

Plant Gene Register

Genomic Nucleotide Sequence of a Wild-Type Shrunken-2 Allele of *Zea mays*¹

Janine R. Shaw and L. Curtis Hannah*

2211 Fifield Hall, University of Florida, Gainesville, Florida 32611

The maize (*Zea mays*) endosperm enzyme, ADP-glucose pyrophosphorylase, which produces ADP-glucose from ATP and glucose-1-phosphate, is an important enzyme in the synthesis of starch. ADP-glucose arising from the action of this enzyme is the major, if not sole, donor of glucose for starch biosynthesis. The enzyme is composed of two dissimilar subunits encoded by the two unlinked genes, *Shrunken-2* (*Sh2*) and *Brittle-2* (*Bt2*) (1, 2). The enzyme is allosterically activated by 3-phosphoglyceric acid and inhibited by phosphate (3). Although it remains an open question whether these allosteric properties are physiologically relevant or whether they simply reflect the evolutionary history of the two structural genes (4), there does exist, nevertheless, much interest in determining whether genetic modification of this enzyme could lead to increased rates of starch biosynthesis in the maize endosperm. Because endosperm starch content comprises approximately 70% of the dry weight of the seed, alterations in starch biosynthesis would clearly affect corn yield.

As a first step toward such experiments, we have isolated and sequenced genomic clones of *Sh2*. The structure of the gene (Table I, Fig. 1) is based on sequence analysis of three overlapping clones isolated from a Black Mexican Sweet genomic library. Exonic sequences were defined by comparison with the cDNA sequenced previously (2) and further sequencing of cDNAs subsequently isolated. These sequences

¹ This work was supported by the U.S. Department of Agriculture competitive grants 88–37234–3917 and 90–37262–5638. This is University of Florida (Agricultural Experiment Station) Journal series No. R-01954.

correct for the fact that the original cDNA contained some non-*Sh2* sequences at its extremities. The exonic genomic and cDNA sequences are nearly 100% identical. The few differences may be due to DNA polymorphisms because these clones were isolated from different corn lines.

Placement of the start of the first exon is based on primer extension experiments. In the absence of dideoxynucleotides, four major bands, differing by one nucleotide, were observed. In the presence of dideoxynucleotides, the three large bands were seen in all four sequencing tracks. In each case, the largest band occurred at the same distance from the primer. It is unknown whether this reflects heterogeneity within the normal population of *Sh2* transcripts or some laboratory artifact. Nevertheless, the start of transcription can be placed within three base pairs.

ACKNOWLEDGMENTS

We would like to thank Jung-Myung Bae for the 5' mRNA sequence. We thank the ICBR DNA Synthesis Core, University of Florida, for the oligonucleotides used for sequencing.

LITERATURE CITED

1. Bae JM, Giroux M, Hannah L (1990) Cloning and characterization of the *Brittle-2* gene of maize. *Maydica* 35: 317–322
2. Bhave MR, Lawrence S, Barton C, Hannah LC (1990) Identification and molecular characterization of shrunken-2 cDNA clones of maize. *Plant Cell* 2: 581–586
3. Dickinson DB, Preiss J (1969) Presence of ADP-glucose pyrophosphorylase in *Shrunken-2* and *Brittle-2* mutants of maize endosperm. *Plant Physiol* 44: 1058–1062
4. Hannah LC, Nelson OE (1975) Characterization of adenosine diphosphate glucose pyrophosphorylase from developing maize seeds. *Plant Physiol* 55: 297–302

Table I. Characteristics of a Shrunken-2 Wild-Type Gene from *Zea mays* L.**Organism:***Zea mays* L. var Black Mexican Sweet**Location on Chromosome:**

3L

Gene Product; Pathway:

A subunit of ADP-glucose pyrophosphorylase; starch biosynthesis

Gene Designation:*Sh2***Techniques:**

λ EMBL3 *Sau*3A partial genomic library screened with *Sh2* cDNA resulting in three overlapping clones; restriction fragment subcloning into pUC19 and pSPORT; double-stranded plasmid dideoxynucleotide sequencing of the overlapping clones (both strands) using various subclones and synthetic oligonucleotide primers; primer extension with and without dideoxynucleotides to determine transcription start.

