In soybean [Glycine max (L.) Merr.], three distinct loci define the \( y_{20} \) Mdh1-n k2 cromosomal region (Chen and Palmer 1996, 1998). The \( y_{20} \) allele conditions the yellow foliage phenotype, the Mdh1-n allele conditions the mitochondrial malate dehydrogenase-1 (MDH1, EC 1.1.1.37) null phenotype, and the \( k_{2} \) allele conditions the tan saddle seed coat phenotype. The null mutants of Mdh1 lack two of the three electrophoretically identifiable mitochondrial malate dehydrogenase isozymes. This is due to the absence of the homodimeric MDH1 and the heterodimeric MDH1/MDH2 enzymes leaving only the active MDH2/MDH2 form (Hedges and Palmer 1992). Mdh1-n, \( y_{20} \), and \( k_{2} \) are inherited as recessive alleles (Chen and Palmer 1996, 1998).

Linkage studies were made possible by independently isolated mutants that showed one, two, or all three of the mutant traits. Table 1 gives the gene symbols and phenotypes of 12 mutants with appropriate near-isogenic wild-type siblings. The \( y_{20} \) mutant with dominant Mdh1 and/or dominant \( k_{2} \) has not been found naturally, by mutagenesis, or from recombinant events as a single mutation. No recombination was observed between the \( k_{2} \) and \( y_{20} \) loci in coupling phase from crosses of T253 (\( y_{20} \) Mdh1-n k2) with wild-type plants from 25,000 \( F_{2} \) and \( F_{3} \) plants (Palmer 1984). Additional recombination tests in coupling phase from \( F_{2} \) data (5400 plants) of T261 (Mdh1-n k2) crossed with wild-type plants gave a recombination value of \( 1 \pm 1.36 \) cm between the Mdh1-n and \( k_{2} \) loci (Chen and Palmer 1996). However, no recombinants were detected between the Mdh1-n and \( y_{20} \) loci in coupling phase with T253 (\( y_{20} \) Mdh1-n k2) and these recombinants (\( y_{20} \) Mdh1-n k2) from the T323, T324, and T325 (all \( y_{20} \) Mdh1-n) mutants crossed with wild-type plants (Chen and Palmer 1998). The recombinants involving the T323, T324, and T325 mutants were isolated in repulsion-phase experiments from cross-pollination with T239 (k2). These data in coupling phase indicate very close linkage of the \( y_{20} \) Mdh1-n k2 loci.

Data from repulsion phase gave recombination values that agree with the results from mutants in coupling phase and data that indicate independent assortment of Mdh1-n and \( k_{2} \). In repulsion phase, no recombinants with tan saddle seed coat and malate dehydrogenase 1 null were identified in 3210 \( F_{2} \) plants from crossing T239 (k2) or L67-3482 (k2) with PI 567.630A (Mdh1-n) (Chen and Palmer 1996). Similarly, no recombinants were identified in re-
Table 1. Description of soybean lines (wild types and mutants)

