plant, plant seed or other plant part may be from a monocotyledon, preferably a grass species, such as *Lolium, Festuca, Cynodon, Bracharia, Paspalum, Panicum, Miscanthus, Pennisetum, Phalaris*, and other forage, turf and bioenergy grasses, more preferably a *Lolium* species such as *Lolium perenne* or *Lolium arundinaceum.*

[0082] In a further aspect of the present invention there is provided a method of modifying lignin biosynthesis in a plant, said method including introducing into said plant an effective amount of a nucleic acid or nucleic acid fragment, genetic construct and/or a vector according to the present invention.

[0083] By "an effective amount" is meant an amount sufficient to result in an identifiable phenotypic trait in said plant,

or a plant, plant seed or other plant part, or plant biomass derived therefrom. Such amounts can be readily determined by an appropriately skilled person, taking into account the type of plant, the route of administration and other relevant factors. Such a person will readily be able to determine a suitable amount and method of administration. See, for example, Maniatis et al, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, the entire disclosure of which is incorporated herein by reference.

[0084] Using the methods and materials of the present invention, plant lignin biosynthesis may be increased, decreased or otherwise modified relative to an untransformed control plant. It may be increased or otherwise modified, for example, by incorporating additional copies of a sense nucleic acid or nucleic acid fragment of the present invention. It may be decreased, for example, by incorporating an antisense nucleic acid or nucleic acid fragment of the present invention or by incorporating a functionally active fragment or variant which is capable of modifying lignin biosynthesis in a plant via sense suppression. In addition, the number of copies of genes encoding for different enzymes in the lignin biosynthetic pathway may be manipulated to modify the relative amount of each monolignol synthesized, thereby leading to the formation of lignin having altered composition.

[0085] Accordingly, in a preferred embodiment of this aspect of the invention there is provided a method of modifying lignin biosynthesis in a plant, said method including introducing into said plant in sense orientation an effective amount of a nucleic acid, genetic construct or vector according to the present invention, such that expression of the corresponding gene is suppressed.

[0086] Preferred functionally active fragments and variants for sense suppression include those hereinbefore described.

[0087] In a further aspect of the present invention there is provided use of a nucleic acid, genetic construct or vector according to the present invention for sense suppression of lignin biosynthesis in a plant.

[0088] In a still further aspect of the present invention there is provided use of a nucleic acid or nucleic acid fragment according to the present invention, and/or nucleotide sequence information thereof, and/or single nucleotide polymorphisms thereof, as a molecular genetic marker.

[0089] More particularly, nucleic acids or nucleic acid fragments according to the present invention, and/or nucleotide sequence information thereof, and/or single nucleotide polymorphisms thereof, may be used as a molecular genetic marker for qualitative trait loci (QTL) tagging, mapping, DNA fingerprinting and in marker assisted selection, and may be used as candidate genes or perfect markers, particularly in ryegrasses and fescues. Even more particularly, nucleic acids or nucleic acid fragments according to the present invention, and/or nucleotide sequence information thereof, may be used as molecular genetic markers in forage and turf grass improvement, eg. tagging QTLs for dry matter digestibility, herbage quality, mechanical stress tolerance, disease resistance, insect pest resistance, plant stature and leaf and stem colour.

[0090] In a still further aspect of the present invention there is provided a substantially purified or isolated polypeptide from a ryegrass (*Lolium*) or fescue (Fustuca) species, selected from the group consisting of the enzymes 4CL, CCR and CAD.

[0091] The ryegrass (*Lolium*) or fescue (*Festuca*) species may be of any suitable type, including Italian or annual ryegrass, perennial ryegrass, tall fescue, meadow fescue and red fescue. Preferably the species is a ryegrass, more preferably perennial ryegrass *L. perenne*).

[0092] In a preferred embodiment of this aspect of the invention, the substantially purified or isolated enzyme 4CL includes an amino acid sequence selected from the group consisting of sequences shown in FIGS. **2**, **3** and **4** hereto (Sequence ID Nos: 2, 4 and 6, respectively); and functionally active fragments and variants thereof.

[0093] In a further preferred embodiment of this aspect of the invention, the substantially purified or isolated enzyme CCR includes an amino acid sequence selected from the group consisting of the sequence shown in FIG. **10** hereto (Sequence ID No: 8); and functionally active fragments and variants thereof.

[0094] In a still further preferred embodiment of this aspect of the invention, the substantially purified or isolated enzyme CAD includes an amino acid sequence selected from the group consisting of the sequence shown in FIGS. **13**, **14**, **26** and **27** hereto (Sequence ID Nos: 10, 12, 15 and 17, respectively); and functionally active fragments and variants thereof.

[0095] By "functionally active" in this context is meant that the fragment or variant has one or more of the biological properties of the enzymes 4CL, CCR and CAD, respectively. Additions, deletions, substitutions and derivatizations of one or more of the amino acids are contemplated so long as the modifications do not result in loss of functional activity of the fragment or variant. Preferably the fragment or variant has at least approximately 60% identity to the relevant part of the above mentioned sequence, more preferably at least approximately 80% identity, most preferably at least approximately 90% identity. Such functionally active variants and fragments include, for example, those having conservative amino acid substitutions of one or more residues in the corresponding amino acid sequence. Preferably the fragment has a size of at least 10 amino acids, more preferably at least 15 amino acids, most preferably at least 20 amino acids.

[0096] In a further embodiment of this aspect of the invention, there is provided a polypeptide recombinantly produced from a nucleic acid or nucleic acid fragment according to the present invention. Techniques for recombinantly producing polypeptides are well known to those skilled in the art.

[0097] In a still further aspect of the present invention there is provided a lignin or modified lignin substantially or partially purified or isolated from a plant, plant seed or other plant part of the present invention.

[0098] Such lignins may be modified from naturally occurring lignins in terms of the length, the degree of polymerisation (number of units), degree of branching and/or nature of linkages between units.

[0099] In a still further aspect, the present invention provides an isolated regulatory element capable of causing expression of an exogenous gene in plant cells. Preferably the regulatory element is isolated from a nucleic acid or nucleic acid fragment encoding OMT, 4CL, CCR or CAD.

[0100] The regulatory element may be a nucleic acid molecule, including DNA (such as cDNA or genomic DNA) and RNA (such as mRNA) that is single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases, and combinations thereof.

[0101] Preferably the regulatory element includes a promoter, more preferably an O-methyltransferase promoter, even more preferably an O-methyltransferase promoter from a ryegrass (*Lolium*) or fescue (*Festuca*) species, more preferably a ryegrass, most preferably perennial ryegrass (*Lolium perenne*).

[0102] In a particularly preferred embodiment of this aspect of the invention, the regulatory element includes a promoter from the caffeic acid O-methyltransferase gene corresponding to the cDNA homologue LpOMT1 from perennial ryegrass.

[0103] Preferably the regulatory element includes a nucleotide sequence including the first approximately 4630 nucleotides of the sequence shown in FIG. **18** hereto (Sequence ID No: 13); or a functionally active fragment or variant thereof.

[0104] By "functionally active" in this context is meant that the fragment or variant (such as an analogue, derivative or mutant) is capable of causing expression of a transgene in plant cells. Such variants include naturally occurring allelic variants and non-naturally occurring variants. Additions, deletions, substitutions and derivatizations of one or more of the nucleotides are contemplated so long as the modifications do not result in loss of functional activity of the regulatory element. Preferably the functionally active fragment or variant has at least approximately 80% identity to the relevant part of the above sequence, more preferably at least approximately 90% identity, most preferably at least approximately 95% identity. Preferably the fragment has a size of at least 100 nucleotides, more preferably at least 150 nucleotides, most preferably at least 200 nucleotides.

[0105] In a particularly preferred embodiment of this aspect of the invention, the regulatory element includes a nucleotide sequence selected from the group consisting of:

[0106] Nucleotides −4581 to −1
[0107] Nucleotides −4285 to −1
[0108] Nucleotides −4020 to −1
[0109] Nucleotides −2754 to −1
[0110] Nucleotides −1810 to −1
[0111] Nucleotides −831 to −1
[0112] Nucleotides −560 to −1
[0113] Nucleotides −525 to −1
[0114] Nucleotides −274 to −1
[0115] Nucleotides −21 to −1
[0116] of FIG. **18** hereto (Sequence ID No: 13);
[0117] or a functionally active fragment or variant thereof.

[0118] In another preferred embodiment the regulatory element includes a 4 coumarate-CoA ligase promoter, even more preferably a 4 coumarate-CoA ligase promoter from a ryegrass (*Lolium*) or fescue (*Festuca*) species, more preferably a ryegrass, most preferably perennial ryegrass (*Lolium perenne*).

[0119] In a particularly preferred embodiment of this aspect of the invention, the regulatory element includes a promoter from the 4 coumarate-CoA ligase gene corresponding to the cDNA homologue Lp4CL2 from perennial ryegrass.

[0120] Preferably the regulatory element includes a nucleotide sequence including the first approximately 2206 nucleotides of the sequence shown in FIG. **38** hereto (Sequence ID No: 17); or a functionally active fragment or variant thereof.

[0121] By "functionally active" in this context is meant that the fragment or variant (such as an analogue, derivative or mutant) is capable of causing expression of a transgene in plant cells. Such variants include naturally occurring allelic variants and non-naturally occurring variants. Additions, deletions, substitutions and derivatizations of one or more of the nucleotides are contemplated so long as the modifications do not result in loss of functional activity of the regulatory element. Preferably the functionally active fragment or variant has at least approximately 80% identity to the relevant part of the above sequence, more preferably at least approximately 90% identity, most preferably at least approximately 95% identity. Preferably the fragment has a size of at least 100 nucleotides, more preferably at least 150 nucleotides, most preferably at least 200 nucleotides.

[0122] In a particularly preferred embodiment of this aspect of the invention, the regulatory element includes a nucleotide sequence selected from the group consisting of:

[0123] Nucleotides −2206 to −1
[0124] Nucleotides −1546 to −1
[0125] Nucleotides −1186 to −1
[0126] Nucleotides −406 to −1
[0127] Nucleotides −166 to −1
[0128] of FIG. **38** hereto (Sequence ID No: 17);
[0129] or a functionally active fragment or variant thereof.

[0130] In another preferred embodiment the regulatory element includes a cinnamoyl-CoA reductase promoter, even more preferably a cinnamoyl-CoA reductase promoter from a ryegrass (*Lolium*) or fescue (*Festuca*) species, more preferably a ryegrass, most preferably perennial ryegrass (*Lolium perenne*).

[0131] In a particularly preferred embodiment of this aspect of the invention, the regulatory element includes a promoter from the cinnamoyl-CoA reductase gene corresponding to the LpCCR1 cDNA from perennial ryegrass.

[0132] Preferably the regulatory element includes a nucleotide sequence including the first approximately 6735 nucleotides of the sequence shown in FIG. **39** hereto (Sequence ID No: 18); or a functionally active fragment or variant thereof.

[0133] By "functionally active" in this context is meant that the fragment or variant (such as an analogue, derivative or mutant) is capable of causing expression of a transgene in plant cells. Such variants include naturally occurring allelic variants and non-naturally occurring variants. Additions, deletions, substitutions and derivatizations of one or more of the nucleotides are contemplated so long as the modifications do not result in loss of functional activity of the regulatory element. Preferably the functionally active fragment or variant has at least approximately 80% identity to the relevant part of the above sequence, more preferably at least approximately 90% identity, most preferably at least approximately 95% identity. Preferably the fragment has a size of at least 100 nucleotides, more preferably at least 150 nucleotides, most preferably at least 200 nucleotides.