Methods of Identification:Exon sequences nearly 100% identical to *Sh2* cDNA**Expression Characteristics:***Sh2* is expressed only in the endosperm of the maize kernel.**Gene Structure:**The coding region of 1913 bp^a is interrupted by 15 introns ranging in size from 68 bp to 1821 bp

Promoter Region:	putative TATA box	TATATAAA at -33 bp to -26 bp
	putative RY repeats	CATGCATG at -926 bp to -919 bp
		CATGCATA at -510 bp to -503 bp
	similarity to animal enhancer	GTGGAAAC at -911 bp to -904 bp
Intron Region:	similarity to animal enhancer	GTGGATAG in 5' part of largest intron

Codon Usage:

43.2% G + C in coding region

No obvious bias in codon usage

Protein Characteristics:

The 1551 bp open reading frame gives rise to a 516 amino acid peptide; *M*, 57,178.98; ADP-glucose pyrophosphorylase is allosterically regulated by phosphoglyceric acid (activation) and phosphate (inhibition).

Antibodies:

Antibodies made against fusion protein expressed in *Escherichia coli* recognizes an endosperm-specific protein. Size and amount of this protein is altered in *sh2* mutants (M. Giroux, L.C. Hannah, unpublished).

Subcellular Localization:

Reported to be in the amyloplasts.

GenBank Accession No.:

M81603

^a Base pairs.

```

-1020 TGATGCTTTTCTGGCCAGGAGAGCTATGAGACGTATGCTCCTCAAAGCCACTTTGCAT
-960 TGTGTAAACCAATATCGATCTTTGTTACTTCACTCATGTCATGAACATTGTTGGAAACATC
-900 TAGCTTACAAGCATTAGTGACAGCTCAGAAAAAGTTATCTCTGAAAGGTTTCATGTGTA
-840 CCGTGGGAAATGAGAAATGTTGCCAACTCAAAACCCCTCAATATGTTGTTGACGGCAAA
-780 CTCTTCTGGAAGAAAGGTGCTAAAACTATGAACGGTTACAGAAAGGTATAAACCAACGG
-720 CTGTGCAATTTGGAAGTATCATCTATAGATGCTGTGTGAGGGGAAAGCCGTCACGCCAAC
-660 TTATTTACTCAGAAACAGCTTCAACACACAGTGTGTCTGTTATATGATGGCATCTCCACCC
-600 AGGCCCCACCATCACTATTCACTATCTCTGTGCCTGTTATTTCTGTCCCTTTCT
-540 GATCATAAAAAATCATTAAAGGTTTGCAAACATGCATAGGCATATCAATATGCTCATTTA
-480 TTAATTTGCTAGCAGATCTCTCTCACTCTTACTTTATTTATTTGTTGAAAAATATG
-420 CCTGCACCTAGGAGCTCGTATACAGTACCAATGCATCTTCAATAATGTAATTCAGA
-360 AAGGAAGTAGGAACTATGAGAGTATTTTCAAATAATAGCGGCTTCTATTATGTTG
-300 ATGACAAAGGCCAAGGCAAAATCGGAACACTAATGATGGTTGGTTGCATGAGCTGTGCG
-240 ATTACTTGAAGAAATGTAACCTTTGTTTCTGTGCGTGGGCATAAAAACAAACAGCTTCT
-180 AGCCTCTTTTACGGTACTTGCATTTGCAAGAAATGGAACCTTTTCTTTTCTGTATGT
-120 GCACATAATGCAAAAGCTCCAGGCTTTTTCATGGTTGATGCTTTTACACAGTTCAT
-60 CTCCACAGTATGCCCTGCTCATACTCTATAAACAACATCAACAGCATCGCAATTAGCC
1  ACAAGTCACTTCGGAGGCAAGTGTGATTTCGACCTGTGCAGCCACCTTTTGGTTCG
61  TTgtaagtatacttccctcaactcttctctgttagtttaattgtaactctgggaagta
121  ttatgggaaagagatgagatgctatctctatctctgcaaatgcaatctgagcttat
181  atgggctgtctcataaatttgaattgctccatctctgcccacaataattgcaaggtat
241  atgctagtctccataaaagttctgtttttctctctctctctctctctctctctctctct
301  ttttgcctagggaaaggaactctggttttctctctctctctctctctctctctctctct
361  atgtccagtttttgaagtaataaactctcaagtttggctataagatgcaatataaagtt
421  caaacatatactcaagttcaactctgtaaatgctcccttttctctctctctctctctctct
481  ctcaattatttgagttgagtgatctctctctctctctctctctctctctctctctctctct
541  CATTGGACACGAACCTCAGTCTCACCAGATAAGATCTTGTGAGGCTGATGGGATGGACA
  A L D T N S G P H Q I R S C E G D G I D
601  GGTGGAATAAATTAAGTATGGGGCCAGAAAGCAGGAGAAAGCTTTGAGAAATAGTGTCT
  R L E K L S I G G R K Q E K A L R L N R C
661  TTGTTGGTAGAGTTGCTGCAACTACAAATGATTTACTTACCTGACGTCTGCTCCGTA
  F G G R V A A T T Q C I L T S D A C P E
721  CTCTTgtaagtatccactcaattattactctctacatggtggttactctctctctctctct
  T L
781  tttcaagggaaattactgtatttttgtgtttgtgggagttctctactctctctctctctct
841  tggttatgtaagattttgtcaaatagggctactcaataattggttgaatctctgggaac
901  tgtgttttcaactggttccagyaagaaaggaattattggttactgcatgaataactattg
961  aaatagacctagagttgctctgtagatatacaaaatcattgctacagatctctataata
1021  gttctcttgcacctgcaatatacaataaactgcaactctctctctctctctctctctctctct
1081  atgcaactctgagaaactcaccaggtgtaactcttctctgtaattggtatttaattgagcat
1141  atgcaactctgtagaactctctctctctctctctctctctctctctctctctctctctctct
1201  gCACTTCAAACACAGTCTCTCTGAAAAAATATGCTGATGCAACCCGATATCTGCTAT
  H S Q T Q G S S R K N Y A D A N R V S A
1261  CATTTTGGCCGAGGCACTGATCTCAGCTCTTCTCCTGACAAAGCACAGAGCTACAGCC
  I L G G G T G A A C T Q L F P L T S T R A T P
1321  TGCTgtaaggtataaactgaaactgcaactggtgattactctctctctctctctctctctct
  A
1381  ctgactcttctgaattattctctctctctctctctctctctctctctctctctctctctctct
1441  agatacaaaattgacataaactgaaaggtgtagtaagacagtgagttgaaagattataa
1501  ttttgggagactctccagcaaatcttctctctctctctctctctctctctctctctctctct
1561  gtcgcgctctctcaactctctctctctctctctctctctctctctctctctctctctctctct
  V P V G
1621  GGATGTTACAGCCTTATGATATCCCTATGAGTAACTGCTTCAACAGTGGTATAAATAAG
  G C Y R L I D I P M S N C F N S G I N K