<table>
<thead>
<tr>
<th>Cultivar</th>
<th>Gene symbol</th>
<th>y20</th>
<th>Mdh1-n</th>
<th>k2</th>
<th>Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harosoy</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>k2</td>
<td>Wild type</td>
</tr>
<tr>
<td>T239</td>
<td>+</td>
<td>+</td>
<td>k2</td>
<td></td>
<td>Green foliage, malate dehydrogenase-I present, tan saddle seed coat</td>
</tr>
<tr>
<td>T253</td>
<td>y20</td>
<td>Mdh1-n</td>
<td>k2</td>
<td></td>
<td>Yellow foliage, malate dehydrogenase-I null, tan saddle seed coat</td>
</tr>
<tr>
<td>Clark</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>Wild type</td>
</tr>
<tr>
<td>Clark-k2 (1L7-3483)</td>
<td>+</td>
<td>+</td>
<td>k2</td>
<td></td>
<td>Green foliage, malate dehydrogenase-I present, tan saddle seed coat</td>
</tr>
<tr>
<td>Mandarin Ottawa</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>k2</td>
<td>Wild type</td>
</tr>
<tr>
<td>T261</td>
<td>+</td>
<td>Mdh1-n</td>
<td>k2</td>
<td></td>
<td>Green foliage, malate dehydrogenase-I null, tan saddle seed coat</td>
</tr>
<tr>
<td>T322</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>Purple and white variegated flowers (w4-m)</td>
</tr>
<tr>
<td>T323 (CD-1)</td>
<td>y20</td>
<td>Mdh1-n</td>
<td>+</td>
<td></td>
<td>Yellow foliage, malate dehydrogenase-I null, yellow seed coat</td>
</tr>
<tr>
<td>T324 (CD-2)</td>
<td>y20</td>
<td>Mdh1-n</td>
<td>+</td>
<td></td>
<td>Yellow foliage, malate dehydrogenase-I null, yellow seed coat</td>
</tr>
<tr>
<td>T325 (CD-3)</td>
<td>y20</td>
<td>Mdh1-n</td>
<td>+</td>
<td></td>
<td>Yellow foliage, malate dehydrogenase-I null, yellow seed coat</td>
</tr>
<tr>
<td>T326 (CD-9)</td>
<td>y20</td>
<td>Mdh1-n</td>
<td>+</td>
<td></td>
<td>Yellow foliage, malate dehydrogenase-I null, yellow seed coat</td>
</tr>
<tr>
<td>Jilin3 (PI 427.099)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>Wild type</td>
</tr>
<tr>
<td>T317</td>
<td>y20</td>
<td>Mdh1-n</td>
<td>+</td>
<td></td>
<td>Yellow foliage, malate dehydrogenase-I null, yellow seed coat</td>
</tr>
<tr>
<td>T361 (JB-yellow)</td>
<td>y20</td>
<td>Mdh1-n</td>
<td>+</td>
<td></td>
<td>Yellow foliage, malate dehydrogenase-I null, yellow seed coat</td>
</tr>
<tr>
<td>T234</td>
<td>y20</td>
<td>Mdh1-n</td>
<td>+</td>
<td></td>
<td>Yellow foliage, malate dehydrogenase-I null, yellow seed coat</td>
</tr>
<tr>
<td>Manchu</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td>Wild type</td>
</tr>
<tr>
<td>Mandelli</td>
<td>+</td>
<td>Mdh1-n</td>
<td>+</td>
<td></td>
<td>Green foliage, malate dehydrogenase-I null, variegated seed coat</td>
</tr>
<tr>
<td>PI 567.391</td>
<td>+</td>
<td>Mdh1-n</td>
<td>+</td>
<td></td>
<td>Green foliage, malate dehydrogenase-I null, variegated seed coat</td>
</tr>
<tr>
<td>PI 567.630A</td>
<td>+</td>
<td>Mdh1-n</td>
<td>+</td>
<td></td>
<td>Green foliage, malate dehydrogenase-I null, yellow seed coat</td>
</tr>
</tbody>
</table>

pulsed phase in 455 F₂ plants from crosses of T317 (y20 Mdh1-n) with T239 (k2) or L67-3483 (k2). A recombination value of 0.34 ± 0.48% was estimated between the Mdh1-n y20 and k2 loci (Chen and Palmer 1998).

In marked contrast to the results of mutants in repulsion phase from PI 567.630A and T317, F₂ progeny from repulsion phase revealed independent assortment between the k2 (T239 or L67-3483) and the y20 Mdh1-n k2 loci in T323, T324, or T325. Recombination values were 61%, 55%, and 56% for T323, T324, and T325, respectively (Chen and Palmer 1998). Chen and Palmer suggested that this is due to chromosome breakage in the presence of the k2 allele from T239 or L67-3483 at meiosis of the F₂ plants. T323, T324, and T325 were suggested to have deletions of the chromosome segment carrying Y20 and Mdh1. In the heterozygote with the k2 Mdh1 Y20 (wild-type) segment, Mdh1 Y20 is proposed to be deleted at a high frequency, thereby creating k2 gametes with a recombinant k2 Mdh1-n Y20 chromosome. The formation of this chromosome would explain the unlinked segregation data. Of interest, T317, which is allelic to mutants T323, T324, and T325, and thus is ascribed the same genotypic formula, does not give unlinked segregation when confronted in a heterozygote with the k2 allele of T239 or L67-3483. T317 and PI 567.630A must therefore have a different molecular constitution from T323, T324, and T325. These latter three mutants were identified in an experiment designed to generate transposable element-induced mutations by using T322, which harbors a putative transposon at the w4-m locus (Palmer et al. 1989). T317 was a spontaneous somaclonal mutant derived from tissue culture of PI 427.099 (Jilin province, China) (Amberger et al. 1992). PI 567.630A is a plant introduction from Henan province, China. In the genetic analyses, not a single case of Mdh1 recombination between Y20 and Mdh1 has been found. Whether the functionally independent mutants are the result of mutations in a single bifunctional gene or two separate genes has not been determined.