[0134] In a particularly preferred embodiment of this aspect of the invention, the regulatory element includes a nucleotide sequence selected from the group consisting of:

[0135] Nucleotides −6735 to −1
[0136] Nucleotides −5955 to −1
[0137] Nucleotides −5415 to −1
[0138] Nucleotides −4455 to −1
[0139] Nucleotides −4035 to −1
[0140] Nucleotides −3195 to −1
[0141] Nucleotides −2595 to −1
[0142] Nucleotides −1755 to −1
[0143] Nucleotides −1275 to −1
[0144] Nucleotides −495 to −1
[0145] Nucleotides −255 to −1
[0146] Nucleotides −75 to −1
[0147] of FIG. **39** hereto (Sequence ID No: 18);
[0148] or a functionally active fragment or variant thereof.

[0149] By an "exogenous gene" is meant a gene not natively linked to said regulatory element. In certain embodiments of the present invention the exogenous gene is also not natively found in the relevant plant or plant cell.

[0150] The exogenous gene may be of any suitable type. The exogenous gene may be a nucleic acid such as DNA (e.g. cDNA or genomic DNA) or RNA (e.g. mRNA), and combinations thereof. The exogenous gene may correspond to a target gene, for example a gene capable of influencing disease resistance, herbage digestibility, nutrient quality, mineral content or drought tolerance or be a fragment or variant (such as an analogue, derivative or mutant) thereof which is capable of modifying expression of said target gene. Such variants include nucleic acid sequences which are antisense to said target gene or an analogue, derivative, mutant or fragment thereof. The transgene may code for a protein or RNA sequence depending the target condition and whether down or up-regulation of gene expression is required. Preferably, the target gene is selected from exogenous coding sequences coding for mRNA for a protein, this protein may be of bacterial origin (such as enzymes involved in cell wall modification and cell wall metabolism, cytokinin biosynthesis), or eukaryotic origin (such as pharmaceutically active polypeptides) or of plant origin (such as enzymes involved in the synthesis of phenolic compounds, cell wall metabolism, sugar metabolism, lignin biosynthesis). Preferably, the target gene is selected from the group comprising O-methyltransferase, 4 coumarate CoA-ligase, cinnamoyl CoA reductase, cinnamyl alcohol dehydrogenase, cinnamate 4 hydroxylase, phenolase, laccase, peroxidase, coniferol glucosyl transferase, coniferin beta-glucosidase, phenylalanine ammonia lyase, ferulate 5-hydroxylase, chitinase, glucanase, isopentenyltransferase, xylanase.

[0151] The plant cells, in which the regulatory element of the present invention is capable of causing expression of an exogenous gene, may be of any suitable type. The plant cells may be from monocotyledons (such as grasses from the genera *Lolium, Festuca, Cynodon, Bracharia, Paspalum, Panicum, Miscanthus, Pennisetum, Phalaris*, and other forage and turf grasses, corn, grains, oat, sugarcane, wheat and barley), dicotyledons (such as *Arabidopsis*, tobacco, legumes, Alfalfa, oak, *Eucalyptus*, maple, *Populus*, canola, soybean and chickpea) and gymnosperms (such as *Pinus*). Preferably the plant cells are from a monocotyledon, more preferably a grass species such as *Lolium, Festuca, Cynodon, Bracharia, Paspalum, Panicum, Miscanthus, Pennisetum, Phalaris*, and other forage, turf and bioenergy grasses, more preferably a *Lolium* species such as *Lolium perenne* or *Lolium arundinaceum.*

[0152] The regulatory element according to the present invention may be used to express exogenous genes to which it is operatively linked in the production of transgenic plants.

[0153] Accordingly, in a further aspect of the present invention there is provided a vector including a regulatory element according to the present invention.

[0154] In a preferred embodiment of this aspect of the invention, the vector may include a regulatory element according to the present invention, an exogenous gene as hereinbefore described, and a terminator; said regulatory element, exogenous gene and terminator being operatively linked, such that said regulatory element is capable of causing expression of said exogenous gene in plant cells. Preferably, said regulatory element is upstream of said exogenous gene and said terminator is downstream of said exogenous gene.

[0155] The vector may be of any suitable type and may be viral or non-viral. The vector may be an expression vector. Such vectors include chromosomal, non-chromosomal and synthetic nucleic acid sequences, eg. derivatives of plant viruses; bacterial plasmids; derivatives of the Ti plasmid from *Agrobacterium tumefaciens*; derivatives of the Ri plasmid from *Agrobacterium rhizogenes*; phage DNA; yeast artificial chromosomes; bacterial artificial chromosomes; binary bacterial artificial chromosomes; vectors derived from combinations of plasmids and phage DNA. However, any other vector may be used as long as it is replicable on integrative or viable in the plant cell.

[0156] The terminator may be of any suitable type and includes for example polyadenylation signals, such as the Cauliflower Mosaic Virus 35S polyA (CaMV 35S polyA) and other terminators from the nopaline synthase (nos) and the octopine synthase (ocs) genes.

[0157] The vector, in addition to the regulatory element, the exogenous nucleic acid and the terminator, may include further elements necessary for expression of the nucleic acid, in different combinations, for example vector backbone, origin of replication (ori), multiple cloning sites, spacer sequences, enhancers, introns (such as the maize Ubiquitin Ubi intron), antibiotic resistance genes and other selectable marker genes [such as the neomycin phosphotransferase (npt2) gene, the hygromycin phosphotransferase (hph) gene, the phosphinothricin acetyltransferase (bar or pat) gene], and reporter genes (such as beta-glucuronidase (GUS) gene (gusA)]. The vector may also contain a ribosome binding site for translation initiation. The vector may also include appropriate sequences for amplifying expression.

[0158] The regulatory element of the present invention may also be used with other full promoters or partial promoter elements.

[0159] As an alternative to use of a selectable marker gene to provide a phenotypic trait for selection of transformed host cells, the presence of the vector in transformed cells may be determined by other techniques well known in the art, such as PCR (polymerase chain reaction), Southern blot hybridisation analysis, histochemical GUS assays, northern and Western blot hybridisation analyses.

[0160] Those skilled in the art will appreciate that the various components of the vector are operatively linked, so as to result in expression of said transgene. Techniques for operatively linking the components of the vector of the present invention are well known to those skilled in the art. Such techniques include the use of linkers, such as synthetic linkers, for example including one or more restriction sites.

[0161] The vectors of the present invention may be incorporated into a variety of plants, including monocotyledons, dicotyledons and gymnosperms. In a preferred embodiment the vectors are used to transform monocotyledons, preferably grass species such as ryegrasses (*Lolium* species) and fescues (*Festuca* species), more preferably perennial ryegrass (*Lolium perenne*) including cultivars for forage and turf applications.

[0162] Techniques for incorporating the vectors of the present invention into plant cells (for example by transduction, transfection or transformation) are well known to those skilled in the art. Such techniques include *Agrobacterium* mediated introduction, electroporation to tissues, cells and protoplasts, protoplast fusion, injection into reproductive organs, injection into immature embryos and high velocity projectile introduction to cells, tissues, calli, immature and

mature embryos. The choice of technique will depend largely on the type of plant to be transformed.

[0163] Cells incorporating the vector of the present invention may be selected, as described above, and then cultured in an appropriate medium to regenerate transformed plants, using techniques well known in the art. The culture conditions, such as temperature, pH and the like, will be apparent to the person skilled in the art. The resulting plants may be reproduced, either sexually or asexually, using methods well known in the art, to produce successive generations of transformed plants.

[0164] In a further aspect of the present invention there is provided a plant cell, plant, plant seed or other plant part, including, eg. transformed with, a vector of the present invention.

[0165] The plant cell, plant, plant seed or other plant part may be from any suitable species, including monocotyledons, dicotyledons and gymnosperms. In a preferred embodiment the plant cell, plant, plant seed or other plant part is from a monocotyledon, preferably a grass species, more preferably a ryegrass (Lolium species) or fescue (Festuca species), even more preferably perennial ryegrass (Lolium perenne), including cultivars for forage and turf applications.

[0166] The present invention also provides a plant, plant seed, or other plant part derived from a plant cell of the present invention.

[0167] The present invention also provides a plant, plant seed or other plant part derived from a plant of the present invention.

[0168] In a still further aspect of the present invention there is provided a recombinant plant genome including a regulatory element according to the present invention.

[0169] In a preferred embodiment of this aspect of the invention the recombinant plant genome further includes an exogenous gene operatively linked to said regulatory element.

[0170] In a further aspect of the present invention there is provided a method for expressing an exogenous gene in plant cells, said method including introducing into said plant cells an effective amount of a regulatory element and/or a vector according to the present invention.

[0171] By "an effective amount" is meant an amount sufficient to result in an identifiable phenotypic change in said plant cells or a plant, plant seed or other plant part derived therefrom. Such amounts can be readily determined by an appropriately skilled person, taking into account the type of plant cell, the route of administration and other relevant factors. Such a person will readily be able to determine a suitable amount and method of administration. See, for example, Maniatis et al, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, the entire disclosure of which is incorporated herein by reference.

[0172] The present invention will now be more fully described with reference to the accompanying Examples and drawings. It should be understood, however, that the description following is illustrative only and should not be taken in any way as a restriction on the generality of the invention described above.

[0173] In the Figures

[0174] FIG. **1** shows plasmid maps of the three cDNAs encoding perennial ryegrass 4CL homologues.

[0175] FIG. **2** shows the nucleotide (Sequence ID No: 1) and amino acid (Sequence ID No: 2) sequences of Lp4CL1.

[0176] FIG. **3** shows the nucleotide (Sequence ID No: 3) and amino acid (Sequence ID No: 4) sequences of Lp4CL2.

[0177] FIG. **4** shows the nucleotide (Sequence ID No: 5) and amino acid (Sequence ID No: 6) sequences of Lp4CL3.

[0178] FIG. **5** shows amino acid sequence alignment of deduced proteins encoded by Lp4CL1 (Sequence ID No: 2), Lp4CL2 (Sequence ID No: 4) and Lp4CL3 (Sequence ID No: 6).

[0179] FIG. **6** shows northern hybridisation analysis of developing perennial ryegrass using Lp4CL1, Lp4CL2 and Lp4CL3 as hybridisation probes. SR: roots from seedlings (3-5 d post-germination), SS: shoots from seedlings (3-5 d post-germination), ML: leaves from 12-week-old plants, MS: stems from 12-week-old plants. Blots were washed in 0.2× SSPE, 0.1% SDS at 65° C. Lp4CL1, Lp4CL2 and Lp4CL3 do not cross hybridise at this stringency. Sizes are given in kb.

[0180] FIG. **7** shows northern hybridisation analysis showing the time course of expression of 4CL mRNA in wounded perennial ryegrass leaves. Sizes are given in kb.

[0181] FIG. **8** shows genomic Southern hybridisation analysis using Lp4CL1, Lp4CL2 and Lp4CL3 as hybridisation probes. 10 μg of digested perennial ryegrass genomic DNA or 20 μg of digested tall fescue genomic DNA were separated on a 1.0% agarose gel, transferred to Hybond N^+ membranes and then hybridised with ^{32}P labelled Lp4CL1, Lp4CL2 or Lp4CL3 probes. The ryegrass Lp4CL1, Lp4CL2 and Lp4CL3 genes reveal homologous sequences in tall fescue and indicate that the ryegrass 4CL genes can be used to isolate and to manipulate the expression of the tall fescue (*Festuca arundinacea*) 4CL genes.

