Nucleotide sequence of the 25S–18S ribosomal gene spacer from *Arabidopsis thaliana*

P.Gruendler, I.Unfried, R.Pointner and D.Schweizer

Institut für Botanik der Universität Wien, Abteilung Cytologie und Genetik, Rennweg 14, A-1030 Vienna, Austria

Submitted June 29, 1989

EMBL accession no. X15550

From an *Arabidopsis thaliana* genomic library constructed in the lambda replacement vector EMBL3 we have isolated a clone containing a complete rDNA repeat unit and subcloned its five EcoRI fragments (1) in the phagemid vector pBlueScript KS⁺ (Promega, Inc.). Three fragments covering the 25S–18S intergenic region (IGR) with flanking regions were sequenced by the dideoxy method after introduction of nested deletions with ExonucleaseIII/ Nuclease S1.

The IGR from *Arabidopsis* contains several regions unusually rich in SalI recognition sites. Similar 'SalI-Boxes' are present in the IGR of *Mus musculus*, where they serve as a transcriptional termination site (2).

```

1  gaattcacca agtgttgat tgctcacca ccaataggga acgtgagctg ggtttagaac gtgtgagac aggttagttt taccctactg atgccecggt
101  cgcgatagta attcaaccta gtaecagagg aacgggtgat tgcacaaatt ggtcatcgcg cttgggttaa aagccagctg ccgaagaeta ccgtgcgctg
201  gattatgact gaagccctct aagtacagat ccggggtaga agcgaagcat ggcgcgcgcg ccagattgce gaccctcagt aggaagcttat ggtccaaagg
301  cactgtcgtt tggctaagtc cgttcggcgg aacgggtgct cggaccgcct tgaattataa ttaccaccca gggcggggta gaatcctttg cagacgactt
401  aaatacgcga cgggggtatg taagtggcag agtggccttg ctgccacgat ccactgagat tcaagccctt gtoctaaaga ttccgccccg ccttaaatca
501  ctccaaaaaa aaacatcccc aattctacac aagtgtttct acactaacaa agcaacagct ccttaacgaa ttcccaactt tacacgagct cgtctctcga
601  ggttaaatgt tattacttgg taagattccg gacctcgcca agtgttttga aaaccccgaa ccgtcgcaaa ttgggtatgt gagaataata agtaagaga
701  cagactttgt caaaaccccc acccagaaag tccatagtga gaagagttag tcaagagata gactttgtcc aaagaaacgc gaagagaaag cgtggggaga
801  ccgtcacgaa ggtgcagat gagaaagata agtcaagaga cagactttgt cgaagaagaa tgcgtggggg tacactcagc aaggtgcata
901  gtgagaagac taactcaaga cagacagctg ttccaaaga aaacaaagag aatgcttggg gagatagagc tctgagatag ttctcaagct aagaagtgt
1001  taanaagctaa gaactagcat caaatgatgg atgaacaca aggtagtgtt tgaanaagta aacacttggt gatatgaaca caaacgttca atatgcacaa