The objectives of these studies were to (1) determine whether or not the MDH1 phenotype arose from deletions in this region; (2) clone and sequence the nuclear-encoded Mdh1 null gene and its surrounding regions from a wild-type soybean line; and (3) determine, if possible, the apparent cause of the observed genetic instability of the y20 Mdh1 null k2 region.

Materials and Methods

Plant Material
Soybean mutants and parents are listed in Table 1. Cultivar T322 is a mutable line with purple and white variegated flowers suspected to contain an active transposable element (Grose et al. 1988). Three Mdh1 null and yellow foliage (y20) mutants, cultivars T323, T324, and T325, were isolated in a gene-tagging experiment using the w4-m mutable allele (Hedges and Palmer 1992; Palmer et al. 1989). T239 and T261 were two spontaneous mutants in the cultivars Harosoy and Mandarin Ottawa, respectively; cultivar T253 is a spontaneous mutant in the T239 background. L67-3483 was generated by X-ray radiation in the cultivar Clark (Rode and Bernard 1975). T234, PI 567.391, and PI 567.630A have no wild-type control.

Isolation of Genomic Soybean DNA
Ten grams fresh weight of leaves were frozen with liquid nitrogen and ground to a fine powder with a mortar and pestle. The frozen powder was transferred immediately to 50 ml of extraction buffer (100 mM Tris-HCL, pH 8.0, 1 mM EDTA, 2% SDS, 100 μg/ml proteinase K) and incubated with gentle agitation for 1 h at room temperature. The suspension was extracted with an equal volume of phenol/chloroform (1:1) and centrifuged at 400g for 15 min at 4°C in order to separate phases. The aqueous phase was extracted with chloroform and the DNA precipitated with ethanol. After centrifugation, the DNA was dissolved in 5 ml of buffer [10 mM Tris-HCL, pH 8.0, 1 mM EDTA, and RNase (10 μg/ml)] and treated for 1 h at 37°C. The DNA was extracted twice with phenol/chloroform and once with chloroform. Finally, the DNA was precipitated with ethanol and dissolved in 10 mM Tris-HCL, pH 8.0, for restriction digestion and Southern blot analysis.

Isolation and Characterization of Soybean MDH cDNA and Genomic Clones
A cDNA library prepared from 10-day-old soybean seedlings (cultivar Resnik), constructed in λgt10, was obtained from Clontech (Palo Alto, CA) and screened with a watermelon mMDH cDNA probe (Gietl et al. 1990). The cDNA inserts of the 10 positive clones were labeled with DigoxigenindUTP (Boehringer, Germany) by amplification with PCR (Saiki et al. 1985) and λgt10-specific primers (Clontech) and used as probes in Southern blot analysis of genomic soybean DNA preparations digested with EcoRI, BamHI, or HindIII. Segments of two soybean cDNA inserts, A2 (AF068689) and D1 (AF068688), were sequenced by the dideoxy chain termination method using T7 DNA polymerase (Sequenase 2.0, U.S. Biochemical Corp.).