[0182] FIG. **9** shows restriction map of LpCCR1. An *L. perenne* seedling cDNA library constructed in Uni-ZAP™ (Stratagene) was screened in a solution containing 10×PIPES, 50% deionised formamide and 10% SDS at 42° C. Filters were washed at room temperature, three times in 0.1% SDS, 2×SSPE and then twice in 0.1% SDS, 0.2×SSPE. The location of the probe used for northern and Southern hybridisation analyses is indicated by the black line labelled LpCCR531.

[0183] FIG. **10** shows the nucleotide (Sequence ID No: 7) and amino acid (Sequence ID No: 8) sequences of LpCCR1.

[0184] FIG. **11** shows Southern hybridisation analysis of DNA from double haploid (DH) perennial ryegrass using LpCCR1 as hybridisation probe. 10 μg of DH genomic DNA was digested with DraI, BamHI, EcoRI, EcoRV, HindIII or XbaI, separated on a 1% agarose gel and then capillary blotted onto nylon membrane (Amersham Hybond-N). The membrane was probed with the digoxigenin (DIG) labelled LpCCR531 fragment at 25 ng/ml in the hybridisation solution. Hybridisation was in 4×SSC, 50% formamide, 0.1% N-Lauroyl-sarcosine, 0.02% SDS, 2% Blocking solution at 42° C. The membrane was washed twice for five minutes in 2×SSC, 0.1% SDS at room temperature, then twice for fifteen minutes in 0.5×SSC, 0.1% SDS at 68° C. Molecular weight was determined by comparison to a DIG-labelled marker (Roche Molecular Biochemicals).

[0185] FIG. **12** shows northern hybridisation analysis of RNA samples from different organs and developmental stages of perennial ryegrass using LpCCR1 probe. Roots from seedlings (3-5 d post-germination), shoots from seedlings (3-5 d post-germination), roots from seedlings (7-10 d post-germination), leaves from seedlings (7-10 d post-germination), roots from 6 and 10 week old plants, leaves from 6

and 10 week old plants, stems from 6 and 10 week old plants, whole seedling from 11 day old Phalaris and 7 day old Festuca.

[0186] Total RNA was isolated using Trizol (GibcoBRL) and 15 μg was separated on a 1.2% Agarose gel containing 6% formamide and then capillary blotted onto nylon membrane (Amersham Hybond-N). The membrane was stained with 0.2% methylene blue/0.3M sodium acetate to visualise the marker and ensure that RNA was evenly loaded. 50 ng LpCCR531 was random-labelled with ^{32}P-dCTP (Amersham Megaprime) and hybridisation conditions were 4×SSC, 50% formamide, 0.5% SDS, 5×denhardt solution, 5% dextrane sulphate, 0.1% Herring sperm DNA at 42° C. over-night. The ryegrass LpCCR1 gene reveal homologous transcripts in tall fescue and Phalaris, thus indicating that the ryegrass CCR gene can be used to manipulate the expression of the tall fescue (Festuca arundinacea) and Phalaris CCR endogenous genes.

[0187] FIG. **13** shows the nucleotide (Sequence ID No: 9) and amino acid (Sequence ID No: 10) sequences of LpCAD1.

[0188] FIG. **14** shows the nucleotide (Sequence ID No: 11) and amino acid (Sequence ID No: 12) sequences of LpCAD2.

[0189] FIG. **15** shows a plasmid map of a cDNA clone encoding perennial ryegrass CAD homologue LpCAD1.

[0190] FIG. **16** shows northern hybridisation analysis of RNA samples from different organs and developmental stages of perennial ryegrass using A) LpCAD1 and B) LpCAD2 as hybridisation probes. Roots from seedlings 3-5 d post-germination, 7-10 d post-germination, 6 weeks and 10 weeks, Shoots from seedlings 3-5 d post-germination and 7-10 d post-germination, Leaves from 6 week old and 10 week old plants, stem tissue from 6 and 10 week old plants. RNA isolated, from *Phalaris* and *Festuca* 11 and 7 day old seedlings. The ryegrass CAD genes reveal homologous transcripts in tall fescue and Phalaris, thus indicating that the ryegrass CAD gene can be used to manipulate the expression of the tall fescue and Phalaris CAD endogenous genes.

[0191] FIG. **17** shows genomic Southern hybridisation analysis. 10 μg of perennial ryegrass genomic DNA digested with a range of restriction enzymes was separated on a 0.8% agarose gel, transferred to Hybond N and then hybridised with a DIG labelled A) LpCAD1, and B) LpCAD2 hybridisation probe.

[0192] FIG. **18** shows the nucleotide sequence of the LpOmt1 promoter (Sequence ID No: 13).

[0193] FIG. **19** shows a plasmid map of plant transformation vector carrying the reporter β-glucuronidase (GUS) gene (gusA) under control of the perennial ryegrass LpOmt1 promoter.

[0194] FIG. **20** (upper image) shows PCR analysis of transgenic tobacco plants containing the gusA gene under the control of the perennial ryegrass LpOMT1 promoter (upper figure). PCR reactions using gusA-specific primers were performed. FIG. **20** (lower images) show histochemical GUS assays, demonstrating xylem-specific gusA expression (A and B) and gusA expression in glandular leaf trichomes (C and D) in transgenic tobacco plants containing the gusA gene under the control of the perennial ryegrass LpOMT1 promoter.

[0195] FIG. **21** shows the isolation of the LpCCR1 genomic clone 1. A) Southern hybridization analysis of CCR genomic clone λLp6.1.1a digested with XbaI, NcoI, SalI, XhoI, XhoI/SalI DNA was separated on a 0.8% agarose gel, transferred to Hybond N and hybridized with a DIG labelled CCR1 probe. B) Map showing the genomic gene organisation of LpCCR1 clone 1 based on sequence results. C) Comparison of plant CCR exon size and number in different plant species (*Lolium perenne*, Lp., *Eucalyptus gunni*, Eg., *Eucalyptus saligna*, Es., *Populus balsamifera*, Pb.)

[0196] FIG. **22** shows the isolation of the LpCCR1 genomic clone 2. A) Southern hybridization analysis of CCR genomic clone λLp6.1.1a digested with XbaI, NcoI, SalI, XhoI, XhoI/SalI DNA was separated on a 0.8% agarose gel, transferred to Hybond N and hybridized with 200 bp of the CCR1 promoter (FIG. **21**B). B) Map showing the promoter region of LpCCR1 clone 2 based on sequence results.

[0197] FIG. **23** shows the isolation of an Lp4CL genomic clone. A) Southern hybridisation analysis of 4CL genomic clone λLp4CL2 digested with BamHI, KpnI or SalI. DNA was separated on a 0.8% agarose gel, transferred to Hybond N and hybridized with a DIG labelled 4CL1 hybridisation probe. B) 10 μl of a standard PCR reaction using forward and reverse oligonucleotides designed to positions outlined on C). The PCR products were separated on a 0.8% agarose gel and stained with ethidium bromide. C) Map showing the genomic gene organisation of λLp4CL2 based on sequence and PCR results.

[0198] FIG. **24** shows the isolation of an Lp4CL genomic clone. A) Southern hybridisation analysis of 4CL genomic clone λLp4CL2 digested with BamHI, KpnI, SalI. DNA was separated on a 0.8% agarose gel, transferred to Hybond N and hybridized with a DIG labelled 4CL1 probe. B) Map showing the genomic gene organisation of Lp4CL2 clone 1 and the promoter region of clone 2.

[0199] FIG. **25** shows plasmid map of plant transformation vector carrying the gusA gene under control of the perennial ryegrass Lp4CL2 promoter (Lp4CL2::gusA).

[0200] FIG. **26** shows nucleotide (Sequence ID No: 14) and amino acid (Sequence ID No: 15) sequences of genomic clone CAD2 cv Barlano (Intron 1 and first 111 by of the coding region are missing).

[0201] FIG. **27** shows nucleotide (Sequence ID No: 16) and amino acid (Sequence ID No:15) sequences of coding sequence deduced from genomic clone CAD2 cv Barlano (region in bold is missing from the genomic clone).

[0202] FIG. **28** shows the isolation of LpCAD2 genomic clone. A) Southern hybridization analysis of CAD genomic clone λLpCAD2 digested with BamHI, EcoRI, KpnI, SalI or XbaI. DNA was separated on a 0.8% agarose gel, transferred to Hybond N and hybridized with a DIG labelled CAD2 hybridisation probe. B) Map showing the genomic gene organisation of λLpCAD2 based on sequence results.

[0203] FIG. **29** shows A) Sense and antisense Lp4CL1, Lp4CL2 and Lp4CL3 transformation vectors under control of the CaMV 35S promoter; B) Sense and antisense Lp4CL1, Lp4CL2 and Lp4CL3 transformation vectors under control of the maize ubiquitin promoter.

[0204] FIG. **30** shows A) Sense and antisense LpCCR1 transformation vectors under control of the CaMV 35S promoter; B) Sense and antisense LpCCR1 transformation vectors under control of the maize ubiquitin promoter.

[0205] FIG. **31** shows A) Sense and antisense LpCAD1 transformation vectors under control of the CaMV 35S promoter; B) Sense and antisense LpCAD1 transformation vectors under control of the maize ubiquitin promoter.

[0206] FIG. **32** shows molecular analysis of Lp4CL1-transgenic tobacco. A) Plasmid map of transformation vector carrying a chimeric sense Lp4CL1 gene. B) PCR analysis of

independent transgenic tobacco clones using Lp4CL1 specific primers. C) Southern hybridization analysis of independent transgenic tobacco plants using an Lp4CL1 specific probe. D) Northern hybridization analysis of independent transgenic tobacco plants using an Lp4CL1 specific probe.

[0207] FIG. **33** shows molecular analysis of LpCCR1-transgenic tobacco. A) Plasmid map of transformation vectors carrying a chimeric sense and antisense LpCCR1 gene. B) PCR analysis of independent sense transgenic tobacco clones using LpCCR1 specific primers.

[0208] FIG. **34** shows protocol for suspension culture-independent production of transgenic perennial ryegrass plants. A) Isolated zygotic embryos, plated on MSM5 medium, day 0; B) Embryogenic callus formation and proliferation, 6-8 weeks after embryo isolation; C) Embryogenic calli arranged on high osmotic MSM3Plus medium prior to biolistic transformation; D) Histochemical GUS assay showing GUS expressing foci 3-4 days post-bombardment of chimeric gusA gene; E) Selection of embryogenic calli on MSM3 medium containing 100 mg/l paromomycin (Pm), 2 weeks after microprojectile bombardment; F) Regeneration of Pm resistant shoots on MSK medium containing 100 mg/l Pm, 4 weeks after microprojectile bombardment; G) In vitro plant regeneration from PM resistant embryogenic calli, 6 weeks after microprojectile bombardment; H) Transgenic perennial ryegrass plants 28 weeks after embryo isolation.