1101  cccatgccaa gataagagaa atgaaacct ggtgattgtt gcggaatcgc tccagcagtc ctgcacaggg acttgaaatc ctccagggga aaatatcggt
1201  tccgaggaat cgtcgatccg gacttggaa ttccgcagaa agttttaccg ctccagagat ttgtccacca ggaagtggaa tctgcagaa aaatctatcg
1301  ggtccagcaa atctgtcagc aggcagagaa atctgtcagc ggggtccagg atttgcgac caggggttga aatcgtcgac caggtccgag acttcatcga
1401  cgggggtcca ggaattcgcg accagcagtc ggaattcgcg agaaaaatct atcgggtccg aggaatctgc caccagagac aggaatctgc caccagagac
1501  gaggatttgt cgaccaggtc cgagacttca tgcaccgggt ccgaggatgc gtccagcagg tccgagactt catcgaccgg gtccaaagtat ttgattttat
1601  tttggaaacc gtgtctcctc agacatttca atgtatgttg gtgccaaagc ggaagagggc tataaagcta tataggggtg ggtgtgtgta tggggatatt
1701  tgcctcgagg cccgcgcgcg aataagaata agaatagata atagaataa gaataagaaa agaaaaaa aaataacaa aaactccata
1801  gagttgtttt ctcaatccgg tccgaggaat cgtcgatctg gacttggaa ttccgcagaa agttttaccg gtccgaaaa ttgtccacca ggaagtggaa
1901  tctgtcagaa aaatctatcg ggtccagaa atctgtcagc aggcagagaa atctgtcagc ggggtccagg atttgcgac caggggttga aatagtcgac
2001  caggtccgag acttcatcga cgggtccga ggaattcgcg accagcagtc ggaattcgcg agaaaaatct atcgggtccg aggaatctgc caccagagac
2101  aggaatctgc cagccgggtc gaggatttgt cgaccaggtc cgagacttca tgcaccgggt ccgaggatgc gtccagcagg tccgagactt catcgaccgg
2201  gtccaaagtat ttgattttat tttggaaacc gtgtctcctc agacatttca atgtatgttg gtgccaaagc ggaagagggc tataaagcta tataggggtg
2301  ggtgtgtgta tggggatatt tgcctcgagg cccgcgcgca atagaataa gaataagaa taagaataa agaaaaaa aaataacaa aaactccata
2401  acaaaaaa aaatccata gagttgtttt ctcaatccgg tccgaggaat cgtcgatctg gacttggaa ttccgcagaa agttttaccg gtccgaaaa
2501  ttgtccacca ggaagtggaa tctgcagaa aaatctatcg ggtccagaa atctgtcagc aggcagagaa atctgtcagc ggggtccagg atttgcgac
2601  caggggttga aatctcagc caggtccga acttcatcga cgggtccga ggaattcgcg accagcagtc tgaattgtgc atgaatctct atcgggttgc
2701  aggaatctgc cagccgggtc tgaatctgc cagccaggtc gagacttcat cgaccgtgtc cagaggatgc tccaggggtt gtccagcagg acgaggaatc
2801  gtccagcagg tccgaggaat tctgcagcag ggggtgaa ttctgcagc cagccaggaat tctgcagca ggaagcggga accttcgacc agcaggaatc
2901  atggcagatc aaatctatc ggggtccagg aatgtcgac caggggttga aatctcagc caggtccgag acttcatcga cgggtccga ggaattcgct