1681  ATATTTGTGATGAGTCACTGTTCAATCTTCACTTCCGCTTAAACCCCATATCTCATGTA
  I F V M S Q F N S T S L N R H I H R T I
1741  CTTGAGGGGGATCAACTTTGCTGATGGATCTGACAGTgatttaccctcatctgtgtg
  L E G G I N F A D G S V Q
1801  atgtgtaactctgtaattaggagtagatttgtgtggagagaataataaacagatgcccag
1861  attcttctctctctctctctctctctctctctctctctctctctctctctctctctctctct
1921  cattedgtcattactttgctcaatgctcaatgctcaatgctcaatgctcaatgctcaatgct
1981  agtgttcaataaaactcaattgctcaactcctgagGTATTAGCGGCTACACAAATGCCCT
  V L A A T Q M P
2041  AAGAGCCAGCTGGATGTTCCAGGATACAGCAGACTCTATCAGAAAAATTTATCTGGGTAC
  E E P A G W F Q G T A D S I R K F I W V
2101  TCGAGgttagtgatattttctctctctctctctctctctctctctctctctctctctctctct
  L E
2161  tctcttgaattttgagttctctctctctctctctctctctctctctctctctctctctctct
  D Y Y S H K S I
2221  GACAACATGTAATCTTGTAGTGGCGATCAGCTTTATCGGATGAATTACATGAACTTTG
  D N I V I L S G D Q L Y R M N Y M E L V
2281  CAGgttagtgattctctctctctctctctctctctctctctctctctctctctctctctct
  Q
2341  actacttttggcatgattatttctctctctctctctctctctctctctctctctctctctct
  K H V E D D A D I T I
2401  TCATGTGCTCTCTGTTGATGAGAGTaatcagttgtttatatacctcaataatgaaatgt
  S C A P V D E S
2461  catcttggatccaacagagatgcataggtctcaatctgctctctctctctctctctctctct
2521  ggaagCCGAGCTTCAAAAATGGGCTAGTGAAGATTGATCATACTGGACGCTGACTTCAA
  R A S K N G L V K I D H T G R V L Q
2581  TTCTTTGAAAAACCAAGGGTCTGATTGTAATCTATGgttagaattctctctctctctctct
  F F E K P K G A D L N S M
2641  ccaactctttgttttctctctctctctctctctctctctctctctctctctctctctctct
2701  ataactgtactgtacttattcagaatgatttctctctctctctctctctctctctctctct
2761  aagctaagtagcattgattgagAGTGTGAGCAACCACTCTGAGCTATGCTATAGAT
  R V E T N F L S Y A I D
2821  GATGCACAGAAATCACCATCTGTCATCAATGGGCAATTTATGCTTCAAGAAAGATGCA
  D A Q K Y P Y L A S M G I Y V F K K D A
2881  CTTTATAGACTTCTCAAgtaactctctctctctctctctctctctctctctctctctctct
  L L D L L K
2941  accottcaacagtgcaactctctctctctctctctctctctctctctctctctctctctct
  S K Y T Q L H D F G S
3001  AAATCTCCCAAGAGCTGACTAGATAGTGTGCAAGTgagttgagttgagttgagttgagtt
  E I L P R A V L D H S V Q
3061  atgtgtctgtaaaactgtaaaactgtaaaactgtaaaactgtaaaactgtaaaactgtaaaact
3121  ctaccaatgcaagttattgatttctctctctctctctctctctctctctctctctctctctct
3181  ttacgtattgtgttagGCACTGATTTTACGGCTATTTGGGAGGATTTGGAACATCA
  A C I F T G Y W E D V G T I K
3241  ATCATTCTTTGATGCAAACTGGCCCTCACTGAGGATctctctctctctctctctctctctct
  S F F D A N L A L T E Q
3301  tgcataatatttaccgtgaaactcaatgcaatgcaatgcaatgcaatgcaatgcaatgcaat
3361  gttttctctcaatgctctctctctctctctctctctctctctctctctctctctctctctct
3421  cctaattgtttctctctctctctctctctctctctctctctctctctctctctctctctct
  P S K F D F Y D P K T P F
3481  CTTCACTGCACCCGATGCTTCCGACCAATTTGGACAAAGTGAAGGCAAGTctctctctctct
  F T A P R C L P P T Q L D K C K
3541  actgagcacaattgttactgagcaagatttggctactgactgactgactgactgactgactgact
3601  TGAAATATGATTTATCTCAGATGTTGCTTCTAGAGAAATGCAACATCGAGCAATTTCTG
  M K Y A F I S D G C L L R E C N I E H S
3661  TGATGGAGTCTGCTCAGCTCAGCTCTGGATGGAECTCAAGgttacatactctctctctct
  V I G V C S R V S S G C E L K
3721  tgtatctactctgtgtagtatactctctctctctctctctctctctctctctctctctctct
3781  agcaacaagacctttagtctcaagcaatttaggttagcttagcttagcttagcttagcttagct
3841  acaaaatgcaagctctctctctctctctctctctctctctctctctctctctctctctctct
3901  tttttgggtataactcaactcaactcaactcaactcaactcaactcaactcaactcaactcaact