[0209] FIG. **35** shows molecular analysis of transgenic perennial ryegrass plants carrying sense and antisense LpOmt1 transgenes. Plasmid maps of vectors used for the co-transformation of perennial ryegrass embryogenic calli; pHP23 carrying a chimeric neomycin phosphotransferase (npt2) selectable marker gene; pUbiomt1 carrying a maize ubiquitin promoter driven sense LpOmt1 gene; pUbitmo1 carrying a maize ubiquitin promoter driven antisense LpOmt1 gene (top). PCR analysis using npt2-specific primers of 5 independent transgenic perennial ryegrass plants from biolistic transformation with sense and antisense LpOmt1 vectors (upper centre). Southern hybridization analysis with an omt1 hybridization probe of 7 independent perennial ryegrass plants co-transformed with sense (lanes 1-3) and antisense (lanes 4-7) LpOmt1 vectors (lower centre left). Southern hybridisation analysis with an npt2 hybridisation probe of independent perennial ryegrass plants (lower centre right). Northern hybridisation analysis of perennial ryegrass plants co-transformed with antisense LpOmt1 vector (bottom). C=negative control untransformed perennial ryegrass; P=positive plasmid control.

[0210] FIG. **36** shows biochemical analysis of LpOmt1-transgenic perennial ryegrass. OMT activity of leaf samples from selected independent LpOmt1-transgenic perennial ryegrass plants (EII8, EII11, EII14 and EII15) was determined and compared to untransformed perennial ryegrass negative control plant *L. perenne* cv. Ellett (wild type). Mean values and standard deviations of replicate assays are shown.

[0211] FIG. **37** shows PCR screening of transgenic ryegrass plants. PCR analysis using npt2-specific primers of 8 independent transgenic perennial ryegrass plants from biolistic transformation with antisense LpUbi4CL2 vector.

[0212] FIG. **38** shows the nucleotide sequence of genomic clone 4CL2 from perennial ryegrass (Sequence ID No: 17).

[0213] FIG. **39** shows the nucleotide sequence of genomic clone CCR1 from perennial ryegrass (Sequence ID No: 18).

[0214] FIG. **40** shows the map location of Lp4CL1, Lp4CL3, LpCAD1, LpCAD2, LpCCR1, LpOMT1 and LpOMT2 (in bold) within the genetic linkage map of perennial ryegrass.

[0215] FIG. **41**. Illustration of the Gateway-derived expression vectors used for generating the constructs for expressing perennial ryegrass lignin biosynthetic genes.

[0216] FIG. **42**. Vector details of Gateway™ Entry clone for the LpCAD3 cDNA.

[0217] FIG. **43**. Vector details of Gateway™ Entry clone for the promoter LpCAD2.

[0218] FIG. **44**. Vector details of Gateway™ Entry clone for the terminator LpCAD2.

[0219] FIG. **45**. Plasmid map of Construct 1, LpCAD2p::LpCAD3::LpCAD2t in vector pAUX3132.

[0220] FIG. **46**. Plasmid map of Construct 2, LpCCR1::LpCCR1::LpCCR1 in vector pAUX3169.

[0221] FIG. **47**. Sequence of LpCCR1 gene (SEQ ID No: 19) and modified forward primer (SEQ ID No: 20) that imparts a single base deletion in the LpCCR1 gene.

[0222] FIG. **48**. Plasmid map of Construct 3, LpCCR1::LpCCR1(fs)::LpCCR1 in vector pAUX3169.

[0223] FIG. **49**. Vector details for pAcH1 construct that was used as the plant selectable marker containing the expression construct Act1D::hph::35S.

[0224] FIG. **50**. Production of transgenic perennial ryegrass from microprojectile bombardment of embryogenic calli derived from immature inflorescences. A) Excised immature inflorescence of perennial ryegrass; 2-3 mm; B-E) Induction and proliferation of embryogenic calli; 1-8 weeks after inflorescence excision. F). Distribution of embryogenic calli on high osmotic medium LP3-OS medium prior to biolisitic transformation; G) Biolistic transformation device, PDS-1000/He; H-I) Growth and development of hygromycin-resistant shoots, 30-75 days post bombardment; J) Growth and development of hygromycin-resistant shoots in vitro; K) Hygromycin-resistant plants established in soil and grown under containment glasshouse conditions.

[0225] FIG. **51**. Production of transgenic perennial ryegrass from microprojectile bombardment of embryogenic calli derived from seedling meristems. A) In vitro shoot culture for basal meristem isolation; regenerated from seedling meristem-derived calli; B) Distribution of basal meristematic material on callus initiation medium; C-E) Induction and proliferation of embryogenic calli from shoot meristems of *Lolium perenne*; F) Distribution of embryogenic calli on high osmotic medium prior to biolistic transformation; G) Biolistic transformation device, PDS-1000/He; H-I) Growth and development of hygromycin-resistant shoots, 30-84 days post bombardment; J) Growth and development of hygromycin-resistant shoots in vitro; K) Hygromycin-resistant plants established in soil and grown under containment glasshouse conditions.

[0226] FIG. **52**. Flowchart describing the transformation method used to generate transgenic perennial ryegrass containing the expression construct of interest and the selectable marker gene (hph).

[0227] FIG. **53**. Amplification of the hygromycin phosphotransferase (hph) gene by Q-PCR in samples of genomic DNA extracted from putative transgenic perennial ryegrass regenerated after co-bombardment with plasmids pAcH1 and pAUX3132-LpCAD2::LpCAD3::LpCAD2.

[0228] FIG. **54**. Southern analysis of genomic DNA digested with Eco R1 (R1) and separated by agarose gel

electrophoresis and the transgene detected with either hph or Ubi promoter probes. All six putative transgenic plants were confirmed to contain both hph and the gene-of-interest, hpLpCCR1.

[0229] FIG. **55**. Mäule staining of cross-sectioned internodes from wild type and transgenic pUbi::hpCCR1::35 S ryegrass at R1 and R2 stage shows a strong decrease of reddish colour in transgenic plants which may suggest a decrease in S lignin content compared to wild type plants.

[0230] FIG. **56**. Total lignin content of perennial ryegrass internodes at the R1 developmental stage shows a progressive reduction in lignin content from internode 1 (base) to internode 5 (top).

[0231] FIG. **57**. Example of a gas chromatogram (GC-MS) showing separation and identification of G-lignin and S-lignin monomers after thioacidolysis derivatisation of lignin extracted from wild type perennial ryegrass.

EXAMPLE 1

[0232] Isolation and characterisation of three 4-Coumarate CoA-Ligase (4CL) cDNAs from *Lolium perenne*

Materials and Methods

Plant Material

[0233] Plants and embryogenic cell suspensions of perennial ryegrass (Lolium perenne L.) cv Ellet and tall fescue (Festuca arundinacea Schreb.) cv Triumph were established and maintained as previously described (Heath et al., 1998). Wounding experiments were performed with 10-day-old seedlings of perennial ryegrass (cv Ellet) as previously described (Heath et al., 1998).

Screening of a cDNA Library

[0234] A cDNA library prepared with RNA isolated from perennial ryegrass seedlings (Heath et al., 1998) was screened with a [^{32}P]dCTP-labelled rice partial 4CL probe. The rice 4CL probe and consisted of a 844 by 4CL specific sequence inserted into PUC119. This insert has 93% sequence identity with a rice 4CL cDNA sequence (Genbank, L43362, bases 453-1300). cDNA inserts were excised and recircularized using the ExAssist helper phage with SOLR strain (Stratagene) as described by the manufacturer.

DNA Sequencing

[0235] cDNA clones were digested with 8 restriction enzymes (BamHI, EcoRI, KpnI, NotI, PstI, SalI, XbaI, XhoI) and selected clones were sequenced on both strands by the dideoxy chain termination method using M13 forward and reverse primers. For sequencing the internal regions of Lp4CL1, Lp4CL2 and Lp4CL3 synthetic oligonucleotide primers were designed from the DNA sequences previously determined. Sequencing was performed using the ABI dye terminator kit and automatic sequencer. Nucleotide sequences were aligned using the SeqEd program (ABI) and further analysis was performed using the HIBIO DNASIS vs2 program (Hitachi Software Engineering).

Genomic DNA Blot Analysis

[0236] Genomic DNA was isolated from single genotype-derived cell suspensions of perennial ryegrass and tall fescue according to Lichtenstein and Draper (1985). Ten of perennial ryegrass DNA and 20 μg of tall fescue DNA was digested with each of the restriction enzymes HindIII and XbaI, separated on 1% agarose gels, and transferred to Hybond N+membranes according to the manufacturer's instructions (Amersham). Probes consisted of BamHI/KpnI fragments of Lp4CL1 (1771 bp), Lp4CL2 (2034 bp) or Lp4CL3 (2080 bp) labelled using the Megaprime labelling kit (Amersham) and [^{32}P]dCTP. Hybridization was performed at 65° C. in 5×SSPE, 5×Denhardt's solution, 0.5% (w/v) SDS, and 200 μg/mL denatured herring sperm DNA. Membranes were washed three times in 2×SSPE, 0.1% SDS for 10 min at 25° C. and then twice in 0.1 X SSPE, 0.1% SDS for 20 min at 65° C.

RNA Blot Analysis

[0237] Total RNA (10 μg) was separated on 1.2% formaldehyde gels and transferred to Hybond N (Amersham) membranes according to the manufacturers instructions. Membranes were stained with 0.2% methylene blue to confirm correct loading and transfer of RNA. Hybridisation was performed at 42° C. in 5×SSPE, 5×Denhart's solution, 0.5% SDS, 50% deionized formamide, 200 μg/mL denatured herring sperm DNA. Preparation of probes and washing of membranes was as for DNA blot analysis except for the tall fescue Northern blot when the final two washes were performed with 0.1×SSPE, 0.1% SDS for 10 min at 42° C.

Results

[0238] Isolation and Sequence Analysis of Perennial Ryegrass 4CL cDNAs

[0239] A cDNA library prepared from RNA extracted from perennial ryegrass seedlings was screened with a rice 4CL hybridization probe and ten cDNAs were isolated from 2×10^5 pfu. The cDNAs were characterised by restriction analysis with 8 restriction enzymes. All clones were full length (approximately 2.0-2.2 kb) with poly(A) tails and could be separated into three groups: Lp4CL1 (four clones) Lp4CL2 (five clones) and Lp4CL3 (one clone). Plasmid maps for Lp4CL1, Lp4CL2 and Lp4CL3 are shown (FIG. **1**). Lp4CL1, Lp4CL2 and Lp4CL3 were fully sequenced (FIGS. **2**, **3** and **4**, respectively).

[0240] Lp4CL1 is 2284 by long with an open reading frame (ORF) of 1710 bp, a 5' noncoding region of 322 by and a 3' noncoding region of 252 by including a poly(A) tail. Lp4CL2 is 1992 by long with an ORF of 1668 bp, a 5' noncoding region of 61 by and a 3' noncoding region of 263 by including a poly(A) tail. Lp4CL3 is 2038 by long with an ORF of 1671 bp, a 5' noncoding region of 112 by and a 3' noncoding region of 255 by including a poly(A) tail.

[0241] Within the coding region, Lp4CL1 has 70% nucleic acid sequence identity with both Lp4CL2 and Lp4CL3, while Lp4CL2 has 79% sequence identity with Lp4CL3. There is little sequence homology in the 3' noncoding regions between clones (52-55%).