```

```

3001 ACCAGGACGG CCGGATGTCC GAGAAAAAAA AATGTTGCCG AATAACTTTC GAAATCATT GGATATGATG CAATGTTTTG TGATCGAATC TCTAAAAATA
3101 CATCAATAAA GAGTTTAGGA TGTCAAATTT GCATCAAAATG TGGCCACGGG GCCCCAACTA GACCATGAAA ATCCGATGTT GTATCAGGTC AAATGACCTA
3201 GCTAGAGGTG TCAAAAAATT ATGAAAAATTT ACCGAAAAAT AGGATTAGT ATCTTTATGA TGCATGCCAA AAGAATTTTT CAATTTCCAA GTATTTCTTT
3301 TTTCTTGCCA CCGGTGTCTC CTCAGACATT TCAATGTCTG TTGGTGCCAA GAGGGAAAAG GGCATTATAG CATATATAGG GGGTGGGTGT TGAGGGAGTC
3401 TGGGCAGTCC GTGGGGAAAC CCGTTTTTTC GTTCGACATT GGGTAGCGAT CGAGGGATGG TATCGGATAT CGGCACGAGG AATGACCGAC CGTCCGGCCG
3501 CCGGGATTTT CGCGGAAAAA CTTTTCCGGG CACTTTTCCG GCGATCGGTT TTGTTGCCTT TTTCCGAGTT TTCTCAGCAG TTCTCGGACA AAAATGCTG
3601 AATCGTCCAG GAGATGGGG TTGCTTGGCT GGGCTGCCAT TAGTCTTTCG AGGCGTAGG GTGGCGGGCG TATAAAGTG TCGGAGTTTT TTCAGCAGTT
3701 CTCGGACAAA AATTGCTGAG TGGCCGAGAA GAATGGGGCT GTCATGCGTG GCGTAGCATG GATTCTTCGA GGCCTAGGGG TGGCGGTATA TAACTTGTTC
3801 GCATGATATT ACCGAGATG CCCCCGGGG ATCTTTTCCG CTCGTGCCCG AAGAGAAATG GCGGTGCATG GCATGGGCTG ACATGGATTC TCCTAGGCCG
3901 TTTGGGTGGC GGTATAGTGG TCTTGGCGAC GAATACCGA GATGTCCCGA TGGGCATCGA TTCCACCGCG CTAGGTTGGA TGGCGGTGCT TCGTCGGAAA
4001 GCATGGATCC GCGTAGGCTG TCCCGAGTGT GAGCGAGGTG TGAGTGTCCG CCATGGGCAT CGACACCTTG CGGCTAGGAA TGGGAACGAG ACGGGTGGCA
4101 AAGATTTCGA GTAGCACTTC ATACTACCGT GGGTTTTTAA AACCTTCGA GTTTTGTGTA TGTATTCCG AGAATTAGCA AACCGTAACG AAGATGTTCT
4201 TGGCAACCAT CTTTGTATGG GAGTCCGGCT GTTCGATAGC CGGCCAAGGG TGATGAACGA AATGTGAACC CTGTCTCGCG CTAGGTTGGA TGGCGGTGCT
4301 TCGTTGGAAA GCATGGATCC GCGTAGGCTG TCCCGAGTGT GAGCGAGGTG TGAGTGTCCG CCATGGGCAT CGACACCTTG CGGCTAGGAA TGGGAACGAG
4401 ACGGGTAGCA AAGATTTCGA GTAGCACTTC ATACTACCGT GGGTTTTTAA AACCTTCGA GTTTTGTGTA TGTATTCCG AGAATTAGCA AACCGTAACG
4501 AAGATGTTCT TGGCAACCAT CTTTGTATGG GAGTCCGGCT GTTCGAAAGC CGGCCAAGGG TGATGAACGA AATGTGAACC CTGTCTCGCG CTAGGTTGGA
4601 TGGCGGTGCT TCGTTGGAAA GCATGGATCC GCGTAGGCTG TCCCGAAGTA TCTCGCGCTT GTACGGCTTT GCGTGGGATT CGTCCGCTTT CTTTCTCTTT
4701 AGCCGAGTAC TTCGGTAGAT TAGTTGGAAc GATTGATGAT TTTGAGTTAA TTGAACGTTT GCGGTATGAG TGGTATCGG ATAGCTAGTG TTCGTAGGCT
4801 CCGTGTCCG GCATCGAACT ACCTACCAAC TATCTTCTC AGTTAATTCA CGGGCGATG TACGCTCGAT GATGAGTTCC GGGGCGTGT TTTCTACCT
4901 AATTTGAAGG AATTGTTGAG TTTGGTTTAC ACCTTTGCCG CGGGCTTCTC CTTGCTGGGG AAGTCTGGGG CTCAAACATC GCGGCTTTCT CACCTCTCGT
5001 CATCGCATTT GTTGGCTTGC TCGCATGGT GAATGAGTTG CGGGTGAAA TCTCGGATGC GGAAGAGTTG TCGACGGTGA CTCGAAGTGA TTCAGTCCCG
5101 CCAAGCTCA TCGGTCTTC GGGCAAAAa TGACGGTCAA GACCTCTGCC TTTCTCTCTT TCCATTGCGT TTGAGAGGAT GTGGCGGGGA ATTGCCGTGA
5201 TCGATGAATC Ctaacgtgtt gatctgtcca gtagtcatat gcttgctcca aagattaagc catgcagtgt taagtatgaa cgaattc

```

Fig.1 Nucleotide sequence of an *Arabidopsis thaliana* 25S - 18S intergenic region. The sense strand relative to the coding regions is shown. The 3' terminus of the 25S gene and the 5' terminus of the 18S gene as determined by homology to lemon 26S rRNA and rice 18S rRNA are printed in lowercase letters. The three Sall-Boxes are underlined.

Acknowledgements: This work was supported by grants P3079 from the "Oesterreichische Nationalbank" and P6577B from the Austrian "Fonds zur Foerderung der wissenschaftlichen Forschung".

References

1. Pruitt, R.E. and Meyerowitz, E.J.M. (1986) J.Mol.Biol. **187**, 169-183
2. Grummt, I., Kuhn, A., Bartsch, I. and Rosenbauer, H. (1986) Cell **47**, 901-911.