

The sequence of the maize *rps19* locus and of the inverted repeat/unique region junctions

William E. McLaughlin and Ignacio M. Larrinua\*

Lilly Research Laboratories, Lilly Corporate Center, Indianapolis, IN 46285, USA  
Submitted March 9, 1987

Accession no. Y00141

We have cloned and sequenced the gene for the small ribosomal subunit protein S19(1,2) from the plastid genome of the monocot *Zea mays*. The *rps19* locus codes for a protein of 93 amino acids with a molecular weight of 10,739. This locus is highly homologous to both the spinach and *E. coli* S19 proteins (1,3). Unlike its dicot homologues (1,2), this locus is entirely within the inverted repeat. Indeed the inverted repeats end 34 bases after the stop codon. Downstream from the gene copy in IRII (Bam 6) (4) is the locus for *rpl22* (McLaughlin, W.E. and Larrinua, I.M. manuscript in preparation). These genes appear to be part of the same operon, a situation similar to that of *E. coli* (3). Downstream from the gene copy in IRI (Bam 8), is the locus for *psbA*. This transcript is on the opposite strand (McLaughlin, W.E. and Larrinua, I.M., unpublished data) and these genes are separated by a strong terminator. The coordinates for the gene in IRII are 110.56 to 110.38 on the negative strand while the copy on IRI runs from 26.19 to 26.37 on the positive strand (4). There is a strong bias in codon usage the %GC content of the three positions being 46.15, 41.75, and 34.06 respectively. This is consistent with the high AT content of this genome.

```

          RPL2      —SD—      RPS 19
1  TTA AAA TGA AAAAAAGGAGTAATCAGCT GTG ACA CGA AAA AAA ACG AAT CCT TTT GTA
   L  K  ?3?                               M  T  R  K  K  T  N  P  F  V
59 GCT C6T CAT TTA TTG GCA AAA ATC GAA AAG GTC AAT ATG AAG GAG GAG AAA GAA ATA
   A  R  H  L  L  A  K  I  E  K  V  N  M  K  E  E  K  E  I
116 ATA GTA ACG TGG TCC CGG GCA TCT AGC ATT CTA CCC GCA ATG GTT GGC CAT ACA ATC
   I  V  T  W  S  R  A  S  S  I  L  P  A  M  V  G  H  T  I
173 GCG ATT CAT AAT GGA AAA GAA CAT ATA CCT ATT TAC ATA ACA AAT CCT ATG GTA GGT
   A  I  H  N  G  K  E  H  I  P  I  Y  I  T  N  P  M  V  G
230 CGC AAA TTG GGG GAA TTC GTG CCT ACT CGG CAT TTC ACG AGT TAT GAA AGT ACA AGA
   R  K  L  G  E  F  V  P  T  R  H  F  T  S  Y  E  S  T  R
                                     INVERTED REPEAT II
287 AAG GAT ACT AAA TCT CGT CGTTAATTGAATTCAGAATGAAAAGATTCAGAATAAACAAA/ATTT
   K  D  T  K  S  R  R  ?1?
          SINGLE COPY REGION BAM 6                                     RPL22
353 ATAGGATTCAAATAATCAATAAAGTAAAGGTAGGCGGGTAATAACCTTATTT ATG ACA AGT
                                     MET THR SER
                                     /6AAA
SINGLE COPY REGION BAM 8
TACCAATATCCTGTTG6AACAAGATATT666TATTTCGGCTTTCCTTCTCAAAA
putative terminator

```

1. Zurawski, G et al. 1982 *Nucl. Acid Res.* 12:6547-6558
2. Sugita, M and Sugiura, M 1983 *Nucl. Acid Res.* 11:1913-1918
3. Zurawski, G and Zurawski, SM 1985 *Nucl. Acid Res.* 13:4521-45264
4. Larrinua, IM et al. 1983 *Pl. Molec. Biol.* 2:129-140

\*To whom correspondence should be addressed