Amino Acid Sequence Comparisons

[0242] The putative proteins encoded by the three cDNAs consist of 570 amino acids [60290 u (Da)] for Lp4CL1, 556 amino acids (59238 u) for Lp4CL2 and 557 amino acids (59735 u) for Lp4CL3. The deduced amino acid sequences of Lp4CL1, Lp4CL2 and Lp4CL3 are shown (FIG. **5**). Lp4CL2 and Lp4CL3 share 79% amino acid sequence identity, Lp4CL1 and Lp4CL2 have 61% amino acid sequence identity, while Lp4CL1 and Lp4CL3 have only 58% amino acid sequence identity. Regions of high sequence homology are

more prevalent in the central and c-terminal regions of the enzyme. For example the sequence identity between amino acids 208 to 568 of each enzyme is 85% for Lp4CL2 and Lp4CL3, 72% for Lp4CL1 and Lp4CL2 and 67% for Lp4CL1 and Lp4CL3.

[0243] Lp4CL1, Lp4CL2 and Lp4CL3 share several common regions with other plant 4CLs. In particular, they contain the putative AMP-binding domain and the conserved GEICIRG motif, except for Lp4CL3 where the second isoleucine has been replaced with valine (FIG. **5**). It has been proposed that domain II is associated with the catalytic activity of 4CL. Also, four Cys residues conserved in plant 4CLs are conserved in Lp4CL1, Lp4CL2 and Lp4CL3 (FIG. **5**). These results suggest that the *L. perenne* cDNAs encode three divergent 4CL enzymes that are likely to have originated from three different 4CL genes.

Expression of Perennial Ryegrass 4CL Genes

[0244] Lp4CL1, Lp4CL2 and Lp4CL3 were used as hybridization probes in Northern blots with RNA prepared from different organs of perennial ryegrass at two developmental stages. All three probes hybridized to a single mRNA species of approximately 2.2-2.3 kb. Lp4CL1, Lp4CL2 and Lp4CL3 were expressed at both seedling and mature stages of development and in all organs tested. For Lp4CL2 and Lp4CL3 the strongest signal was found in RNA samples from seedling roots and mature stems (FIG. **6**).

[0245] Lp4CL1, Lp4CL2 and Lp4CL3 were also used as hybridization probes in Northern blots with RNA prepared from tall fescue. All three probes hybridized to a similar mRNA species (2.3 kb) as that in perennial ryegrass (FIG. **6**). The strongest signal was found in RNA samples from mature stems with weaker signals in RNA from roots and seedling shoots. No expression of Lp4CL1, Lp4CL2 or Lp4CL3 was observed in leaves. The three probes varied in their ability to hybridize to the corresponding homologues in tall fescue, with Lp4CL3 resulting in the highest signal and Lp4CL1 hybridizing only weakly.

[0246] To determine whether 4CL could be induced under stress conditions, leaves of perennial ryegrass seedlings were wounded. No increase in the transcript level upon wounding was observed with Lp4CL1, Lp4CL2 or Lp4CL3 (FIG. **7**).

Genomic organization of perennial ryegrass 4CL genes

[0247] Perennial ryegrass DNA was digested with two restriction enzymes, HindIII or XbaI. Restriction sites for these enzymes are not present in the cDNA sequence of Lp4CL1, Lp4CL2 or Lp4CL3. When Lp4CL1, Lp4CL2 or Lp4CL3 was used as a probe, several DNA hybridizing fragments of varying intensity were revealed (FIG. **8**). Each probe hybridized to a unique set of fragments, suggesting that Lp4CL1, Lp4CL2 and Lp4CL3 represent three different genes. Furthermore, Lp4CL1 and Lp4CL2 hybridized to 2 to 3 major fragments per digest which may represent either alleles of the same gene or indicate the presence of more than one gene in each class. The Lp4CL1, Lp4CL2 and Lp4CL3 probes also revealed several different size hybridizing DNA fragments in genomic Southern blots from tall fescue under high stringency conditions (FIG. **8**), suggesting that three similar 4CL genes are present in *F. arundinacea.*

EXAMPLE 2

Isolation and Characterisation of a Cinnamoyl CoA Reductase (CCR) cDNA From *Lolium Perenne*

[0248] A total of 500,000 phage were screened from a cDNA library constructed from ten-day-old etiolated *L. perenne* seedlings using a maize CCR probe. Ninety-three positive plaques were observed in the primary screen and five were subsequently analysed by restriction enzyme digestion. Four out of the five were identical. One of the four identical cDNAs, LpCCR1, was selected for further analysis (FIG. **9**).

Nucleic Acid Sequence Analysis of Perennial Ryegrass CCR cDNA

[0249] The full nucleotide sequence of LpCCR1 was obtained and the amino acid sequence predicted (FIG. **10**). LpCCR1 is a 1395 by cDNA with 149 by of 5' non-coding region and 160 by of 3' non-coding region. An open reading frame of 1086 by encodes a protein of 362 amino acids. The composition of the coding region was found to be 68% G+C rich. Codon usage was also examined and found to be biased towards XXC/G codons (94%), with XCG and XUA codons accounting for only 9% and 0.55% respectively. G+C richness and bias towards G and C in the third position of a codon triplet are previously reported characteristics of monocot genes.

Genomic Organization of Perennial Ryegrass CCR Gene

[0250] The number of CCR genes present in the ryegrass genome was determined by Southern blot analysis of genomic DNA from double haploid plants, using as probe a fragment of the LpCCR1 cDNA (LpCCR531, FIG. **9**). Double haploid DNA reduces the complexity associated with allelic variation. Genomic DNA was cut with enzymes that do not cut the cDNA internally; DraI, BamHI, EcoRI, EcoRV, HindIII and XbaI, and the membrane was hybridised and washed under medium-stringency conditions. A single strongly hybridising band was evident in each lane (FIG. **11**) indicating that there is a single copy of the LpCCR1 gene in the perennial ryegrass genome.

Expression of Perennial Ryegrass CCR Gene

[0251] To investigate the expression profile of the CCR gene in ryegrass, northern hybridisation analysis was carried out with total RNA extracted from roots and shoots at seedling growth stages (0.5-1 cm and 4-6 cm shoots) and roots, stem and leaves at mature growth stages (6 and 10 weeks). Seedlings were grown on filter paper in the dark at 25° C. and then transferred to soil and glasshouse conditions (25° C.) until the 6 and 10-week stages. Whole seedling total RNA from Festuca and Phalaris was included in the northern analysis. Hybridisation with LpCCR531 (FIG. **9**) was performed at medium-stringency and the membrane was then washed at high-stringency. A transcript of approximately 1.5 kb was detected in all tissues, the level of expression varying with maturity and from one tissue type to another (FIG. **12**). The LpCCR1 transcript appears to be more abundant in roots and stem than shoots and leaves. In the stem, transcript abundance increases from 6-weeks to 10-weeks; indicating that transcription in stem tissue is up-regulated as the plant matures. Expression was found predominantly in tissues such as stems and roots that are forming secondary cell walls indicating that LpCCR1 is constitutively involved in lignification.

EXAMPLE 3

Isolation and Characterisation of Cinnamyl Alcohol Dehydrogenase (CAD) cDNAs from *Lolium Perenne*

[0252] A 558 by cinnamyl alcohol dehydrogenase (CAD) fragment was amplified from cDNA synthesised from total

RNA prepared from perennial ryegrass seedlings. The conserved amino acid domains between *Pinus radiata, Medicago sativa, Aralia cordata, Eucalyptus botryoides* and *Arabidopsis thaliana* CADs were used to design oligonucleotides for the amplification of the perennial ryegrass CAD. The forward oligonucleotide was designed to the conserved amino acid domain CAGVTVYS and the reverse oligonucleotide to the conserved domain DVRYRFV. The 551 by PCR fragment was cloned and sequenced to confirm that it corresponded to a perennial ryegrass CAD PCR fragment. A cDNA library prepared from RNA extracted from perennial ryegrass seedlings was screened with the 551 bp PCR fragment specific for perennial ryegrass CAD. Eight cDNAs were isolated and separated into six groups by restriction digest analysis. One representative clone each from two groups (LpCAD1, LpCAD2) were selected for further characterisation.

Nucleic Acid Sequence Analysis of Perennial Ryegrass CAD cDNAs

[0253] The complete sequence of the perennial ryegrass CAD homologue LpCAD1 was determined (FIG. **13**). The 1325 by clone had a poly (A) tail, typical start and stop codons and the open reading frame (ORF) of this clone coded for a putative protein of 408 amino acids.

[0254] The complete nucleotide sequence of the perennial ryegrass CAD homologue LpCAD2 was also determined (FIG. **14**).

Expression Of Perennial Ryegrass CAD Genes

[0255] A northern hybridisation analysis with RNA samples isolated from perennial ryegrass at different developmental stages hybridised with the full length LpCAD1 1325 by cDNA (FIG. **15**) was performed to determine patterns of organ and developmental expression. The probe hybridised to a single mRNA species of approximately 1.6 kb. The LpCAD1 transcript was expressed in all tissue tested: roots, shoots, stem and leaves (FIG. **16**A). The LpCAD1 transcript was most abundant in root tissue and the mature stem, this expression pattern is typical of a gene involved in the lignification of plant cell walls. Intergeneric homologies were revealed in Festuca and Phalaris.

[0256] A similar northern hybridisation analysis was performed with LpCAD2 (FIG. **16**B), however the transcript was found to be most abundant in mature stem tissue and the shoots.

Genomic Organization of Perennial Ryegrass CAD Genes

[0257] A Southern hybridisation analysis using DNA samples isolated from a perennial ryegrass double haploid plant digested with DraI, BamHI, EcoRI, EcoRV, HindIII and XbaI and hybridised with a 500 by LpCAD1 probe was performed. The hybridisation pattern at high stringency revealed the presence of two prominent bands for most digests indicating that LpCAD1 belongs to a small gene family and exists a muliticopy gene in perennial ryegrass (FIG. **17**A).

[0258] A similar Southern hybridization analysis was performed with LpCAD2 (FIG. **17**B) the hybridisation pattern at high stringency revealed the presence of one or two prominent bands for most digests indicating that LpCAD2 exists as a single copy gene or a member of a small gene family in perennial ryegrass (FIG. **17**B).

EXAMPLE 4

Isolation and Characterisation of Genomic Clones and Promoters for O-Methyltransferase (OMT), Cinnamoyl-CoA Reductase (CCR), 4 Coumarate CoA-Ligase (4CL) and Cinnamyl Alcohol Dehydrogenase (CAD) from *Lolium Perenne*

[0259] Genomic clones and promoters of O-methyltransferase (OMT), cinnamoyl-CoA reductase (CCR), 4 coumarate CoA-ligase (4CL) and cinnamyl alcohol dehydrogenase (CAD) were isolated from a perennial ryegrass genomic library using the corresponding cDNAs as hybridisation probes.

Isolation and Characterisation of Genomic Clones and Promoters for Perennial Ryegrass O-Methyltransferase (OMT)

[0260] A perennial ryegrass genomic library was screened with the cDNA clone, LpOmt1, (Heath et al. 1998) encoding O-methyltransferase (OMT). The sequence of the 5' untranslated region and the coding region was found to be identical to that of the LpOmt1 cDNA previously isolated. The entire 4.8 kb genomic clone was fully sequenced (FIG. **18**).

[0261] To further characterise the promoters, transcriptional fusions of the promoter sequence to the β-glucuronidase (GUS) coding sequence (gusA) have been generated (FIG. **19**). Direct gene transfer experiments to tobacco protoplasts were performed with the corresponding chimeric genes to transgenically express them in a heterologous system for in planta expression pattern analysis by histochemical GUS assays. A set of transgenic tobacco plants carrying a chimeric gusA gene under the control of the 5' regulatory region of the LpOmt1 promoter was generated to assess the potential use of the LpOmt1 promoter for xylem-specificity and targeted downregulation of genes encoding key lignin biosynthetic enzymes.