A genomic library of 9–23 kb Sau3A partial restriction fragments of soybean DNA (cultivar Williams 82), constructed in the
Genomic DNA of the wild-type soybean line T322 (Table 1) was double-digested with EcoRI and XbaI and cloned into the EcoRI site of the λZAPII vector (Stratagene). Preliminary experiments had shown that the mMDH sequences should not contain an internal XbaI site. The library was screened by hybridization to the watermelon mitochondrial malate dehydrogenase cDNA clone (Gietl et al. 1990). Two positive clones were isolated: clone c1 with an insert size of 6050 bp, and clone c2 with an insert size of 7500 bp. These λZAPII phage clones were converted into the plasmid pBluescript SK (−) by in vivo excision according to the manufacturer’s instructions and partially restriction mapped. Clone 1 had a 5700 bp EcoRI-XbaI fragment coding for part of the Mdh1 gene that is not deleted in the mutants (AF068687) and a flanking 350 bp XbaI-EcoRI fragment. Clone 2 contained an internal 4.8 kb XbaI fragment coding for the C-terminal segment of geranylgeranyl hydrogenase and the complete coding sequence of the nodule-enhanced malate dehydrogenase (AF068686). This 4.8 kb XbaI fragment was flanked by a 2.5 kb EcoRI-XbaI fragment and a 0.45 kb EcoRI-XbaI fragment. Due to our cloning strategy, we assumed that these flanking EcoRI-XbaI fragments belonged to different parts of the genome and became linked to the internal XbaI site as cloning artifacts. Nucleotide sequences were determined by PCR walks by the Iowa State University DNA Sequencing Facility. Southern blot analysis of genomic DNA and phage DNA was carried out according to Sambrook et al. (1989).

Results

Mdh1-n Mutations Correlate with the Deletion of Specific Genomic Restriction Fragments that Encode a Mitochondrial Malate Dehydrogenase

Genomic DNAs isolated from the malate dehydrogenase null mutants (Mdh1-n) T323, T324, T239, and T317 were digested with EcoRI and probed with a cDNA encoding the watermelon mitochondrial malate dehydrogenase (Gietl et al. 1990). The mutant DNA preparations consistently lacked a 5.5 kb EcoRI fragment that was present in all corresponding wild-type parental lines (Figure 1). Subsequently genomic DNAs from wild-type T322 and mutant T323 were digested with HindIII or BamHI. A comparison of the HindIII digests showed that the mutant lacked a specific 6.0 kb fragment, whereas in BamHI digests the mutant lacked a specific 3.0 kb fragment (data not shown).

To isolate the cDNA fragment encoding the missing mMDH1 subunit, a commercial Agt10 cDNA library (Clontech) was screened with the watermelon mMDH cDNA. Ten cDNA clones were obtained, two of which (A2 and D1) recognized on a genomic Southern blot the expected restriction fragments (i.e., EcoRI, 5.5 kb; HindIII, 6.0 kb; BamHI, 3.0 kb) that were missing in the mutant T323. When EcoRI digests of the other soybean Mdh1-n mutant genomic DNAs (see Table 1) were probed with the A2 clone, the 5.5 kb fragment was not detected (autoradiogram not shown). The A2 clone was sequenced (AF068689) and revealed an open reading frame that encoded a putative peptide of 65 amino acids. This peptide showed 85% identity at the amino acid level to a domain of the watermelon mMDH (gil126896). The soybean cDNA D1 clone was sequenced (AF068688) and revealed an open reading frame that encoded a peptide of 87 amino acids with 93% identity to the internal peptide of the watermelon mMDH containing part of the characteristic NADH-binding site and amino acids belonging to the catalytic pocket. When compared with the coding regions of watermelon (Gietl 1990) or soybean glyoxysomal MDH (Guez et al. 1995), identity at the amino acid level for A2 and D1 was only 66% and 54%, respectively. Thus it was concluded that the insertions of clones A2 and D1 were parts of the mitochondrial Mdh1 gene of soybean that was deleted in the Mdh1-n mutants.

Isolation of Genomic Clones with the Complete Mdh1 Gene

Using the soybean cDNA clone A2 as a probe, 24 positive lambda clones were isolated from the Stratagene FIX II soybean library. Three genomic clones (11, 32, and 33) contained the 5.5 kb EcoRI fragment and the 3.0 kb BamHI fragment expected for a genomic clone containing the wild-type Mdh1 gene (data not shown). The A2 probe also highlighted the expected 6 kb HindIII band in clone 33. On the other hand, a 23 kb HindIII fragment hybridized with the probe when tested on the insert of clone 11, whereas a 15 kb HindIII fragment hybridized with the probe when ap-
plied to the insert of clone 32. The latter two fragments resulted from the absence of HindIII recognition sites between the insert and the respective arms of the vector. Inserts of the other 21 lambda clones did not produce the restriction fragment patterns expected for the gene deleted in the mutants.