3961  cccatctactctctctctctctctctctctctctctctctctctctctctctctctctctct
4021  ggcatgttcaaacactcaactcaactcaactcaactcaactcaactcaactcaactcaactcaact
4081  ctcaactcaactcaactcaactcaactcaactcaactcaactcaactcaactcaactcaactcaact
4141  atgcaacatcagcatttagcaacatttagctgtaactgtaactgtaactgtaactgtaactgtaact
4201  acatattgcaacatacaactgtagctatcaactcaactcaactcaactcaactcaactcaact
4261  ctatgttagtctctctctctctctctctctctctctctctctctctctctctctctctctct
4321  tgctttgattctctctctctctctctctctctctctctctctctctctctctctctctctct
4381  taatattggaataacattctctctctctctctctctctctctctctctctctctctctctct
4441  gctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctct
4501  ctaagtcccggttctctctctctctctctctctctctctctctctctctctctctctctctct
4561  tctgtctctctctctctctctctctctctctctctctctctctctctctctctctctctctct
4621  ttgcaaggtgagcagttgtagctctctctctctctctctctctctctctctctctctctctct
4681  cggccgggacagtttctctctctctctctctctctctctctctctctctctctctctctctct
4741  gctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctct
4801  tcaactctctctctctctctctctctctctctctctctctctctctctctctctctctctctct
4861  tgtaaaaagcacaacaaatgtaactcctctctctctctctctctctctctctctctctctctct
4921  agaaaaaagtaaggtcaaaatgtaactctctctctctctctctctctctctctctctctctctct
4981  atctgtatctctctctctctctctctctctctctctctctctctctctctctctctctctctct
5041  tgagttcaactgtaactctctctctctctctctctctctctctctctctctctctctctctct
5101  cactctctctctctctctctctctctctctctctctctctctctctctctctctctctctctct
5161  ctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctct
5221  taagatcccacacaaaagagtttgaactcaagaaactctggaactctggaactctggaactct
5281  aatctctctctctctctctctctctctctctctctctctctctctctctctctctctctctctct
5341  tcaactctgtttagcaagtgaaactctctctctctctctctctctctctctctctctctctctct
5401  cccccccagggccagatataccagacatgaatacaagaatattgaaacccagatctaga
5461  gttgtgtgactgttgaatactggtgaaactctctctctctctctctctctctctctctctctct
5521  gacagGACTCCGTGATGATGGAGGGGACCACTATGAACTGAAGAAAGAGCTTCAAAGC
  D S V M M G A D T Y E E A S K
5581  TACTGTAGCTGGGAAGTCCCGTGGAAATGGAAGAACCAAGATAAGgttagatL
  L L L A G K V P V G I G R N T K I R
5641  ggtatggaaccaccgggttagttcccaaaatctactcactgatactgagttatctctctctct
  N C I I D M N A R I G K N V V
5701  ctgattattttcaagGAACGTATCATTGACATGAATGAGTGGGAAGAAAGCTGGTG
  N C I I D M N A R I G K N V V
5761  ATCACAAACAGTAAGgtgagcagccactcaactgaggtgagaaactctctctctctctctct
  I T N S K
5821  tctctaaatccgtaattctctctctctctctctctctctctctctctctctctctctctctct
5881  ctctgtgacagGGCTCCAAGAGGCTGATCACCCGGAAGAGGTACTACTAAGGCTGTG
  G I Q E A D H P E E G Y Y I R S
5941  GAATCTGGTGATCTGGAAGAATGCAACCATCAACGATGGGCTCTGATATAGATCGGCT
  G I V V I L K N A T I N D G S V I -
6001  GCGTGTGCTCTACAAAACAAGAACCTCAATGGTATGTCATCGATGGATCGTGAACCT
6061  TGATGGTAAGAGCCGCTTGACAGAAAGTGGAGCTCCGGCCAGAGTCGCTAGCTCTGGC
6121  ATGCTGCTCTGACCAATTTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT
  GCAGCAAACTTTTATGAACTTTGATTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT
6181  GTCATCTATATATTCAAAATTTTACGCTGTTTTCTAAATTCGGGTGCTGTTTTGGAT
6241

```

Figure 1. Nucleotide sequence and deduced amino acid sequence of a *Zea mays* wild-type *Sh2* allele. Base number 1 is the transcription start site. Introns are indicated by lowercase letters. Arrow indicates 3' end of cDNA. Putative TATA, RY dyad, and enhancer sequences are underlined.