[0262] The transgenic tobacco plants generated using the LpOmt1 promoter driven chimeric gusA transformation vector were screened by PCR and histochemical GUS assays.

[0263] A PCR screening was undertaken using gusA specific primers for the initial identification of transgenic tobacco plants (FIG. **20**). PCR positive tobacco plants were screened by histochemical GUS assays for in planta expression pattern analysis (FIG. **20**).

Isolation and Characterisation of Genomic Clones and Promoters for Perennial Ryegrass Cinnamoyl-CoA Reductase (CCR)

[0264] A CCR genomic clone from perennial ryegrass was isolated containing 6.5 kb of promoter and the entire gene organisation (intron/exon boundaries). The CCR promoter can be used for targeted expression of foreign genes in transgenic plants.

[0265] A perennial ryegrass genomic library was screened with the cDNA clone LpCCR1 which codes for the lignin biosynthetic enzyme, cinnamyl-CoA reductase (CCR). Four different genomic clones were identified based on restriction digest analysis. Clone 6.1.1a was selected for further analysis. A 6.42 kb XhoI fragment from clone 6.1.1a, which hybridized strongly to the LpCCR1 cDNA probe, was subcloned into pBluescriptSK (FIG. **21**A). Sequence analysis revealed that the 6.42 kb XhoI fragment contained the entire LpCCR1 gene and 200 by of promoter region. The intron/exon boundaries are illustrated in FIG. **21**B, the location and

the size of the exons appear to be conserved in other CCRs from different species (FIG. **21**C).

[0266] To isolate the promoter region of LpCCR1, the Southern blot containing digested phage genomic DNA isolated from clone λLp6.1.1a was reprobed with the 200 bp promoter region. The probe hybridized strongly to a 6.5 kb SalI fragment. This genomic fragment LpCCR1 clone 2, was subcloned into pBluescriptSK and sequenced (FIG. **22**A). Sequence results revealed that the 6.5 kb SalI fragment contained 6.5 kb of promoter (FIG. **22**B). The full sequence of LpCCR1 genomic clone containing the promoter and entire gene sequence (exons and introns) was obtained and is shown on FIG. **39**.

Isolation and Characterisation of Genomic Clones and Promoters for Perennial Ryegrass 4 Coumarate CoA-Ligase (4CL)

[0267] A 4CL2 genomic clone from perennial ryegrass was isolated containing 2.5 kb of promoter and partial gene organisation (intron/exon boundaries). The 4CL2 promoter can be used for targeted expression of foreign genes in transgenic plants. The 2.5 kb promoter has been fused to the reporter gene gusA for expression analysis.

[0268] A perennial ryegrass genomic library was screened with an Lp4CL cDNA probe. After tertiary screening positive 4CL genomic clones were obtained and characterised by restriction digest and Southern hybridisation analysis (FIG. **23**A).

[0269] Sequence analysis revealed that the isolated 4CL genomic clone (4CL2) from perennial ryegrass had 100% nucleotide identity to the Lp4CL2 cDNA clone. To further characterise this 5 kb λLp4CL2 genomic clone and to confirm that it corresponds to the cDNA of Lp4CL2, a number of PCR reactions using primers designed to the cDNA were used. PCR results confirmed that the 5 kb genomic fragment was a partial genomic clone corresponding to the Lp4CL2 cDNA (FIG. **23**B). Using primer combinations F1 and R1 the entire 4.8 kb genomic fragment was amplified. To determine the location of introns additional PCR reactions using the primer combinations F1/R2 and F2/R1 were performed, a 1 kb and 3.5 kb bands were amplified respectively. The location and size of the introns could be determined from these results, and further confirmed by sequence analysis. This large 5 kb genomic fragment contains 4 small exons representing the coding sequence of Lp4CL2 between 508 by and 1490 by (FIG. **23**C).

[0270] The genomic clone 1, Lp4CL2 contained no promoter region. To isolate the promoter region of Lp4CL2, the Southern blot containing digested phage genomic DNA isolated from clone λLp4CL2 was reprobed with a 300 by EcoRI/BglI isolated from the 5' end of the cDNA clone Lp4CL2. The 300 by probe hybridised strongly to a 2.5 kb BamHI fragment. This genomic fragment Lp4CL2 clone 2, was subcloned into pBluescriptSK and sequenced (FIG. **24**A). Sequence results revealed that the 2.5 kb BamHI fragment contained the 508 by of the 5' ORF of Lp4CL2 missing from genomic clone 1 and 2.0 kb of promoter region (FIG. **24**B). The full sequence of the Lp4CL2 genomic clone containing the promoter and partial gene sequence (exons and introns) was obtained and is shown on FIG. **39**.

[0271] The promoter from Lp4CL2 was thus isolated and used for the production of a chimeric gusA reporter gene (FIG. **25**).

Isolation and Characterisation of Genomic Clones and Promoters for Perennial Ryegrass Cinnamyl Alcohol Dehydrogenase (CAD)

[0272] A CAD genomic clone from perennial ryegrass was isolated containing the gene organisation (intron/exon boundaries) minus intron 1 containing the first 111 by of the CAD coding region. The genomic clone has allowed the identification of a G at position 851 by in the coding region of the CAD2 genomic clone isolated from perennial ryegrass cv. Barlano which is absent in the CAD2 cDNA clone isolated from perennial ryegrass cv. Ellett. The SNP (single nucleotide polymorphism) found to exist between the 2 cultivars has the potential utility as a molecular marker for herbage quality, dry matter digestibility, mechanical stress tolerance, disease resistance, insect pest resistance, plant stature and leaf and stem colour.

[0273] Results below show the isolation of the genomic clone and sequence analysis of deduced coding sequence from the genomic clone CAD2 from perennial ryegrass cv. Barlano compared to the truncated cDNA CAD2 from the cv Ellett. The missing G in the perennial ryegrass cv. Ellett has been highlighted (FIGS. **26** and **27**).

[0274] A perennial ryegrass genomic library was screened with a probe corresponding to the 5' end of the LpCAD2 cDNA clone, which codes for the lignin biosynthetic enzyme cinnamyl alcohol dehydrogenase. Ten positive plaques were identified and isolated in the primary library screening. After a secondary and tertiary screening, two positive plaques were obtained and corresponding positive genomic clones were further characterised by restriction digest and Southern hybridization analyses. Both genomic clones were found to be identical based on restriction digest analyses. One clone, named λLpCAD2 was chosen for further Southern hybridization analyses. A 4.5 kb BamHI fragment which hybridized strongly to the LpCAD2 cDNA probe was subcloned into pBluescriptSK and sequenced (FIG. **28**A). Sequence analysis revealed that the 4.5 kb BamHI fragment was a partial genomic clone of LpCAD2. This large 4.5 kb genomic fragment contains 4 small exons representing the coding sequence of LpCAD2 between 213 by and the stop codon at 1213 bp, and the location of the intron/exon boundaries are illustrated in FIG. **28**B.

EXAMPLE 5

Development of Transformation Vectors Containing Chimeric Genes with 4CL, CCR and CAD cDNA Sequences from Perennial Ryegrass

[0275] To alter the expression of the key enzymes involved in lignin biosynthesis 4CL, CCR and CAD, through antisense and/or sense suppression technology and for over-expression of these key enzymes in transgenic plants, a set of sense and antisense transformation vectors was produced. Transformation vectors containing chimeric genes using perennial ryegrass 4CL, CCR and CAD cDNAs in sense and antisense orientations under the control of either the CaMV 35S or the maize ubiquitin promoter were generated (FIGS. **29**, **30** and **31**).

EXAMPLE 6

Production and Characterisation of Transgenic Tobacco Plants Expressing Chimeric 4CL, CCR and CAD Genes from Perennial Ryegrass

[0276] A set of transgenic tobacco plants carrying chimeric 4CL, CCR and CAD genes from perennial ryegrass were produced and analysed.

[0277] Transformation vectors with Lp4CL1, Lp4CL2 and Lp4CL3 full length cDNA sequences in sense and antisense orientations under the control of either the CaMV 35S or the maize ubiquitin promoters were generated. Transformation vectors with LpCCR1 cDNA in both sense and antisense orientation under the control of either the CaMV 35S and maize ubiquitin promoters were generated. Transformation vectors with 1325 by full length LpCAD1 cDNA in sense and 1051 by partial LpCAD1 cDNA in antisense orientation under the control of either the CaMV 35S and maize ubiquitin promoters were generated.

[0278] Direct gene transfer experiments to tobacco protoplasts were performed using these transformation vectors.

[0279] The production and molecular analysis of transgenic tobacco plants carrying the perennial ryegrass Lp4CL1 and LpCCR1 cDNAs under the control of the constitutive CaMV 35S promoter is described here in detail.

[0280] A set of transgenic tobacco plants generated using the Lp4CL1 sense transformation vector was screened by PCR and subjected to Southern and northern hybridization analyses.

[0281] A PCR screening was undertaken using npt2 and Lp4CL1 specific primers for the initial identification of transgenic tobacco plants. Independent transgenic tobacco plants were identified to be co-transformed with both the selectable marker npt2 and the Lp4CL1 chimeric genes (FIG. **32**).

[0282] Southern hybridisation analysis was performed with DNA samples from PCR positive transgenic tobacco plants to demonstrate the integration of the chimeric Lp4CL1 transgene in the tobacco plant genome. Independent transgenic tobacco plants carried between 1 and 5 copies of the Lp4CL1 transgene. No cross-hybridization was observed between the endogenous tobacco 4CL gene and the perennial ryegrass hybridization probe used (FIG. **32**).

[0283] Northern hybridization analysis using total RNA samples prepared from the transgenic tobacco plants carrying the chimeric sense Lp4CL1 transgene and probed with the Lp4CL1-specific hybridization probe revealed the presence of a 1.2 kb Lp4CL1 transcript strongly expressed in one Lp4CL1-transgenic tobacco plant analysed (FIG. **32**).

[0284] The sense and antisense transformation vectors of LpCCR1 under the control of the CaMV 35S promoter were introduced into tobacco protoplasts via direct gene transfer. A set of transgenic tobacco plants was generated and screened by PCR with specific primers to identify transgenic tobacco plants carrying chimeric LpCCR1 transgene. The molecular analysis of LpCCR1-transgenic tobacco plants is shown (FIG. **33**).

EXAMPLE 7

Production and Characterisation of Transgenic Perennial Ryegrass Plants Expressing Chimeric OMT, 4CL, CCR and CAD Genes from Perennial Ryegrass

[0285] An improved transformation method was developed for the production of transgenic perennial ryegrass plants by biolistic transformation of embryogenic cells. Transgenic perennial ryegrass plants were generated using chimeric OMT, 4CL, CCR and CAD genes from perennial ryegrass and the improved transformation method.

Improved Method for the Production of Transgenic Perennial Ryegrass Plants

[0286] This improved procedure utilises embryogenic calli produced from mature seed-derived embryos as direct targets for biolistic transformation without requiring the establishment of embryogenic cell suspensions. The protocol relies on a continuous supply of isolated zygotic embryos for callus induction. Transgenic ryegrass plants can be regenerated 24-28 weeks after embryo isolation (FIG. **34**). Isolated embryos are plated onto MSM5 medium to produce embryogenic calli suitable as targets for biolistic transformation within 8 weeks. The embryogenic calli, treated on high-osmoticum medium MSM3 Plus prior to microprojectile bombardment, are selected on MSM3 medium containing 100 mg/l paromomycin (Pm) for 2 weeks before being transferred onto MSK with 100 mg/l Pm for further 4 weeks until differentiation of Pm resistant shoot appear. Regenerated shoots are transferred on to fresh selective media MSK with 100 mg/l Pm for a further 4 weeks (FIG. **34**).