Wild-type (T322) and mutant (T323) DNAs were digested with BamHI or EcoRI and probed with the labeled 3.0 kb BamHI fragment of lambda clone 33, which contains several exons of the mMDH1 gene that is deleted in the mutants, was used as a probe.

Sequence Analysis of Clones 11, 33, and 32
The T3 end of clone 11 (i.e., bp number 1) extends upstream 4636 bp from the T7 end of clone 33. Clone 33 is comprised of 11,970 bp and contains the entire Mdhl gene. The T7 end of clone 32 is distal to the T3 end of clone 33 and extends downstream 10,982 bp (Figure 3). Thus the three clones span a unique sequence of 27,588 bp (AF180335).

Clone 11 contains an inverted repeat of 164 bp (65% identity) between bp 981 and 1135. Also, clone 11 contains a 235 bp region (4113–4345 bp) with 69% identity to an Arabidopsis gene encoding an AMP-binding protein (CA10181), and a 138 bp sequence shows 73% identity to a Brassica napus gene encoding an AMP-binding protein (Z72151). A 200 bp direct repeat with 81% identity extends from bp 5039 to 5253 and reoccurs at bp 5865–6065.

Clone 33 contains the entire Mdh1 gene. The AUG start codon appears at 7254 bp and the UAA termination codon at 10,702–04 bp. The coding region contains six introns (see AF180335). At 11,875–12,241 bp, a 368 bp region occurs that is duplicated at 15,742–16,109 bp. This duplicated sequence occurs near the end of clone 33 and is similar to a barley retrotransposon (HVU76261).

Clone 32 possesses several unique properties. Immediately past the T3 end of clone 33 (i.e., at 16,851–17,206 bp) a 372 bp region occurs with 65.6% identity to the antisense of soybean catalase (Z12021). A second region with similarity to soybean catalase exists at 17,865–18,156 bp. Two copies of a rare 8 bp palindromic repeat (i.e., TGGTAATCGATTAC) were found at 18,120–18,135 bp and at 18,156 bp. Two copies of a rare 8 bp palindromic repeat (i.e., TGGTAATCGATTAC) were found at 18,120–18,135 bp and at 18,156 bp. A truncated repeat occurs at 17,920–17,931 bp. At the T7 end of clone 32 (i.e., 26,349–27,579 bp) occurs a 1233 bp region with 59% identity to the Cyclops retrotransposon (AJ000640; Chavanne et al. 1998). A reading frame that encodes a putative 238 amino acid peptide containing one termination codon resides within this 1233 bp region (Table 2). Over a 191 amino acid residue segment, a 40% identity was found between the putative soybean peptide and the putative Cyclops gag-pol protein (AB007467).

Genomic Clones that Contain a Partial mMDH2 Gene and a Nodule-Enhanced Malate Dehydrogenase
Using the watermelon mMDH cDNA as a probe, two genomic clones, designated c1 (i.e., MHD2) and c2 (i.e., mMDH), were isolated from an EcoRI-XbaI-digested DNA library of wild-type soybean T322. Subsequently a 1607 bp HindIII/PstI fragment was isolated from the 5770 bp c1 clone. This fragment was labeled and used as a probe in the Southern blot analysis of both wild-type (T322) and mutant (T323) DNAs. A 7.0 kb EcoRI fragment (data not shown) and a 9.0 kb HindIII fragment (data not shown) were detected in both DNAs. Correspondingly a 1300 bp EcoRI fragment (data not shown) and 20 kb EcoRI fragments from both wild-type and mutant DNA (data not shown). This clearly demonstrated that these two MDH genes, which hybridized with the watermelon cDNA coding for mMDH, were present in the mutants and hence might represent two MDH genes with different primary structures.

The restriction site maps of the 5770 bp mMDH1 region deleted in the mutants.