Production of Transgenic Perennial Ryegrass Plants Expressing Chimeric OMT, 4CL, CCR and CAD Genes from Perennial Ryegrass

[0287] Transgenic perennial ryegrass (*Lolium perenne*) plants were generated using chimeric ryegrass OMT, 4CL, CCR and CAD genes by biolistic transformation of embryogenic calli. Examples of the production and detailed molecular analysis of these transgenic ryegrass plants are described.

[0288] Transgenic perennial ryegrass plants for OMT down-regulation were produced using biolistic transformation of embryogenic calli and plant transformation vectors pUbiomt1 and pUbitmo1 carrying LpOmt1 cDNA sequence in sense and antisense orientation under control of the constitutive maize ubiquitin promoter. These transgenic perennial ryegrass plants for down-regulated OMT activity were regenerated from paromomycin resistant calli obtained from biolistic transformation using microprojectilies coated with two plasmids; pHP23 (carrying the chimeric npt2 gene as the selectable marker) and either the sense or antisense LpOmt1 transformation vector driven by the maize Ubi promoter.

[0289] Transgenic perennial ryegrass plants were subjected to a polymerase chain reaction (PCR) screening using npt2-specific primers. Independent npt2 PCR-positive transgenic perennial ryegrass plants obtained from biolistic transformation of embryogenic calli-generated from approximately 60,000 isolated mature seed-derived embryos—using LpOmt1 sense (pUbiomt1) and LpOmt1 antisense (pUbitmo1) transformation vectors were identified [16 pUbiomt1 transgenic plants and 27 pUbitmo1 transgenic plants] (FIG. **35**).

[0290] Southern hybridization analysis was performed with undigested and HindIII -digested DNA samples prepared from the PCR positive transgenic perennial ryegrass plants, to demonstrate their transgenic nature and the integration of the chimeric npt2 and LpOmt1 transgenes. Independent transgenic perennial ryegrass plants co-transformed with both, the selectable marker npt2 gene and LpOmt1 chimeric genes, were identified (FIG. **35**). In most instances, the transgenic perennial ryegrass plants recovered contained multiple copies of the selectable marker gene including rearranged transgene copies. No npt2-hybridizing bands were detected in the untransformed negative control.

[0291] Samples of HindIII-digested genomic DNA were included in the analysis when the LpOmt1 gene-specific hybridization probe (omt1) was used. The omt1probe hybridized to a number of bands in DNA samples corresponding to both, the transgenic plants and the untransformed negative control. The omt1-hybridizing bands shared in all samples correspond to endogenous LpOmt1 gene sequences represented as a small multigene family in the perennial ryegrass genome (Heath et al. 1998). The different omt1-hybridizing bands evident in the samples from the transgenic plants and absent in the untransformed negative control sample correspond to antisense (tmo1) and sense (omt1) LpOmt1 transgene integration events (FIG. **35**).

[0292] Northern hybridization analysis using strand-specific LpOmt1 probes allowed the identification of transgenic perennial ryegrass plants expressing the antisense LpOmt1 transgene (FIG. **35**).

[0293] The OMT activity of selected antisense and sense LpOmt1 transgenic perennial ryegrass plants was determined. Biochemical assays for OMT activity were initially established in untransformed plants (such as tobacco and perennial ryegrass). The assays utilise radiolabelled S-adenosylmethionine as the methyl donor for the OMT-catalysed conversion of caffeic acid into ferulic acid. The production of radioactive ferulic acid is measured and allows the OMT activity to be determined.

[0294] The OMT activity of selected LpOmt1-transgenic perennial ryegrass plants (*L. perenne* cv. Ellett) was determined. Significantly altered OMT activity in individual transformation events was observed (FIG. **36**). The manipulation of OMT activity in transgenic perennial ryegrass plants due to the expression of the chimeric ryegrass LpOmt1 gene was thus demonstrated.

[0295] Transgenic perennial ryegrass plants were recovered, using biolistic transformation of embryogenic calli, for the manipulation of the expression of genes encoding the key lignin biosynthetic enzyme, 4CL. The plant transformation vectors pUbi4CL2 and pUbi2LC4 carrying chimeric Lp4CL2 cDNA sequences in sense and antisense orientation, respectively, driven by the constitutive maize ubiquitin (Ubi) promoter were used. Perennial ryegrass plants for 4CL manipulation were regenerated from Pm-resistant calli obtained from biolistic transformation of embryogenic calli using microprojectiles coated with the plasmids pHP23, carrying a chimeric npt2 gene as selectable marker gene and the antisense pUbi2LC4.

[0296] Transgenic perennial ryegrass plants were subjected to a polymerase chain reaction (PCR) screening using npt2-specific primers. Independent npt2 PCR-positive transgenic perennial ryegrass plants were obtained from biolistic transformation of embryogenic calli (FIG. **37**).

[0297] Transgenic perennial ryegrass plants were also recovered, using biolistic transformation of embryogenic calli, for the manipulation of the expression of genes encoding the key lignin biosynthetic enzymes, CCR and CAD.

EXAMPLE 8

Genetic Mapping of Perennial Ryegrass OMT, 4CL, CCR and CAD Genes

[0298] Lp4CL1, Lp4CL3, LpCAD1, LpCAD2, LpCCR1, LpOMT1 and LpOMT2 clones were PCR amplified and radio-labelled for use as probes to detect restriction fragment length polymorphisms (RFLPs). RFLPs were mapped using 110 progeny individuals of the p150/112 perennial ryegrass reference population restricted with the enzymes described in Table 3 below.

TABLE 3

Mapping of RFLPs

Clones	Polymorphic in p150/112	Enzyme mapped with	Locus	Linkage group
Lp4CL1	Y	DraI	Lp4CL1	2
Lp4CL3	Y	EcoRV	Lp4CL3	6
LpCAD1	Y	EcoRV	LpCAD1	2
LpCAD1.2.1	Y	EcoRI	LpCAD2a	7
			LpCAD2b	—
			LpCAD2c	2
LpCCR1	Y	EcoRI	LpCCR1	7
LpOMT1	Y	DraI	LpOMT1	7
LpOMT2	Y	EcoRV	LpOMT2	6

[0299] Lp4CL1, Lp4CL3, LpCAD1, LpCAD2, LpCCR1, LpOMT1 and LpOMT2 loci mapped to the linkage groups as indicated in Table 3 and in FIG. **40**. These gene locations can now be used as candidate genes for quantitative trait loci for lignin biosynthesis associated traits such as herbage quality, dry matter digestibility, mechanical stress tolerance, disease resistance, insect pest resistance, plant stature and leaf and stem colour.

EXAMPLE 9

Sense Suppression

DNA Sequence Elements and Construct Production

[0300] Three constructs were engineered for development of transgenic perennial ryegrass with modified lignin biosynthesis, using sense suppression technology. The individual components of the sequence elements are listed in Table 4. The promoters and terminators used in construct production originated from perennial ryegrass genomic sequences. The genes were derived from perennial ryegrass cDNA sequences. The origin of the pAUX plasmid vectors has been described previously (Goderis et al., 2002).

TABLE 4

Components used in the generation of constructs for perennial ryegrass transformation.

Construct No.	Vector backbone	Promoters	Genes	Terminators
1	pAUX3132	LpCAD2	LpCAD3	LpCAD2
2	pAUX3169	LpCCR1	LpCCR1	LpCCR1
3	pAUX3169	LpCCR1	LpCCR1(fs)	LpCCR1

[0301] The constructs were produced using Gateway™ technology (Invitrogen). The Gateway™ cloning system consists of one vector backbone and several auxiliary vectors based on pUC18 (Goderis et al., 2002). The multisite recombination cassette was assembled in the auxiliary vectors utilizing the multi-cloning site, flanked by homing endonuclease sites (FIG. **41**). Homing endonucleases are rare cutting restriction enzymes minimising the risk of accidental restriction within the expression cassettes if excision of the expression cassette is required.

[0302] The respective promoter, cDNA and terminator sequences were amplified by PCR using primers incorporating the appropriate AttB recombination sequences and cloned into separate Gateway™ Entry vectors. For example, three Entry clones were required for the generation of the LpCAD expression vector (Construct 1); the LpCAD3 cDNA (FIG. **42**), the LpCAD2 promoter (FIG. **43**) and the LpCAD2 terminator (FIG. **44**). These were then combined with pAUX3132 for the multi-recombination reaction and generation of the expression cassette pAUX3132-LpCAD2::LpCAD3::LpCAD2 (FIG. **45**).

[0303] For Construct 2, Entry clones with the individual components; LpCCR1 promoter, LpCCR1 cDNA, and LpCCR1 terminator were generated using the same PCR cloning strategy. The Entry clones were combined in a recombination cloning reaction with base vector pAUX3169 to produce the final construct pAUX3169-LpCCR1::LpCCR1::LpCCR1 (FIG. **46**).

[0304] For Construct 3, an alternative silencing strategy was employed involving a frame-shift based approach. This method involves the deletion of a single base pair, just downstream of the start site, which is introduced using a forward primer which has the single base deletion (FIG. **47**). This construct works via sense suppression, as the transcript produced will not encode the correct protein and no functional protein will be produced.

[0305] The Entry clones with individual components LpCCR1 promoter, LpCCR1(fs) cDNA, and LpCCR1 terminator were generated and combined in a recombination cloning reaction with base vector pAUX3169 to produce the final construct pAUX3169-LpCCR1::LpCCR1(fs)::LpCCR1 (FIG. **48**).

[0306] The plant selectable marker which facilitates selection of putative transgenic ryegrass on the antibiotic Hygromycin B is contained on a separate plasmid, pAcH1. This plasmid utilizes the rice Actin1D promoter to drive in planta expression of the hygromycin phosphotransferase (hph) gene. The pAcH1 plasmid has been used previously in the transformation of forage grasses (Spangenberg et al., 1995).

Transformation Protocols

[0307] The protocol developed and established is based on the biolistic transformation of embryogenic calli induced from immature inflorescences isolated from an in planta maintained vernalised collection of perennial ryegrass, or seedling meristems derived from in vitro seedling cultures. Illustrations of the different stages in both processes, from the isolation of explants for the induction and proliferation of embryogenic calli for genetic transformation to the recovery of transgenic plants are shown in FIGS. **50** and **51**. Both genetic transformation methods allow for a sustainable, readily-available source of donor plant materials which are highly competent for plant regeneration and genetic transformation and are compatible with biolistic transformation techniques. A general outline of the process involved in transformation is described in FIG. **52**.

Molecular Analysis of Putative Transgenic Plants

[0308] Molecular analysis of putative transgenic perennial ryegrass plants has been conducted using primers for Q-PCR. The following primers were designed:

[0309] 1. Primers specific for the hph gene

[0310] 2. Primers across the CAD2 promoter-CAD3 gene junction

[0311] 3. Primers specific for the pAUX3169 vector (as primers specific for the CCR1 junctions could also amplify endogenous genomic sequences).

[0312] An example of Q-PCR run for detection of hph in extracted genomic DNA is shown in FIG. **53**.

[0313] The results summarising the number of transgenic perennial ryegrass plants for each Construct is shown in Table 5.

TABLE 5

Summary of transformation progress for perennial ryegrass lines harbouring constructs for the modification of lignin biosynthesis.

Construct	Vector	No. Putative Transgenics	No. hph positive	No. GOI positive
1	pAUX3132-LpCAD2::LpCAD3-LpCAD2	180	65	25
2	pAUX3169-LpCCR1::LpCCR1::LpCCR1	90	67	38
3	pAUX3169-LpCCR1::LpCCR1(fs)::LpCCR1	322	185	141
Total		592	317	204

Down-Regulation of CAD and CCR Expression by RNA Interference and Sense Suppression

[0314] In order to modify the expression level of LpCCR1 in perennial ryegrass, an RNA-mediated posttranscriptional gene silencing strategy was employed (RNA interference). The maize Ubiquitin (Ubi) promoter was used to drive expression of a LpCCR1 hairpin (hp) construct containing the variable region of 3' UTR in transgenic perennial ryegrass. Immature inflorescence-derived calli of perennial ryegrass were used as a target for biolistic transformation. hpLpCCR1 transgenic ryegrass plants were confirmed by Southern analysis (FIG. **54**).

[0315] In the same manner, CAD and CCR expression is modified in perennial ryegrass using constructs 1, 2 and 3 (sense suppression).

Analysis of Lignin in Transgenic Plants

[0316] Lignin content and composition is visualised by specialized staining methods, including Mäule histochemical staining which can differentiate between G-lignin and S-lignin monomers (Moore et al., 1991). Mäule staining of flowering stems from different internodes was conducted for wild type and Ubi::hpLpCCR1 transgenic perennial ryegrass. The

results demonstrate that there is significantly less lignin accumulating in stems at both the early reproductive (R1) and mid-reproductive (R2) stages (FIG. **55**). Furthermore, there is an acropetal (base to apex) decrease in the relative amount of total lignin in the different internodes.

[0317] In the same manner, lignin content and composition is analysed in transgenic perennial ryegrass lines harbouring constructs 1, 2 and 3.

[0318] Additional lignin analytical methods includes isolation of cell wall material by successive hot water, ethanol and chloroform/methanol extractions (Fukushima and Hatfield, 2001) followed by determination of total lignin content/dry weight, using acetyl bromide method (Liyama and Wallis, 1990) (FIG. **56**).

[0319] Further lignin monomer analysis to determine the G/S ratio is performed by thioacidoylysis cleavage method (Rolando et al., 1992) and quantification using a gas chromatography (GC-MS) (FIG. **57**).

REFERENCES

[0320] Fukushima, R. S. and R. D. Hatfield (2001). "Extraction and isolation of lignin for utilization as a standard to determine lignin concentration using acetyl bromide spectrophotometric method." *J. Agri. Food Chem.* 49: 3133-3139.

[0321] Goderis, I., M. De Bolle, I. Francois, P. Wouters, W. Broekaert and B. Cammue (2002). "A set of modular plant transformation vectors allowing flexible insertion of up to six expression units." *Plant Mol Biol* 50: 17-27.

[0322] Heath et al (1988) cDNA cloning and differential expression of three caffeic acid O-methyltransferase homologues from perennial ryegrass (*Lolium perenne*). Journal of Plant Physiology 153:649-657

[0323] Lichtenstein, C, And J. Draper (1985) Genetic engineering of plants. In: D. M. Glover (ed.), DNA Cloning, Vol. 2, pp. 67-119, IRL Press, Washington.

[0324] Liyama, K. and A. F. A. Wallis (1990). "Determination of lignin in herbaceous plants by an improved acetyl bromide procedure." *J Sci Food Agric* 51: 145-161.

[0325] Moore, K. J., L. E. Moser, K. P. Vogel, S. S. Waller, Johnson B. E. and P. J. F. (1991). "Describing and quantifying growth stages of perennial forage grasses." *Agron. J.* 83: 1073-1077.

[0326] Rolando, C., B. Monties and C. Lapierre (1992). Thioacidolysis. *Methods in Lignin Chemistry* S. Y. Lin and C. W. Dence, Springer-Verlag: pp. 334-349.

[0327] Spangenberg, G., Z. Y. Wang, X. L. Wu, J. Nagel, V. A. Iglesias and I. Potrykus (1995). "Transgenic tall fescue and red fescue plants from microprojectile bombardment of embryogenic suspension cells." *J Plant Physiol* 145: 693-701.

[0328] Finally, it is to be understood that various alterations, modifications and/or additions may be made without departing from the spirit of the present invention as outlined herein.

[0329] It will also be understood that the term "comprises" (or its grammatical variants) as used in this specification is equivalent to the term "includes" and should not be taken as excluding the presence of other elements or features.

[0330] Documents cited in this specification are for reference purposes only and their inclusion is not an acknowledgement that they form part of the common general knowledge in the relevant art.

SEQUENCE LISTING

The patent application contains a lengthy "Sequence Listing" section. A copy of the "Sequence Listing" is available in electronic form from the USPTO web site (http://seqdata.uspto.gov/?pageRequest=docDetail&DocID=US20100287660A1). An electronic copy of the "Sequence Listing" will also be available from the USPTO upon request and payment of the fee set forth in 37 CFR 1.19(b)(3).

1-26. (canceled)

27. A substantially purified or isolated nucleic acid comprising a fragment or variant of a gene encoding a lignin biosynthetic enzyme, said nucleic acid being capable of modifying lignin biosynthesis in a plant via sense suppression;

wherein said fragment or variant comprises a frame shift mutation relative to the gene upon which the fragment or variant is based, resulting in a loss or substantial reduction in enzymatic activity in the encoded polypeptide; and

wherein said frame shift mutation is a mutation that inserts or deletes a number of nucleotides that is not evenly divisible by three within a short distance of the ATG start codon of the gene upon which the fragment or variant is based.

28. The nucleic acid according to claim **27**, wherein said lignin biosynthetic enzyme is selected from the group consisting of 4CL, CCR, CAD, C3H, C4H, CCoAOMT, COMT, F5H and PAL.

29. The nucleic acid according to claim **27**, from a forage, turf or bioenergy grass species.

30. The nucleic acid according to claim **29**, from a *Lolium, Festuca, Cynodon, Bracharia, Paspalum, Panicum, Miscanthus, Pennisetum*, or *Phalaris species*.

31. The nucleic acid according to claim **27**, wherein said gene encodes 4CL and comprises a nucleotide sequence selected from the group consisting of the Seq, ID Nos: 1, 3, 5, 21, 25, 29, 33, 37, and 235-243; and wherein said nucleic acid is capable of modifying lignin biosynthesis in a plant via sense suppression of a gene encoding 4CL in said plant.

32. The nucleic acid according to claim **27** wherein said gene encodes CCR and comprises a nucleotide sequence selected from the group consisting of sequences shown in Sequence ID No: 7, 117, 121 and 244 to 251 and wherein said nucleic acid is capable of modifying lignin biosynthesis in a plant via sense suppression of a gene encoding CCR in said plant.

33. The nucleic acid according to claim **27**, wherein said gene encodes CAD and comprises a nucleotide sequence selected from the group consisting of the sequences shown in Sequence ID Nos: 9, 11, 14, 16 53, 57, 252 to 269 and 361; and wherein said nucleic acid is capable of modifying lignin biosynthesis in a plant via sense suppression of a gene encoding CAD in said plant.

34. The nucleic acid according to claim **27**, wherein said gene encodes C4H and comprises a nucleotide sequence selected from the group consisting of Sequence ID Nos: 41, 45, 49, and 270-273; and wherein said nucleic acid is capable of modifying lignin biosynthesis in a plant via sense suppression of a gene encoding C4H in said plant.

35. The nucleic acid according to claim **27**, wherein said gene encodes CCoAOMT and comprises a nucleotide sequence selected from the group consisting of Sequence ID Nos: 89, 93, 274 to 294 and 362; and wherein said nucleic acid is capable of modifying lignin biosynthesis in a plant via sense suppression of a gene encoding CCoAOMT in said plant.

36. The nucleic acid according to claim **27**, wherein said gene encodes COMT and comprises a nucleotide sequence selected from the group consisting of Sequence ID Nos: 133, 137, 141, 145, 295 to 342, 360, 362 and 362; and wherein said nucleic acid is capable of modifying lignin biosynthesis in a plant via sense suppression of a gene encoding COMT in said plant.

37. The nucleic acid according to claim **27**, wherein said gene encodes F5H and comprises a nucleotide sequence selected from the group consisting of Sequence ID Nos: 173, 177, and 343- to 346; and wherein said nucleic acid is capable of modifying lignin biosynthesis in a plant via sense suppression of a gene encoding F5H in said plant.

38. The nucleic acid according to claim **27**, wherein said gene encodes PAL and comprises a nucleotide sequence selected from the group consisting of Sequence ID Nos: 181, 185, 189, 193, 197, 201, 205, 209, 213, 217, 220, 224 and 347 to 358; and wherein said nucleic acid is capable of modifying lignin biosynthesis in a plant via sense suppression of a gene encoding PAL in said plant.

39. The nucleic acid according to claim **27**, wherein said gene encodes C3H and comprises a nucleotide sequence shown in SEQ ID No: 359; and wherein said nucleic acid is capable of modifying lignin biosynthesis in a plant via sense suppression of a gene encoding C3H in said plant.

40. The nucleic acid according to claim **27**, comprising a nucleotide sequence selected from the group consisting of SEQ ID Nos: 23, 27, 31, 35, 39, 43, 47, 51, 55, 59, 63, 67, 71, 75, 79, 83, 87, 91, 95, 99, 103, 107, 111, 115, 119, 123, 127, 131, 135, 139, 143, 147, 151, 155, 159, 163, 167, 171, 175, 179, 183, 187, 191, 195, 199, 203, 207, 211, 215, 218, 222, 225, 229 and 233.

41. The nucleic acid according to claim **27** encoding a polypeptide comprising a sequence selected from the group consisting of SEQ ID Nos: 24, 28, 32, 36, 40, 44, 48, 56, 60, 64, 68. 72, 76, 80, 84, 88, 92, 96, 100, 104, 108, 112, 116, 120, 124, 128, 132, 136, 140, 144, 148, 152, 156, 160, 164, 168, 172, 176, 180, 184, 188, 192, 196, 200, 204, 208, 212, 216, 219, 223, 226, 230 and 234.

42. A genetic construct or vector comprising the nucleic acid according to claim **27**.

43. A transformed plant, plant cell, plant seed or other plant part or transformed plant biomass comprising the nucleic acid according to claim **27**.

44. A method of modifying lignin biosynthesis in a plant, said method including introducing into said plant in a sense orientation an effective amount of a nucleic acid according to claim **27**, such that expression of the corresponding gene is suppressed.

45. The method according to claim **44**, wherein said plant is selected from the group consisting of *Lolium, Festuca, Cynodon, Bracharia, Paspalum, Panicum, Miscanthus, Pennisetum, Phalaris*, and other forage and turf grasses, corn, grains, oat, sugarcane, wheat, barley, *Arabidopsis*, tobacco, legumes, Alfalfa, oak, *Eucalyptus*, maple, *Populus*, canola, soybean, chickpea and *Pinus*.

46. The method according to claim **44**, wherein said nucleic acid is part of a genetic construct or vector.

* * * * *