**Table 2.** Composition of the putative Cyclops-like gag-pol peptide encoded by the complementary strand of AF180335 (26,350–27,063 nucleotides) downstream from the Mdh1 gene

<table>
<thead>
<tr>
<th>Residue</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SRSQAREKK</td>
</tr>
<tr>
<td>51</td>
<td>MPFIDALQGM</td>
</tr>
<tr>
<td>101</td>
<td>HKDPPSWVTIP</td>
</tr>
<tr>
<td>151</td>
<td>TLQIADCSFT</td>
</tr>
<tr>
<td>201</td>
<td>FILTASCVID</td>
</tr>
</tbody>
</table>

* The protease active site is underscored.
* Designates a termination codon.

---

**Figure 1.** Southern blot analysis of soybean genomic DNA (wt, wild-type T322; mu, mutant T323; see Table 1) digested with EcoRI or BamHI. The 3.0 kb BamHI fragment of lambda clone 33, which contains several exons of the mMDH1 gene that is deleted in the mutants, was used as a probe.

**Figure 2.** Southern blot analysis of soybean genomic DNA (wt, wild-type T322; mu, mutant T323; see Table 1) digested with EcoRI or BamHI. The 3.0 kb BamHI fragment of lambda clone 33, which contains several exons of the mMDH1 gene that is deleted in the mutants, was used as a probe.

**Figure 3.** Restriction site maps of lambda clones 11, 32, and 33, which contain the mMDH1 region deleted in the mutants.
phenotypically, T323 is the most yellow, T324 is a moderate yellow, and T325 is a greenish-yellow. We hypothesize that at least one of the break points occurring among these mutants. To date, we have isolated 29 spontaneous Mdh1-n mutations (Chen et al. 1999; Chen and Palmer 1996, 1998), suggesting that the Mdh1 chromosomal region may contain an instability factor. We examined the 27,588 bp region that encodes and surrounds the Mdh1 gene for possible causes of this putative instability. The AUG start codon for the MDH1 protein is found at 7254 bp and its termination codon at 10,704 bp. The six introns within the Mdh1 gene did not reveal an apparent cause of instability. Upstream from the Mdh1 coding region, a 150 bp repeat region was found at 981–1135 bp. Downstream from the termination codon several possible instability factors were found. Sequences similar to a barley retrotransposon were found at 11,875–12,241 bp and at 15,742–16,109 bp. Within this latter region, two copies of a rare 8 bp palindromic repeat (TGTTAATCGGATTACCA) were found. Finally, near the distal end of the sequenced region (i.e., 26,340–27,538 bp), a 1233 bp DNA segment was found that showed a 59% identity at the DNA level to a segment of the Cyclops retrotransposon (AJ000640) from Pisum sativum (Chavanne et al. 1998). The 1233 bp segment contained a reading frame that may encode a peptide similar to a segment of the large gag-pol protein thought to be produced by Cyclops (Chavanne et al. 1998). Translation of the complementary strand from 26,350 to 27,063 nucleotides produced a putative peptide containing 238 amino acid residues, one of which (number 80) was a termination codon (Table 2). At the amino acid level, a 40% identity and 68% similarity was found over a 191 amino acid residue region of the 238 residue polypeptide and a putative Cyclops protein from Vicia faba (AB007467). Likewise, a 35% identity and 57% similarity was found over a 225 amino acid residue region of the putative protein shown in Table 2 and a putative protein from Arabidopsis (AC006250). The short putative 238 residue peptide contained the active site (i.e., LVDLGAS) of the protease portion of the gag-pol protein (Chavanne et al. 1998). The Pisum sativum genome is reported to contain approximately 5000 copies of Cyclops, whereas the Glycine max genome contains perhaps 1000 copies (Chavanne et al. 1998). With this extent of genetic duplication, these retrotransposons could promote genetic rearrangements deleting the Mdh1 region as documented in maize (Zhang and Peterson 1999).

References
Miller SS, Driscoll BT, Gregerson RG, Gantt JS, and Vance CP, 1998. Alfalfa malate dehydrogenase (MDH):


Received June 6, 2000
Accepted February 14, 2001
Corresponding Editor: Irwin Goldman